LAPPEENRANTA-LAHTI UNIVERSITY OF TECHNOLOGY LUT
School of Business and Management
Master's Programme in International Marketing Management (MIMM)
Master's thesis

Arla Behm

THE ANTECEDENTS OF BIG DATA ANALYTICS: INTEGRATING RESOURCE-BASED THEORY AND KNOWLEDGE MANAGEMENT PERSPECTIVE

17th of October 2019

1st Supervisor: Associate Professor Anssi Tarkiainen

2nd Supervisor: Assistant Professor Joel Mero

ABSTRACT

Author's name:	Behm, Arla
Title of the thesis:	The antecedents of big data analytics: Integrating resource-based theory and knowledge management perspective
Faculty:	LUT School of Business and Management
Degree Programme:	Master's Programme in International Marketing Management (MIMM)
Year:	2019
Master's thesis university:	LUT University. 66 pages, 4 figures, 1 table and 1 appendix
Examiners:	Associate Professor Anssi Tarkiainen & Assistant Professor Joel Mero
Keywords:	Big data, Big data analytics, Knowledge management, Organisational resources, Resource-based theory

The purpose of this research is to identify the antecedents of big data analytics and to study the ways in which efficient resource and knowledge management can help companies to capitalise on big data analytics. Therefore, this research aims to illustrate possible unique insights regarding organisational resources and knowledge management processes that have a major role in supporting big data analytics practices. To achieve this objective, the study applies qualitative research method in the form of a case study.

The findings of this research conclude a comprehensive overview of the antecedents of big data analytics. The findings emphasise the importance of open communication in the organisations as well as other individual skills, such as curiosity and proactiveness, of the employees. Additionally, the findings depict the importance of establishing an environment and an organisational culture that supports and enhances successful big data analytics. By identifying and discovering the relevant matters from organisational resources and knowledge management perspective, the findings provided by this research can be utilised for establishing sustainable and successful big data analytics practices in organisations.

TIIVISTELMÄ

Tekijä:	Behm, Arla
Tutkielman nimi:	The antecedents of big data analytics: Integrating resource-based theory and knowledge management perspective
Tiedekunta:	Kauppatieteellinen tiedekunta
Maisteriohjelma:	Master's Programme in International Marketing Management (MIMM)
Vuosi:	2019
Pro gradu -tutkielma:	LUT-Yliopisto. 66 sivua, 4 kuvaa, 1 taulukko ja 1 liite
Tarkastajat:	Apulaisprofessori Anssi Tarkiainen & Apulaisprofessori Joel Mero
Hakusanat:	big data, big data analytiikka, tietämyksenhallinta, yrityksen resurssien hallinta, resurssiperusteinen teoria

Tämän tutkimuksen tarkoituksena on tarkastella big data -analytiikkaa edesauttavia tekijöitä, joissa yritysten resurssien hallinnalla sekä tietämyksenhallintaprosesseilla on merkittävä vaikutus. Tutkimuksen tavoitteena on havainnollistaa ainutlaatuisia ominaisuuksia yrityksen resursseista sekä tietämyksenhallinnasta, jotka mahdollistavat ja tukevat big data -analytiikan hyödyntämistä yrityksen liiketoiminnassa. Työ on toteutettu kvalitatiivisena tutkimuksena, jossa lähestymistapana on hyödynnetty tapaustutkimusta.

Tutkimuksen löydökset korostavat avoimen vuorovaikutuksen ja keskustelun merkitystä yrityksissä sekä muiden työntekijöiden yksilöllisten ominaisuuksien ja taitojen, kuten uteliaisuuden sekä proaktiivisuuden, tärkeyttä. Lisäksi tutkimus osoittaa, että kannustavan, osallistavan ja avoimen työilmapiirin ja -kulttuurin vakiinnuttamisella on tärkeä rooli big data -analytiikkaa edesauttavana tekijänä.

Tunnistamalla ja tutkimalla merkittäviä tekijöitä sekä yrityksen resurssien että tietämyksenhallintamenetelmien näkökulmasta, tämän työn tarjoamia löydöksiä ja tuloksia voidaan hyödyntää onnistuneen ja kestävän big data -analytiikan vakiinnuttamiseksi yrityksissä.

ACKNOWLEDGEMENTS

I would like to thank all of those who have helped me during this thesis project. Firstly, I would like to thank my supervisor, Assistant Professor Joel Mero, for the invaluable guidance and support he provided throughout this thesis project.

I also want to thank all the interviewees for their effort and taking the time to participate in this project. Additionally, the excitement and interest the interviewees expressed on the topic was highly appreciated and motivated me during the project.

Finally, I want to thank my family for encouraging me and my best friend for believing in me and for all the support I received during the entire thesis project.

Table of contents

1 INTRODUCTION

The amount of generated data in the entire world is growing exponentially. During the last couple of years, 90 percent of all the data in the world was generated (Marr 2018). As stated by Gandomi & Haider (2015), the role of technology as an enabler of generating data is undeniable and the technology today has led us to a point where data is described as *big data*. Furthermore, as technology is constantly developing and advancing, the amount of data can be expected to exceed to even greater dimensions than of today. Organisations across industries are already constantly confronted with the enormous amount of complex data flowing in from various sources. Utilising data to gain business intelligence is a common practice for business operators alike, yet many organisations are already struggling with handling bigger amounts of data (Intezari & Gressel 2017). The rapid development of data requires organisations to adapt to the changes and adjust their data analytics tools, processes and resources accordingly. Only then, are organisations able to meet the challenges data will present in the future.

Numerous studies present the marvellous opportunities and notable changes big data and big data analytics provide to the world of business as we know it. Balducci & Marinova (2018) claims big data will reshape business practises in different industries notably, whereas Merendino, Dibb, Meadows, Quinn, Wilson, Simkin & Canhoto (2018) state, how big data would reduce risks in decision-making processes and improve strategic decision-making by allowing the management level a more holistic view. Furthermore, according to Côrte-Real, Ruivo, Oliveira & Popovič (2019) big data analytics is a notable differentiator between high-performing and low-performing organisations. Therefore, big data analytics and the positive impacts, numerous opportunities and possibilities it enables, are and have been a subject of interest for both academics and corporate leaders – and thus a relevant topic for research.

1.1 Aim of the study

The aim of this study is to identify the antecedents of big data analytics where efficient resource management and knowledge management processes have a notable role. Under inspection are the ways in which organisational resources as well as knowledge management processes support the exploitation of big data analytics. This research provides a conceptual framework that can be utilised by practitioners to help in identifying the antecedents of big data analytics by integrating both organisational resources and knowledge management processes into one framework. By studying the matter in an alternative perspective, this research can provide integrative insights that may have been unnoticed in previous studies.

1.2 Research problem and research questions

The research problem is that given the numerous prosperous opportunities big data provides, it is still not exploited by companies successfully due to different reasons. According to the literature, only a few organisations analyse the data available or obtain any benefit of big data analytics. The reasons can be due to lack of internal competencies to conduct analysis processes, lack of necessary knowledge on big data, due to lack of necessary resources or due to lack of necessary cooperation within the company (Beach & Schiefelbein 2014; Gandomi & Haider 2015; Côrte-Real et al. 2019; Berinato 2019). These reasons can be deemed interrelated in corporate environments. Additionally, big data and resource-based theory as research subjects have been studied comprehensively as well as knowledge management. Nevertheless, integrative studies that combine resource-based theory, big data analytics and knowledge management are scarce, which generates a gap in the literature regarding the topic. Utilising data to gain business intelligence is not an innovative practise as companies have been gathering and capitalising on data for years (Intezari & Gressel 2017). However, the data generated today is entirely different regarding volume, velocity and variety and thus it requires advanced and innovative processes, resources, tools and technologies to be used (Gandomi & Haider 2015). Therefore, in order to conduct effective data analytic practices with the data of today and in the future, the antecedents of the practises need to be studied comprehensively. The antecedents, in this case, being company’s resources and knowledge management processes.

To gain a comprehensive understanding of the antecedents of big data analytics the following research questions are made. Firstly, the specific resources that are required to conduct big data analytics are identified by answering the first research question. Secondly, to gain an understanding of the processes supporting big data analytics, another research question is established. Hence, the research has two research questions that are equally important for this study. Both of the research questions are presented below.

What resources does the exploitation of big data analytics require?

What knowledge management processes support the exploitation of big data analytics?

After having understood the ways in which organisational resources and knowledge management processes support big data analytics practices, relevant and thorough insights about successful big data analytics can be drawn.

1.3 Theoretical framework and key concepts

The theoretical framework of this study is based on resource-based theory and on knowledge management. In this paper, knowledge management is studied as the core process which is enabled by the company's resources. The role of knowledge management in big data analytics is significant as the information, that can be used for company's benefit, is in fact knowledge that is extracted from the data. Although these entities of resource-based theory and knowledge management are discussed separately, they in fact are related as knowledge can be deemed as a company's resource. Additionally, they both are studied as antecedents that enable the execution of successful big data analytics. Nevertheless, the insights on the role of knowledge management in big data analytics will be discussed as a separate, yet related entity to resource-based theory. The theoretical framework of the study is presented visually in Figure 1.

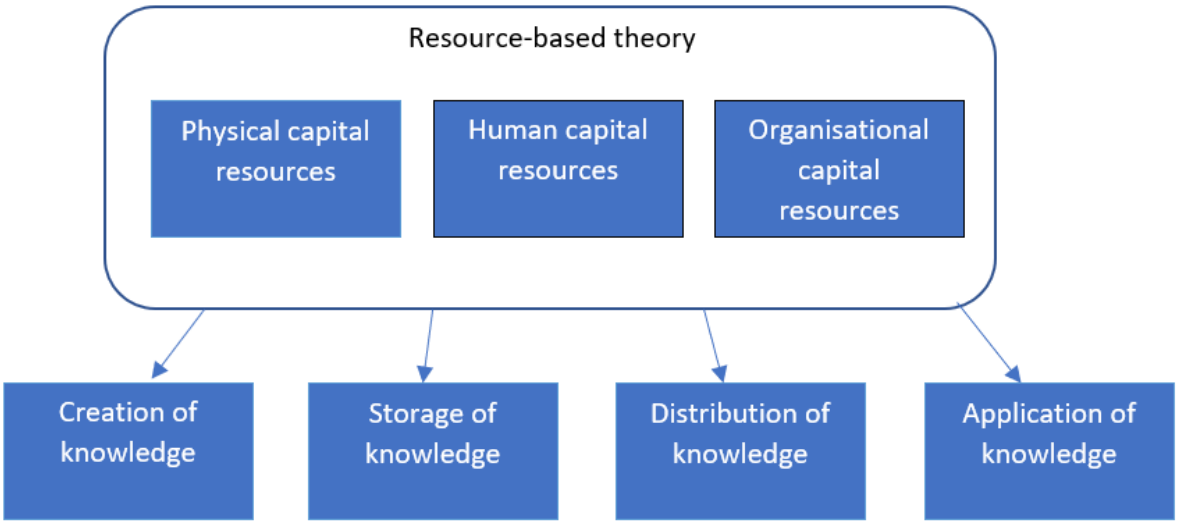

Figure 1. Theoretical framework of the study.

Combining resource-based theory with knowledge management will help in identifying the possible factors that prevent or are of great importance for companies to successfully capitalise on big data. Therefore, the theoretical framework and the findings of the study will provide a comprehensive outlook on the antecedents and processes that are likely to enable companies to conduct and to successfully exploit big data analytics.

Key concepts of this study are organisation’s resources, knowledge management and big data analytics. Resources are defined as according to Barney (1991), the physical capital resources, human capital resources, and organisational capital resources. Additionally, the concept consists of resources that are both tangible and intangible in nature. Worth noting is that knowledge is also deemed as a company’s intangible resource that possesses a critical strategical purpose (Grant 1996b). Therefore, the analysis of knowledge and knowledge management are relevant when studying company’s resources comprehensively. Knowledge management is defined as a systematic process that consists of practises like creating, sharing and implementing knowledge (Intezari & Gressel 2017).

The term big data in itself is defined as a massive amount of complex and ever-altering data, flowing in from multiple sources that exceeds the analytic capabilities of traditional technologies and systems (Rajaraman 2016). In this study, successful big data analytics refers to the orchestration of organisational resources and knowledge management processes as an efficient and an agile entity that generates sustainable competitive advantage (Gupta & George 2016). Therefore, the focus of the study is to present the antecedents for successful big data analytics that will enable the company to discover, create and organise valuable knowledge from the data which will provide an opportunity to enhance the organisation’s competitiveness.

1.4 Research methodology

This study is conducted as a qualitative research as the aim is to describe phenomena or to understand certain empirical activity rather than establish statistical statements (Eskola & Suoranta 1998, 13-14). The qualitative research method selected for this study is a case-study, as it is best used to thoroughly describe a phenomenon within its real-world context and to understand the related contextual conditions (Yin 2014, 16-17). The empirical material was collected by interviewing employees of a case company where data is actively gathered, managed and utilised. Furthermore, additional secondary data was collected in the form of public documents produced by the case company.

1.5 Structure and delimitations of the study

This study will not analyse resource-based theory and knowledge management as separate entities rather as a combination where both entities are related and support each other. Hence, to provide insights and deeper understanding by combining these two disciplines and studying them in parallel with big data analytics. This study will not concentrate on the precise big data analytics methods, tools or systems, rather focus comprehensively on the antecedents of big data analytics and generally on the ways in which efficient resource and knowledge management can help companies to capitalise on big data analytics and to gain a competitive edge.

The study is structured as follows: first the introduction section where the aim of this study, research problem and questions, theoretical framework and key concepts, research methodology and limitations of the study are presented. Then in the following chapters the theoretical literature is reviewed in a logical order. Starting with analysing the key concepts of the study in precision, following by a review of the constructs of the theoretical framework. After the theoretical literature review, the methodology of the research is presented to clarify the research process. In the following chapter the findings of this study are presented by combining the results of the empirical data and the theoretical framework of this study. Lastly, in the discussion and conclusion chapter, the significance of this research, the research data and findings are discussed, and the conclusions of this research are presented.

2 BIG DATA ANALYTICS

As its name depicts, big data refers to an enormous amount of extremely complex and constantly altering data that is generated through a myriad of sources. A comprehensive definition was necessary in understanding the complex nature of the data and therefore, in 2001 Laney established three dimensions or the *three V's* of big data to help in defining and understanding the concept. Although, De Mauro, Greco & Grimaldi (2014) specify in their article that these dimensions are used to characterise the information involved in big data analytics. What De Mauro et al. (2014) also describe as features associated with big data are the specific technology and analytic methods that big data requires. Chen, Chiang & Storey (2012) also agree, that big data poses advanced and unique requirements for storage, management analysis and visualisation technologies. Additionally, the principal way big data analytics affect organisations and society by providing insights that result in creation of economic value is also deemed as characteristics of big data analytics. The three dimensions have been and still are the most commonly used framework when attempting to encompass and define big data (Davenport, Barth & Bean 2012; Erevelles, Fukawa & Swayne 2016; Gandomi & Haider 2015). Nevertheless, by combining these definitions a summary that would attempt to comprehensively define big data can be constructed. As Rajaraman (2016) summarises, big data can be defined as data that is massive in its volume, tends to vary and alter constantly, requires analysis executed by novel and advanced tools for obtaining insights from its implications as well as requires enormous computing resources.

The three V's used to describe big data are volume, velocity and variety (Laney 2001). Volume referring to the size of data. As stated by Gandomi & Haider (2015), the concept of big data volumes is relative and prone to vary over time and type of data. What was deemed big data in the beginning of 2000 is completely different from the current perception of size. Therefore, the current perception of big data is not likely to meet the threshold in the future. Davenport et al. (2012) agree that the current database and analytics technologies as well as storage capacities are constantly increasing and improving, thus allowing even bigger data sets to be captured. Furthermore, as the type of data affects notably to the data's perception of size, it is illogical as well as impractical to define any strict or specific thresholds for big data volumes (Gandomi & Haider 2015).

Variety refers to the structural heterogeneity, the complex nature and diverse richness of a dataset that the multiple sources of big data provide (Erevelles et al. 2016; Gandomi & Haider

2015). Variety as a characteristic of data is not novel per se – already in the 1960's variety in data sets existed as the predominant types of data were numerical and textual (Rajaraman 2016). Organisations have been collecting different types of data from internal and external sources for business intelligence activities during several years. However, the shift of big data has been from collecting and from being exposed to mainly structural data to also unstructured data. As presented by Erevelles et al. (2016), the technological advances enable average consumers to generate not only traditional, structured data but also more contemporary, complex and unstructured behavioural data. Not only do the technological advances enable consumers to generate more data but also allow organisations to leverage these various types of structured, semi-structured, and unstructured data in business processes (Gandomi & Haider 2015).

Velocity as a character describes the rate at which data are generated and the speed at which it should be analysed and acted upon. As stated by Hilbert (2016), one good example of a source that generates enormous volumes of complex data at a rapid pace is social networks. Furthermore, as mobile devices are becoming increasingly ubiquitous, they also generate a variety of data constantly and thus have become a universal data source. This expansion of portable digital devices has caused a remarkable increase of real-time data creation that naturally require rapid analytics and evidence-based planning. The fast-paced nature of data generation shortens its life-span – the generated data can become irrelevant quickly (Rajaraman 2016). This kind of high-frequency data generated from a myriad of streaming data sources is constantly growing and thus the demand for advanced analytic techniques increases. Naturally, the traditional data management systems are not built for handling these kinds of massive data feeds. Therefore, the need for advanced big data technologies that enable organisations to constantly harvest relevant information from high volumes of data and act upon the insights is imminent (Erevelles et al. 2016; Gandomi & Haider 2015).

Over the years, couple of additional dimensions have been added to the framework to support in the analytics processes regarding big data. These additional dimensions, as presented by Gandomi & Haider (2015) are veracity, variability and value. Veracity refers to the variety of credibility and reliability of different data sources. Therefore, one must be aware of data quality since the possibility of gaining inaccurate and even irrelevant data is likely when handling big data. Furthermore, as stated by Abbasi, Sarker & Chiang (2016), the occurrence of spam and false information in social media channels is imminent which affects data quality. Variability represents the variation in the data flow rates and the multiple sources through which big data is produced (Gandomi & Haider 2015). This character makes big data very complex in nature. Lastly, value refers to the level of value regarding the volumes of the data analysed (Erevelles et al. 2016). The key success factor is to promptly eliminate irrelevant data and focus on the remaining, useful data that will help in creating relevant insights and work as a basis for decision-making. Which is, due to the immense dimensions of big data, one of the biggest challenges of big data analytics.

In an article by Côrte-Real, Oliveira & Ruivo (2017), big data analytics is defined as technologies and architectures that are superior to previous generations and that are designed to efficiently extract value from enormous volumes of complex data, by enabling high velocity discovery, capturing and analysis. Whereas, Fosso, Akter, Edwards, Chopin & Gnanzou (2015) define big data analytics as a holistic approach to managing, analysing and processing big data in respect to its attributes and dimensions. The aim of big data analytics is to generate ideas that can be implemented for delivering sustained value, measuring performance and naturally for establishing competitive advantages. Nevertheless, as Gandomi & Haider (2015) claim, big data is valuable only when used to support decision-making. Therefore, efficient processes to convert high volumes of rapid and complex data into relevant and meaningful insights is the basis for facilitating better informed and evidence-based decision-making and thus, basis for successful exploitation of big data analytics that provide sustainable competitive advantages (Mikalef, Boura, Lekakos & Krogstie 2019a).

Labrinidis & Jagadish (2012) have divided extracting of insights from big data analytics into five different steps that are acquisition and recording; extraction, cleaning and annotation; integration, aggregation and representation. These steps define the activities that can be deemed as data management activities. The last two steps are modelling and analysis, and interpretation, those steps can be deemed as analytics activities. Additionally, Gandomi & Haider (2015) claim data management activities to consist of processes and supporting technologies that are used to acquire and store data. Furthermore, they are used to prepare and retrieve the data for analysis. Analytics activities refer to the techniques that are used in analysing and acquiring relevant intelligence from big data. Therefore, big data analytics can also be deemed as a sub-process of the general process of extracting insights from big data.

2.1 Theoretical framework

The theoretical framework of this study consists of two main disciplinaries: resource-based theory and knowledge management which are tied together with big data analytics perspective to form a coherent framework. Resource-based theory is used to identify the antecedents of big data analytics. As stated by Vidgen, Shaw & Grant (2017), success in analytic activities depends on the organisation's ability to continuously and simultaneously manage organisational resources alongside with data, and to deploy these to generate a competitive advantage that is sustainable and valuable. Furthermore, resource-based theory is an efficient tool in describing the relationship between organisational resources and performance (Gupta & George 2016). Therefore, organisational resources and their relationship with big data analytics are analysed in this study. The purpose of introducing resource-based theory is to analyse the role of resources regarding big data analytics. Additionally, the core-process of conducting big data analytics is analysed in knowledge management's perspective. Therefore, the organisational resources and knowledge management processes are studied as antecedents of big data analytics. The reason why knowledge management is studied as a central process of big data analytics is, as stated by Pauleen & Wang (2017), because human knowledge has been the main developer of the capabilities of big data analytics. Moreover, the authors claim that human knowledge has a major role in deciding the ways in which the information generated from big data analytics is used.

An illustration of the resources and knowledge management processes of big data analytics identified from the literature is presented in Figure 2. The connections between the constructs depict the ways how the two disciplinaries interlace in regards of big data analytics as presented in the literature review. Additionally, the most notable matters that emerge from the literature are presented in each construct. Since prior studies that combine the two disciplinaries alongside with big data are scarce, the illustrations are based on notable matters emerging from the literature regarding these two disciplinaries separately.

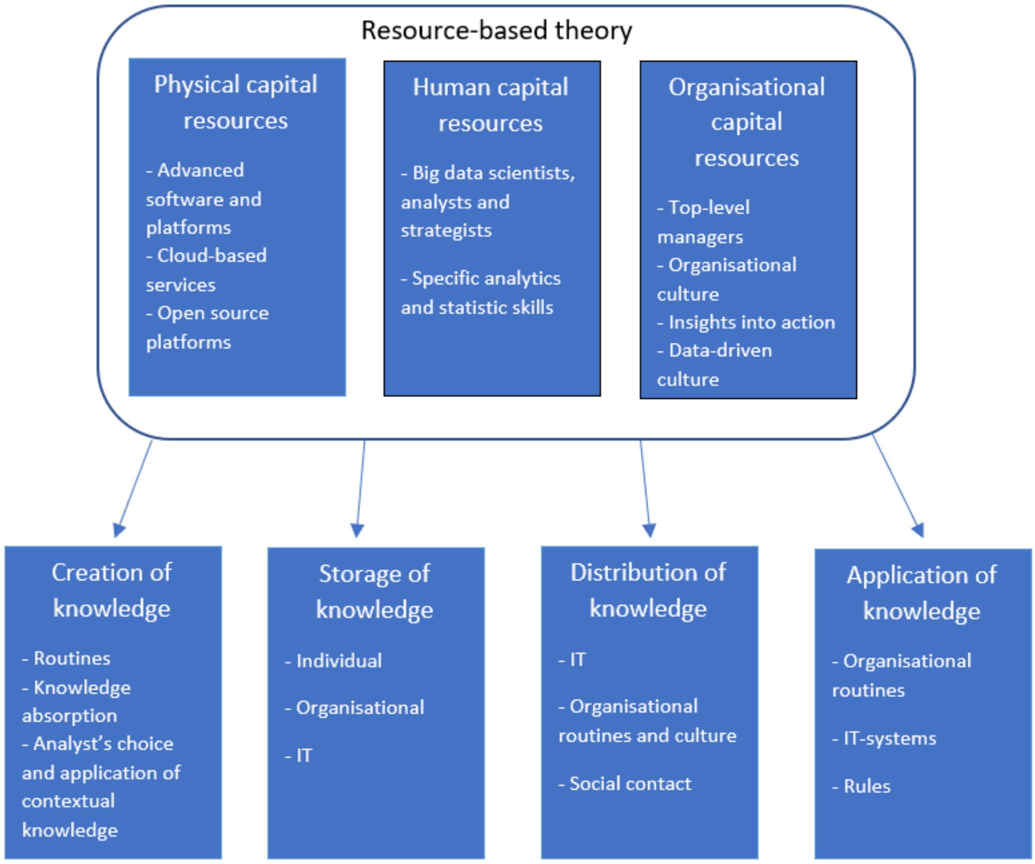

Figure 2. Fundamental resources and knowledge management processes of big data analytics according to literature.

Having established a thorough overview of the big data analytics concept in the previous chapter, the following chapters will illustrate the two disciplinaries and their relationship with big data analytics. The aim of the following chapters is to depict the ways in which organisational resources and knowledge management processes connect to and relate with big data analytics. When studying the resources, the impact of big data analytics will be studied regarding physical capital, human capital and organisational capital resources. As knowledge management is studied as a process consisting of four phases, hence the connections between these phases and big data will be analysed.

2.2 Resource-based theory and big data analytics

Resource-based theory studies firm's resources in a holistic way, meaning it considers both tangible and intangible resources. As presented by Morgan (2012) arrays of tangible assets include the organisation's equipment, factories, buildings as well as inventory. Whereas, intangible assets include the nonphysical matters like knowledge, patents, brand reputation and recognition. Nevertheless, both tangible and intangible resources define assets available to the organisation. According to the resource-based theory presented by Barney (1991), resources hold the potential of facilitating competitive advantage when the resource is either valuable, rare, imperfectly imitable, or non-substitutable.

According to Barney (1991), resource can be deemed valuable when it improves the company's effectiveness and efficiency. A valuable resource usually provides something of value to customers that competitors cannot achieve. In cases where the valuable resource is also unique, it will also be deemed rare and thus generate the company a competitive advantage. Whereas, inimitable resource indicates that the resource cannot easily be copied by competitors. Non-substitutable resources indicate that a resource should not be strategically equivalent to another resource. Therefore, two non-substitutable resources cannot be utilised separately to implement the same strategy. Having non-substitutable resources can be seen as a key factor in generating sustainable competitive edge for a company.

As Mikalef et al. (2019a) argument, big data analytics capability comprehensively includes organisational resources that are significant in transforming harvested data into actionable insights and enabling execution of those insights into action through operational and strategic decision-making. Therefore, the orchestration of organisational resources as an efficient and an agile entity enables organisations to successfully capitalise on and generate sustainable competitive advantage (Gupta & George 2016). Thus, making organisational resources the antecedents of big data analytics upon which organisations are capable of generating value.

2.2.1 Physical capital resources of big data analytics

According to Barney (1991) resources of a company include physical capital resources, human capital resources, and organisational capital resources. Additionally, another categorisation of the resources has been introduced in an article by Kozlenkova, Samaha & Palmatier (2014) where the main resources are classified as physical, financial, human and organisational. Nevertheless, physical capital resources consist of the technology used in an organisation. Furthermore, the organisation's facilities and equipment as well as its geographical location and access to raw materials are considered as physical capital resources. Financial resources include all the money in its numerous forms an organisation possesses or is accessible to. As presented by Scarpellini, Marín-Vinuesa, Portillo-Tarragona & Moneva (2018), the organisation’s access to capital via credit institutions, venture capital or individual funds as well as the possible availability of public funds are considered as company’s financial resources.

In the era of big data, physical capital resources include the advanced software or a platform that is used to collect, store, or analyse big data (Erevelles et al. 2016). The specific technology required for big data utilisation is already tightly associated with the term big data which depicts its importance in conducting successful big data analytics (De Mauro et al. 2014). As stated by Abbasi et al. (2016), the challenges posed by big data have nudged organisations’ and IT-departments’ to focus on distributed storage architectures that can handle enormous quantities of complex, unstructured data. Furthermore, the volume and velocity of big data have pushed a shift from physical on-premises data centres to cloud-based offerings. Davenport et al. (2012) also point out the delivery of big data capabilities through cloud-based services as a disruptive force of the ways big data is changing the technology. They also emphasize big data analytics dependency on extensive storage capacities and processing power, that in today’s constantly altering world needs also to be flexible and easily reconfigured according to different needs. This dependency is also one factor driving towards the cloud-based services and offerings and creating and operating through flexible platforms. What many of the authors name as novel and innovative products for dealing with big data are open source platform software systems, that are designed solely to support and process the enormous quantities of data generated and managed (Davenport et al. 2012; De Mauro et al. 2014; Rajaraman 2016).

As Chen et al. (2012) argument, the increasing amount of vast and complex information available in the internet for gathering and for organising and visualising requires specific and novel text and web mining techniques and systems. These systems must be integrated with mature and scalable techniques in text and web mining as well as in social network analysis. What De Mauro et al. (2014) also name as a fundamental element for technologies during the big data era, is the ability to store increasing quantities of data on smaller physical devices. Although, the storing capacities of computers are constantly growing, big data storing requires innovative methods and systems (Rajaraman 2016). Davenport et al. (2012) name virtual data-marts that enable an efficient way of sharing existing data as well as data hubs as systems for big data storing. Therefore, the ways in which big data analytics is changing and requiring from organisation’s physical capital resources is the capability of storing immense amounts of data, the power to process it and the ability to collect it, while being an agile and a flexible service that is designed for discovering patterns and opportunities while also being easily reconfigured. Thus, driving the processes mostly into cloud-based platforms and databases.

2.2.2 Human capital resources of big data analytics

Barney's (1991) definition of human capital resources include the abilities, activities and cognitive functions of the individuals working in an organisation. Therefore, the training activities and relationships as well as judgement, intelligence, experience and insights of individuals are considered as human capital resources. The individual employees and employers of an organisation and their abilities may enable organisations to achieve and to construct value-increasing strategies. Nevertheless, these attributes will not function optimally if the organisational capital resources are hindering the value-creating processes (Gonzalez & Martins 2017). In the context of big data, human capital resources include the data scientists, analytics and strategists that handle and analyse big data. They are experienced in capturing information from consumer activities and managing and extracting relevant insights from the data at hand for the company to capitalise on (Erevelles et al. 2016). Additionally, these kinds of resources are used to discover and to create opportunities and thus to enhance the company's dynamic capabilities (Cepeda-Carrion, Martelo-Landroguez, Leal-Rodríguez & Leal-Millán 2017; Teece 2007).

Having data-savvy analytical professionals that handle, work with and process data as organisation's human capital resources has been normal throughout the years. Nevertheless, as Davenport et al. (2012) point out, the requirements for data analytics support personnel are entirely different now in the era of big data. The interaction with and handling of the data itself as well as obtaining, structuring and extracting it is critical with big data, hence the personnel handling big data must have substantial and creative IT-skills. De Mauro et al. (2014) also agree that the sole process of analysing extensive quantities of data and the demand for identifying valuable information from complex data content require data processing methods that are notably advanced and demanding when compared to the traditional statistical techniques, and therefore require specific skills from the organisation's human capital resources.

Since the main objective of big data analytics is to capitalise it in the organisation's decision-making processes, the managerial level of an organisation is also affected by big data. As stated by Gupta & George (2016) the human resources that are specific to big data analytics are technical and managerial skills that both are of great importance when building successful and sustainable big data analytics processes. Top-level managers also possess the power to hinder or to enhance the organisation's tendency to use big data and in the creation of data-driven culture. As Mikalef, Boura, Lekakos & Krogstie (2019b) have found that resistance towards conducting data-driven decision-making as opposed to traditional and previous ways of making decisions has a notable impact on the organisation. This kind of resistance dwelling in and originating from the organisation's managerial level has an immense negative effect on the efficiency and success of building big data analytics capabilities.

2.2.3 Organisational capital resources of big data analytics

As stated by Barney (1991), organisation's formal reporting structure, controlling as well as planning and coordinating systems are considered as organisational capital resources. Furthermore, the interrelationships between groups and group dynamics in an organisation as well as the relationships between a company and external partners within its environment are structures of organisational resources. Organisational capital resources include the organisational structure that enables the transformation of insights into action. That is, to nourish such an organisational culture that encourages and engages the company to act upon insights. Cepeda et al. (2017) state that an organisation should be capable of reconfiguring its resources to establish sustainable competitive edge without compromising other changes occurring in the organisation. According to the framework presented by Erevelles et al. (2016), to successfully incorporate big data and big data analytics into an organisation's processes and to gain sustainable competitive advantage requires that all the resources are used to transform consumer activities into an advantage at different stages. By doing so, the company would enjoy of a sustainable competitive advantage with valuable, rare, inimitable and non-substitutable resources that all are generated by successful handling of big data

Barney (1991) arguments that although the main resources are interrelated and affected by each other, it is only some attributes that enable the company to implement effective strategies. As some attributes may even hinder the company from implementing valuable strategies while others may have no impact on any of the company's strategizing activities. As Teece (2007) arguments, successfully conducting value-creating processes, through which a sustainable advantage can be achieved, requires more than owning valuable, rare, imitable and non-sustainable resources. As stated by Fahy (2000), value is gained when an organisation effectively arranges resources in its product-markets. Therefore, the emphasis is on strategic choice and on efficient management of resources. Thus, leaving the responsibility of identifying, developing and arranging key resources efficiently, also in regard of resources needed for big data analytics, to the organisation's managers.

As Gupta & George (2016) illustrate, organisation's top-level managers should not only focus on establishing a data-driven organisational culture but also aim to maintain and enhance organisational learning when pursuing successful big data analytics capabilities. By data-driven culture, the authors define an organisational culture where decisions are based on data rather than simply on intuitions. Whereas organisational learning consists of the abilities to explore, store, share and apply knowledge that are possessed by the individuals of the organisation. What Davenport et al. (2012) also emphasise, is the possibility of development of organisations into information ecosystems, where information is constantly shared by internal and external service networks, communication about results is open and the mutual aim of the organisation is to generate new insights.

2.3 Knowledge management and big data analytics

Knowledge itself is an ambiguous and an abstract concept and the definition of knowledge has been a debate amongst academics across disciplinaries for many years (Gao, Chai & Liu 2018). Nonaka (1994) defines knowledge as a dynamic human process of justifying personal beliefs to attain truth. Therefore, knowledge stems and is directly related to human mind, making it also intangible in nature. Gonzalez & Martins (2017) state that knowledge is a result of an evolutionary cycle occurring in human minds. The cycle of knowledge is a flow that begins with data, which develops into information, that will develop into realisation. The following level is action and reflection based on the realisation and the cycle ends with the individual gaining wisdom. Nevertheless, knowledge is a process taking place in the minds of humans.

Although, knowledge is intangible in nature, can it be divided into two categories: tacit knowledge and explicit knowledge. Explicit knowledge represents knowledge that can be codified and documented in a tangible form (Jasimuddin, Klein & Connell 2005). Therefore, it can be expressed in words and numbers and shared in the form of data and manuals (Roberts 2000). Explicit knowledge can be easily and systematically communicated between individuals. Nevertheless, as this feature of explicit knowledge makes it easily available for even large numbers of people, it also makes the knowledge itself easy for competitors to imitate or copy. According to Nonaka & Konno (1998) tacit knowledge, on the contrary, is knowledge that is possessed by people. It is affected by emotions, values and experiences as it also is intertwined to the individual’s actions. Therefore, making tacit knowledge highly personal and hard to formalise or to document. It also complicates the communication or transmitting of tacit knowledge with other people. As Nonaka (1994) claims, tacit knowledge is acquired through experience. Tacit knowledge encompasses subjective insights, intuitions and hunches which can be described as the cognitive dimension of tacit knowledge. The other dimension of tacit knowledge is more technical, it is the informal skills a person possesses or the know-how skills

It is important to distinguish information from knowledge, although the two terms may occasionally be used correspondently. Knowledge relates to human action and it can be deemed as the skill, vision, experience and concept that is organised and created by information flows (Nonaka 1994; Gao et al. 2018). As also stated by Intezari & Gressel (2017) knowledge is the concept that provides a broad and deep understanding of data and information. Knowledge also presents a framework for evaluating and incorporating new experiences and information. This is due to fact that knowledge combines experience, values, contextual information, and other insights. The entire process is a myriad of complex cognitive processes occurring in the human mind. Therefore, information can be defined as processed and meaningful facts or flows of meanings that may add to, restructure or change knowledge (Nonaka 1994). Whereas data is, in its basics, a set of facts and therefore cannot be defined as knowledge nor information.

Nevertheless, in the era of big data even the data sets are complex, enormous and constantly altering. Therefore, big data challenges the process of obtaining the meaningful information from the data sets from which the deeper and broader understanding of the facts can be created (Intezari & Gressel 2017). Hence, providing valuable knowledge for the organisation to capitalise on. Furthermore, as the technologies in the world are constantly advancing and becoming more powerful, the role of collecting, generating and managing of information and knowledge increases. As knowledge is originated in the minds of people and with effective management practices it can bring strategical value to the organisation. (Hota, Upadhyaya & Al-karaki 2015)

When analysing knowledge management and big data, one notable impact is on the knowledge management systems. As stated by Intezari & Gressel (2017), knowledge management systems are certain information systems that are designed and implemented for managing organisational knowledge. Whereas Dayan & Evans (2006) describe knowledge management as a systematic effort to comprehensively manage knowledge assets inside an organisation. The knowledge management effort should be integrated in the organisation's operational and business objectives, and to be conducted in a measurable manner, to achieve innovativeness and competitive advantages. The aim of knowledge management is to identify the assets and expertise available within the organisation and thus to increase its value. As mentioned by Dayan & Evans (2006), the most essential and valuable assets of an organisation are likely possessed by the personnel who have the knowledge, not the organisation's products nor services. Additionally, knowledge management aims to promote the flow of knowledge inside the organisation (Gonzalez & Martins 2017). Knowledge management itself involves social and cultural facets, yet it relies on information-technologies as its enabler. As Alavi & Leidner (2001) claim, knowledge management systems are IT-systems that facilitate and support in creating, circulating and implementing knowledge in organisations. Traditionally, knowledge management systems are used to identify, share and capitalise on knowledge. Additionally, they are used to incorporate knowledge into processes where problems are identified and solved in business environments. As big data is changing the nature of data entirely, the volume, velocity and variety of data avalanching into the companies is immense. Therefore, the management of the data becomes more challenging as does the ways in which it is processed into valuable information and knowledge.

Naturally, the current traditional knowledge management practices or knowledge management systems are inadequate to handle big data. Therefore, companies need to establish more advanced practices and systems for knowledge management regarding big data. This would, according to Intezari & Gressel (2017) potentially mean incorporating advanced knowledge management systems that do not simply link knowledge repositories to data storages, rather aim to incorporate big data into the organisation's strategical decisions. Having advanced knowledge management systems for big data, the organisation can increase value by providing immediate performance feedback and more objective decision-making by incorporating algorithms into the decision-making processes. Furthermore, big data enables companies, with necessary resources, to operate cost-efficiently, effectively and with agility. Nevertheless, all of the decisions made upon big data utilisation require valuable and relevant knowledge. As stated by Pauleen & Wang (2017), human knowledge is the main decider of how the obtained information from data is used in the organisation to gain strategical advantages. Only then can the big data be exploited efficiently, and sustainable competitive advantages achieved.

Concept of knowledge management as a process and its structure has been discussed by many researchers throughout the years. Being a rather complex and an abstract term, the idea of the concept as well as the structure vary depending on the study (Gao et al. 2018). Nevertheless, many of the studies have identified four steps of knowledge management process that appear to be more fundamental than others. Those steps are creation, storage, distribution and use of knowledge (Gao et al. 2018; Gonzalez & Martins 2017; Durst & Edvardsson 2012). Nevertheless, as the focus of this study is to analyse knowledge management as the core-process of utilising big data analytics, the framework of knowledge management process is altered slightly. As stated already in the beginning of this chapter, the value derived from big data analytics is generated by the insights, knowledge, and relevant information extracted and created from big data analytics with the support of human individuals. Therefore, rather than focusing solely on the general management of knowledge, this study encompasses big data as the source of knowledge. Hence, the following figure illustrates the knowledge management process, where the focus is to manage the knowledge namely derived from big data. The process of knowledge management in the big data era is illustrated in Figure 3.

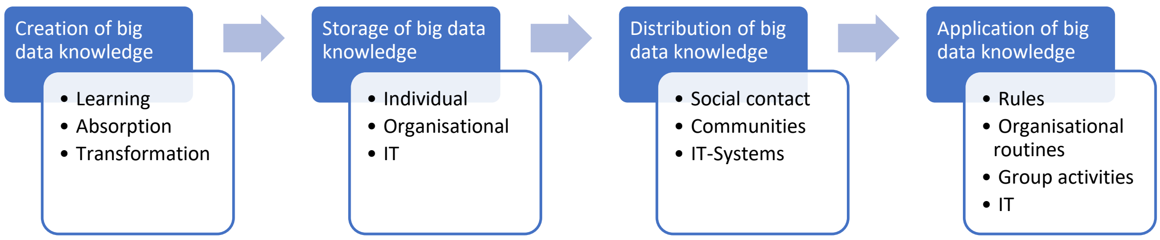

Figure 3. Knowledge management process and stages (Adapted from Gonzalez & Martin 2017).

As stated by Gao et al. (2018) knowledge creation phase consists of the processes where novel knowledge is created. Big data analytics can be a means of developing novel knowledge through acquisition of new content or by replacement activities when existing content is replaced by tacit and explicit knowledge. The next phase, storing of the knowledge developed from big data, describes the process of storing and recording knowledge in the organisation's storage systems. Such repositories are archives, databases and filing systems – usually knowledge that is stored in repositories is explicit in nature. The aim of knowledge storage process is to enable the transmitting of the knowledge for other people to be applied for and used. According to Argote & Ingram (2000) the distribution of knowledge, or the transfer of knowledge, is critical phase in the knowledge management process. Basically, it refers to the process of distributing knowledge and experience between organisational units. Therefore, one organisational unit is influenced by another unit's experience or knowledge. Hence, knowledge distribution process generates change in the recipient units and thus in the knowledge base of an organisation. Lastly, the knowledge application phase which refers to actualising of knowledge (Gao et al. 2018). By actualising the knowledge, the organisation can capitalise on it and use it for strategic purposes. Each of these phases are analysed more thoroughly in the following chapters.

2.3.1 Creation of big data knowledge

In their article Gonzalez & Martin (2007) identify fundamental matters regarding knowledge acquisition process. These matters are organisational learning, transforming the organisational knowledge and knowledge absorption. According to Zollo & Winter (2002), organisational learning process stems from two different activities – from routines and from experience accumulation. Organisational routines encompass the operational activities and the functionality of an organisation. Therefore, routines are patterns of behaviour that illustrate organisational reactions to either internal or external incentives. The routines can be procedures that are already known patterns of increasing the organisation's revenue or procedures where the routines are altered, and thus new patterns to increase the organisation's competitive advantage are established. The first method can be defined as utilising the organisational capabilities whereas the second can be deemed as an activity that enhances the organisation's dynamic capabilities (Gonzalez & Martin 2007). The way in which big data analytics affect the organisational routines is mostly by altering them and by creating new patterns. As previously the decision-making routines of the organisation's top management level were mainly based on intuitions, the decision-making routines with big data are transformed into data-based decision-making (Ferraris, Mazzoleni, Devalle & Couturier 2018). Furthermore, as stated by Davenport et al. (2012) the routines of analysts, IT-specialists and data-handlers are changed entirely when interacting with big data.

Experience accumulation refers to the process of improving organisational routines by accumulation of experience and specifically to accumulate tacit knowledge. This process is not dependent on the nature of the knowledge (big data knowledge vs traditional data knowledge) nor experience (experienced big data analyst vs traditional data analyst), the process itself and its objectives remain the same, only the means of accumulation may vary. As Nonaka (1995) claims, successful knowledge creation is identifying tacit insights, intuitions and hunches of individual employees and making those insights available for utilisation by the organisation. Furthermore, Zollo & Winter (2000) agree that experience accumulation is a critical learning process for developing operating routines. Basically, experience accumulation refers to activities that enable the individuals of an organisation to gather and discuss their respective experiences and beliefs with the aim of sharing tacit knowledge and as a result to improve the organisational operating routines. Not only is the focus on facilitating experience accumulation but also on absorption of experimental wisdom. Knowledge absorption refers to the organisation's ability to identify and comprehend the value of certain knowledge and to accommodate it to achieve competitive advantage (Cohen & Levinthal 1990). To which Pauleen & Wang (2017) agree that when newly identified organisational knowledge from big data analytics is reused as a part of contextual knowledge, the organisation can successfully manage its knowledge and gain value from it.

According to Pauleen & Wang (2017), the ways in which novel knowledge is created through big data analytics is through the analyst's choice and application of contextual knowledge. When the analyst chooses the specific analytic tools for identifying new knowledge, the human knowledge, experience and even innovativeness the analyst possesses impact the resulting knowledge that is generated from the big data. The resulting new knowledge from the analytics process will become a solution for previously defined problems or to initiate subsequent organisational actions to improve the performance. Ferraris et al. (2018) also point out that the value extracted from big data is not only dependent on the quality of the data but also on the quality of different processes in which the data is collected and analysed.

2.3.2 Storage of big data knowledge

Organisation's knowledge storage process can be thought to rely on three different entities – individual, organisational and IT (Gonzalez & Martins 2017). Individual knowledge refers to the tacit and explicit knowledge possessed by an individual in an organisation. Naturally, this knowledge is affected by personal beliefs, motives, emotions as well as experience. As stated by Grant (1996a), all tacit knowledge and most of explicit knowledge are stored in individuals. Nevertheless, most of the knowledge is created within an organisation and thus, is specific to the organisation. The environment, in which the individual is, affects notably to the ways in which individual knowledge is developed, increased and shared. Pauleen & Wang (2017) argument that to create an environment for data collection at an operational level, the managers and professionals of an organisation need to establish an infrastructure and organisational systems parameters that are based on contextual knowledge.

According to Alavi & Leidner (2001), organisational refers to organisational culture, structure of an organisation, internal processes and procedures as well as internal and external information archives. Organisational culture is a mean of storing and transmitting organisational knowledge through norms, beliefs and values that are commonly established and agreed upon by the groups and individuals of an organisation (Gonzalez & Martins 2017). In the context of big data, the organisational culture should be data-driven, for successful exploitation of big data analytics (Gupta & George 2016). Furthermore, since big data carries significantly different attributes compared to traditional data – the internal and external information archives of an organisation are likely to change to more advanced systems. Additionally, organisational knowledge consists of codified human knowledge stored in expert systems and tacit knowledge obtained by individuals and groups of an organisation. Lastly, as presented by Alavi & Leidner (2001), IT-systems support storing both individual and organisational knowledge for the benefit of the organisation. Big data will have most notable impact on the knowledge storage IT-systems, simply due to the complexity and volume of the data. IT-storage systems and technologies such as digital databases, intranets, and repositories in general where all relevant information and knowledge of an organisation can be stored to enhance, develop and increase the organisational and individual knowledge.

2.3.3 Distribution of big data knowledge

According to Argote & Ingram (2006) knowledge transfer or distribution refers to the process where one unit, a unit can be a group, department or division, is affected by another unit's experience or knowledge on big data and big data analytics. One important aspect of knowledge distribution is that it generates changes in the knowledge or performance of the recipient units. This change in knowledge or performance can also be used to measure knowledge distribution. Nevertheless, due to different features of knowledge the measuring of knowledge distribution is also facing some challenges. As the knowledge organisations acquire may be tacit in nature, it may not be entirely captured through verbal communication, that usually is used to measure knowledge. As stated by Davenport et al. (2012) the data scientists who work closely with big data must possess not only advanced analytics skills but also the ability to communicate effectively with decision-makers. Thus, to ensure effective and fluent distribution of knowledge and experience extracted and gained from big data analytics. Another challenge, named by Argote & Ingram (2006), regarding measuring of knowledge distribution is caused by knowledge residing in multiple repositories and to measure the distribution of knowledge, the changes in all different repositories must be captured. Such knowledge repositories are for example the organisation's individual members and the roles and organisational structures - repositories where knowledge resides in organisations.

As Gao et al. (2018) claim, there are three aspect through which knowledge distribution can be analysed. These aspects consist of the exchange of experiences and knowledge between individuals through social contact, sharing knowledge through communities of practice and distribution of explicit knowledge supported by IT. As stated, explicit knowledge can be distributed by IT-systems, but also social interactions can be means of transferring explicit knowledge. By sharing knowledge, the people can contribute in establishing a knowledge network, that is supported by IT. Alavi & Leidner (2001) agree by claiming that IT can increase knowledge distribution process by extending the individuals reach beyond formal communication boundaries. Usually, knowledge sources are limited to immediate colleagues with whom an individual is in regular and routine contact. Furthermore, these immediate work networks tend to consist of individuals that possess similar information and thus are not likely to offer the individual new knowledge.

On the contrary, IT-systems such as computer networks and discussion groups or repositories provide a space, where the individual looking for knowledge and the people who have access to or possess the required knowledge can contact each other (Alavi & Leidner 2001). Additionally, Gao et al. (2018) present communities of practice to define groups of individuals who actively exchange knowledge. They also develop a common identity and own social context that facilitate the knowledge sharing process. They tend to manifest themselves through behavioural uniqueness and by reflecting a specific community, where knowledge can be easily shared. Therefore, knowledge distribution as a process requires use of IT-systems to distribute explicit knowledge that is supported by organisational routines and culture that enhances and enables social contact between individuals and groups to distribute the existing tacit knowledge.

2.3.4 Big data knowledge application

In their article, Gao et al. (2018) define knowledge application as an ability of the organisation's individuals to discover, identify and utilise the knowledge that is stored in the organisation. Additionally, Alavi & Leidner (2001) claim that knowledge application is the source of organisation's competitive advantage. The aim of knowledge application, according to Gao et al. (2018), is to develop new knowledge through integration, innovation and extension of existing knowledge base, as well as to be used in decision making. In the context of big data, the new and novel knowledge extracted from the big data analytics, through innovative and advanced analytics tools and human decisions will extend the existing knowledge base and work as a basis for executing data-based decisions or organisational activities (Pauleen & Wang 2017). Grant (1996b) has presented mechanisms to integrate knowledge to gaining competitive advantage. These mechanisms are rules, organisational routines and group solving and decision making. Rules define and are an essential construct of human interaction and rules regulate the interaction between individuals. Such rules are standards and instructions that are developed as tacit knowledge possessed by a competent individual is converted into explicit and integrated knowledge for individuals and groups, who lack the knowledge, to be easily communicated to and thus used for (Alavi & Leidner 2001).

According to Grant (1996b), organisational routines are defined as complex patterns of behaviour generated by slight signals or choices. The resulting behaviour is seemingly recognisable and conducted in a fairly automatic manner. Routines support interaction between individuals in situations where rules and directives as well as verbal communication are astray. Therefore, routines allow individuals to integrate and implement the knowledge they possess even without articulating or communicating their knowledge to others. Additionally, as stated by Alavi & Leidner (2001), knowledge application can be enhanced by technology, as it enables embedding of knowledge into the organisational routines. Organisational and culturally specific procedures can be integrated into IT-systems which will then depict the organisational norms in an efficient and clear manner that is easily accessible by all. Lastly, group problem-solving and decision-making define groups of individuals that possess necessary knowledge for solving complex, unusual and important matters (Grant 1996b; Alavi & Leidner 2001). During the era of big data, the individuals that possess the necessary knowledge for solving emerging problems are not necessary the top management level groups of individuals, who tend to execute intuition-based decisions rather the data scientists who, with the help of big data analytics, are capable of conducting efficient, rapid and effective solutions and decisions based on data (Ferraris et al. 2018). Therefore, usually in the context of big data analytics, the application of knowledge is executed through group problem-solving where the group consists of individuals who are experienced in interacting with big data.

3 RESEARCH DESIGN AND METHODS

This study's empirical part is based on and conducted through qualitative research methods. The aim of the empirical part is to provide answers to the research questions as well as illustrate insights that are relevant to the study's framework. Additionally, the focus of the research is on a case company; hence the empirical part aims to illustrate the case company's perspective on the subject. Case study was chosen as a research method to investigate a phenomenon thoroughly within its real-world context and to understand the related contextual conditions (Yin 2014, 16-17). For the study, employees from the case company with different responsibilities were interviewed to gain relevant and multifaceted data.

Following this chapter, the methodology as well as the selection of the conducted research will be explained. To continue, the data collection methods and practices are presented in more detail as well as the precise execution of the analysis of the collected data. Subsequently, the reliability and validity of the research are analysed and examined. Lastly, a short description of the case company is provided.

3.1 Methodology

This research was conducted as a case study, one form of qualitative research, to comprehensively study and analyse the research topic. As stated by Hirsjärvi Remes & Sajavaara (2007, 157) qualitative research methods help in understanding, describing and analysing comprehensively the target of the research as well as a phenomenon in real-life environment. Therefore, when studying the relationship between big data analytics, organisational resources and knowledge management (the phenomenon) the environment would naturally be organisations where this phenomenon occurs. Thus, the case-study method was deemed appropriate for this research to thoroughly understand the interrelationships of the subjects under inspection as well as their impact on the surrounding environment. Through this, the aim is to answer this study's research problem and to illustrate new insights to the subject.

As Yin (2014, 16-17) also states, the selected case company should be related to the study's theory. Therefore, the case company for this research was selected based on its active operations and comprehensive actions with data. By studying an organisation that actively interacts with data, it enables the analysis of the contextual environment where the data analytics processes enabled by organisational resources that allow the extraction of knowledge from the data and how that knowledge is then managed, that are of great importance and relevance to this research. Furthermore, both the phenomenon and the environment are related and thus, it is relevant to study both thoroughly. As organisational resources and organisational culture are tightly intertwined and relevant parts of the phenomenon, it is suitable to study the organisational environment in its entirety in order to understand the contextual conditions (Alavi & Leidner 2001). As stated by Yin (2014, 16-17), the most efficient way to understand the phenomenon and the contextual conditions is through a case-study.

3.2 Data collection

As stated by Yin (2014, 118), the data collection of a case study should be constructed upon multiple sources of evidence. Therefore, the data for the empirical part was collected through interviewing the employees of the case company as well as through secondary data from the case company’s public documents. The public documents and information are used to provide a description of the case company and to clarify organisation-specific terms and processes that came up during the interviews. Interviewing as a data collection method was selected to get a deeper understanding of the case company’s contextual environment where the employees act as active parties and to integrate meaning into the research (Hirsjärvi et al. 2007, 200). Furthermore, as the focus of this study is to analyse knowledge management’s perspective in big data analytics, it is logical to use the stems of knowledge, human minds as Nonaka (1994) claims, as the main source of the data. As stated by Hirsjärvi et al. (2007, 207) interviews enable a more profound overview of the cognitive sides of humans – through conversation the interviewees can express their emotions, thoughts, feeling and beliefs more naturally. When studying a phenomenon (data analytics practices) occurring in a particular environment (case company), gaining profound data that contains not only factual information, but also cognitive features and depictions of personal experiences of the employees will help in conducting a thorough analysis that accounts relevant aspects more profoundly. The factual information concerning the company’s business and operations regarding data analytics was collected through the public documents to support the data and insights gained from the interviews.

The interview was constructed following theme interview guidelines. Therefore, the interview questions are based on the framework of this research and hence are divided into three categories: big data analytics, organisational resources and knowledge management. Nevertheless, big data analytics is present in both latter categories to gain a more thorough outlook on the subject. The interview questions are presented in Appendix 1. Theme interview as a data collecting method was suitable for this research as it provides the necessary and logical structure to the interview without restricting the conversation, rather enabling a free flow of natural conversation under the selected theme (Eskola & Suoranta 1998, 86). The nature of the interviews allowed posing additional questions and ask for clarifications while also maintaining a relaxed and a natural atmosphere. All interviews were conducted face-to-face in closed meeting-rooms and recorded, with the interviewees consent, to ensure efficient storage and analysis of the collected content. The first interview took place in the end of July 2019 and the last was done in the end of August 2019, all of the interviews were therefore conducted during three months' time. The meeting-rooms provided a clear sound environment and also tranquillity to focus on the on-going interview as the room excluded external distractions. The recordings of the interviews were transcribed which resulted in a separate 36-page document (font size 11 and spacing 1.0) that was used as a basis for the data analysis of this research. All interviewees were delighted to participate in the research and expressed interest in the topic. Interviews were conducted and the transcriptions were written in Finnish, except for the one interview that was conducted and transcribed in English. Direct quotes presented in this research are therefore English translations of the interview transcriptions.

To gain relevant insights regarding the topic of big data analytics, organisational resources and knowledge management, the selected interviewees were ones who actively interact and work with data in a data-driven environment. Furthermore, as the aim of this research is to study the relationship of the three constructs, organisational resources, knowledge management and data analytics, it was important not to restrict the interviews to simply focus on data analysts but also to broaden the scope of selection. The versatility of the interviewees enhances the collected data, gives more perspectives about the topic and enables the analysis of matters such as group dynamics and other organisational culture aspects which are relevant to the study. Hence providing a deeper outlook on the phenomenon taking place in the environment. Therefore, the selection of interviewees consists of employees with varying titles and job descriptions, nevertheless data impacts each interviewee's every-day work. The interviewees represent two units and are presented below in Table 1. Some are employees of Aller Media's Analytics & Business Development team while others are employees of Data Refinery which is Aller Media Finland's subsidiary. Nevertheless, all of the interviewees work in the same office and also quite often work together.

Table 1. The selection of interviewees and interview durations

Interviewees	Job title	Unit	Duration
Interviewee 1	VP, Tech & Development	Data Refinery	00:23:20
Interviewee 2	Junior Data Analyst	Data Refinery	00:24:53
Interviewee 3	Account Director	Data Refinery	00:33:21
Interviewee 4	Project Manager	Data Refinery	01:03:21
Interviewee 5	Lead Data Scientist	Aller Media	00:26:11
Interviewee 6	Data Analyst	Aller Media	00:41:19
Interviewee 7	Junior Data Analyst	Aller Media	00:31:54

3.3 Data analysis

The aim of the data analysis is to provide clear and meaningful information by creating coherent content from the dispersed data (Eskola & Suoranta 1998, 137). Therefore, content analysis was conducted by comparing the empirical findings to the content presented in this study's theoretical part. Hence, the analysis consisted of finding themes that emerged from the data as well as are identifiable from the theoretical framework (Eskola & Suoranta 1998, 174). By comparing themes from data with the theory, the structure of the entire research will remain consistent and logical as well as help in providing relevant and coherent results. The themes which were used for coding the data and upon which the interview questions were built are as presented in the theoretical part and in the theoretical framework of this study. The main categories were:

- Big data and organisational resources (human, physical and organisational capital)
- Big data and knowledge management (creation, storage, sharing and application)
- Challenges and important factors regarding successful and sustainable big data analytics

After having transcribed all the recordings of the interviews into one data pool, the data was analysed accordingly and systematically. Fortunately, as the interviews followed the theme interview guidelines and the questions were constructed upon the theoretical framework of this study, they both provided good basis for the data analysis. Primarily, the transcriptions were read thoroughly and categorised by intuition. In other words, common themes and meaningful matters were identified from the interviews as well as differences between the interviewees' insights were noted. Secondly, each construct of the theoretical framework of this study was separately under inspection during the data analysis. Hence, the data was coded according to the themes provided by the theoretical framework of this research and then analysed systemically. During each coding round, the emerging insights of the particular theme where highlighted from the entire document. This procedure was executed systematically to each theme in turn. After the initial coding rounds, multiple rounds were rerun to identify possible subthemes from the data. After having coded thoroughly the data into different themes, the main ideas from each section was summarised to provide clear and concise observations.

3.4 Reliability and validity

As presented by Hirsjärvi et al. (2007, 227), when studying a phenomenon, more so when its occurrence is related to human and cultural aspects, it is relevant to ponder the meaning of reliability and validity, as the observations regarding the phenomenon and all its related features are, at least to some degree, unique. Furthermore, all of the variables are prone to alter in the course of time. As stated, this study aims to provide an outlook of the current state of big data analytics by studying a case company that operates actively in the field. Therefore, for example the selection of the interviewees, as presented in this study, may be different in possible future studies due to natural employee turnover.

Regarding the validity of this study, all of the interviews were executed face-to-face, which enabled the researcher to pose further questions and the interviewees were also able to ask for clarifications when necessary, both of which prevented possible misunderstandings and misinterpretations. Additionally, the interviewees were asked to read the analysis to prevent any misinterpretations by the researcher. As for the credibility of the collected data, all of the interviewees are in a position where data affects their work at some level. Nevertheless, it is understandable that for the employees who do not use data analytic tools and systems in their every-day work life nor have gained any prior experience about them, the question for example regarding physical capital resources of big data analytics might have seemed slightly out of their field of expertise. In this kind of situation, the interviewee could impress his/her uncertainty about the question and thus prevent compromising the sampling with uncertain information. Nevertheless, all uncertain ponderings were also recorded and analysed to prevent missing any noteworthy insights. The answers for example for the physical capital resources was received in great detail from the analysts and thus, sufficient and credible data was successfully collected.

The reliability of a research indicates that in a situation where the research would be executed again, it would produce similar results (Eskola & Suoranta 1998, 213). Therefore, the research must be constructed in such a way that it is easily repeatable, and all the phases and details of the research are clearly presented. As for the reliability of this research, all of the information necessary for repeating this study are presented in the earlier chapters, the used methods are described to the detail and the interview questions are presented in the appendices to increase transparency. Naturally, as the interviews were conducted following the theme interview guidelines where the questions provided a structure for the conversation, the exact form of and order of the questions varied, which is natural when having a conversation about a certain subject. Nevertheless, all of the themes presented in the research questions were addressed during each interview and the relaxed conversation-like situation resulted in good and credible data.

3.5 Case company description

Aller Media Finland is an organisation where content, in its multiple forms, is created. The organisation focuses on creating quality media and marketing content in Finland. All marketing and media operations are data-driven, and the organisation is a pioneer of digital and data business operations (Aller 2019c). Aller Media Finland's network consists of online services like Suomi24, Treffit, Seiska, Katso and Telvis. Through this network nearly 72 million pageviews are generated monthly as well as other massive and complex data (Aller 2019a). Not only does the organisation focus on digital and online operations but also to offline data that is generated from their business and from their customers. For example, Aller Media administers and publishes one of the most popular magazines in Finland (Seiska) that contributes to a massive offline database. As recently measured by Media Audit Finland, the reach of the published Seiska-magazine and the Seiska.fi -website is 677 000 weekly readers and thus, it has the biggest reach amongst its competitors (Aller 2019b). Hence, the organisation unarguably offers a data-driven environment that is suitable as a case company and relevant for analysis for this research. To gain a thorough overview of the entire data operations of the organisation, the subsidiary Data Refinery is also included in this research.

Data Refinery is Aller Media Finland's subsidiary and part of the marketing services provided by Aller. Data Refinery is a data house that has specialised in providing services and products that are entirely based on efficient and innovative utilisation of data. They enrich customer data and offer data-based customer understanding as well as data-based target groups for both digital and traditional channels (Data Refinery 2019a). In its operations the company utilises data from Aller Media Finland's websites and subscriber database. Furthermore, with Data Refinery's vast partnership networks that include for example data sets from official national databases, they manage a database through which almost every Finnish person can be reached (Data Refinery 2019c). From all of this data, the company can create multiple data-based online and offline target groups that are enriched with high-quality data (Data Refinery 2019c). Enriching of data means importing individual demographics information, forecasts, or data about interests derived from online behaviour to a database to increase its value and effectiveness (Data Refinery 2019b). Additionally, the company offers other data-based services like profile reporting, customer segmentation and prediction modelling. To conclude, Data Refinery is a company that possesses great expertise in interacting and operating with massive, complex and constantly altering data.

4 EMPIRICAL FINDINGS AND ANALYSIS

In this chapter the findings, analysis and results from the empirical data will be presented. The results are presented as according to the theoretical framework of this study, to maintain a consistent and a logical structure throughout this research. Therefore, each construct of the theoretical framework will be under inspection. Subsequently, findings regarding the research questions of this study are presented and analysed.

The main points of each construct of the theoretical framework are studied and analysed in detail in the following chapters. Additionally, the main findings regarding the antecedents of big data analytics, meaning organisational resources and knowledge management processes, are illustrated in Figure 4 below.

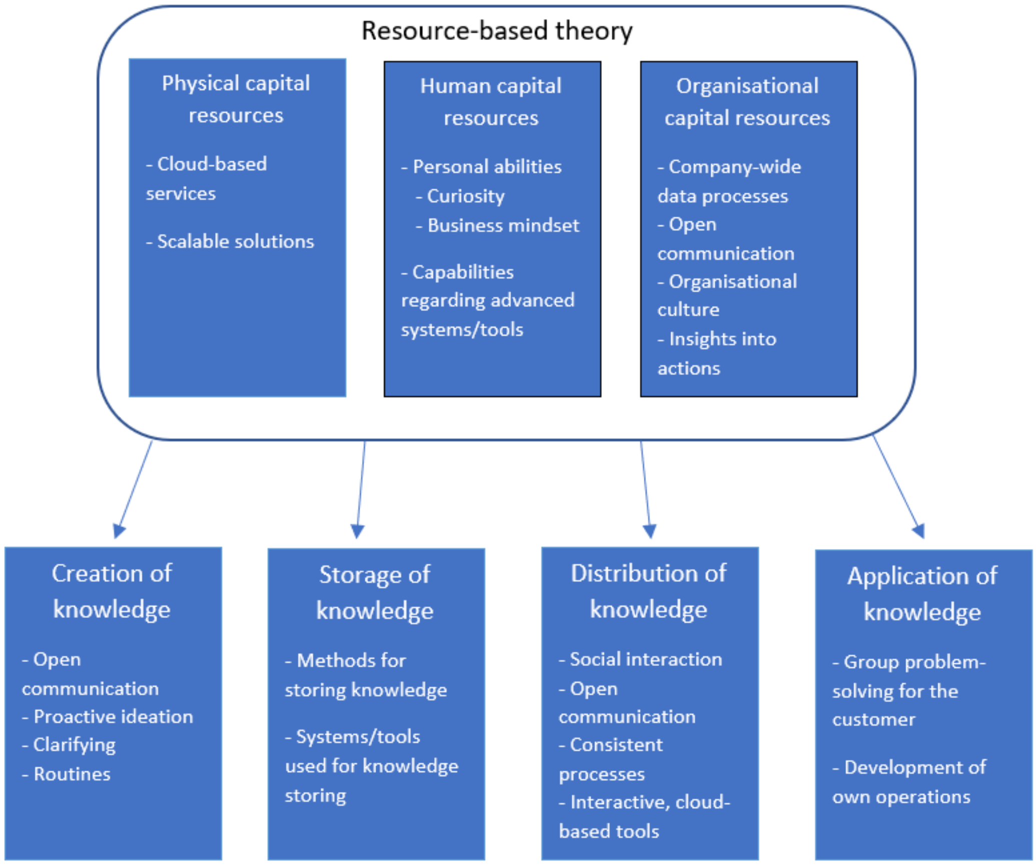

Figure 4. Fundamental resources and knowledge management processes of big data analytics according to the empirical data.

4.1 Big data and organisational resources

Human capital resources

Regarding the human capital resources of an organisation, there were two main themes that emerged from the research data. These two themes are personal abilities or know-how the analyst should possess, and capabilities regarding advanced systems and tools. An analyst or a data scientist that possesses these abilities and knowledge would, according to the research data, have the necessary competencies to prolifically interact with big data. As for the personal abilities and know-how of the analyst, some of the abilities tend to be quite fundamental and distinctive for an analyst or a data scientist. Such fundamental characteristics are abilities that have been important since the early days of data analytics and will most likely be important in the future. Such basic abilities are logical thinking, analytical mindset and precision. Additionally, coding and basic knowledge of data-analytics practices can be deemed as fundamental abilities for the position. For example, as coding is used to create commands for retrieving required information from the data. Nevertheless, the analyst must understand the basics of data-analytics to understand what kind of activities or commands are needed to acquire the necessary information and where to possibly find it. When analysing the future challenges and requirements for data analysts or scientists, the personal abilities and know-how advance.

The practices of comprehending, managing and interacting with big data require specific and advanced competencies. An analyst who possesses good mathematical and statistical abilities is likely to achieve a more comprehensive overview of the big data, which allows improved managing and handling of the data. While data volumes are constantly growing the analysts should possess advanced analytical and statistical mindset and understand the nature of data. Then working with huge masses of constantly altering data will not astound the analyst rather despite the commotion, the analyst is able to direct the focus on identifying slight and relevant signals or patterns from the data. Curiosity as an important ability for an analyst or a data scientist emerged from the research data. Curiosity would drive the analyst to proactively browse through the enormous databases to find new insights, or to develop the analytic processes and activities further to being more efficient and prolific. In case of constantly evolving, growing and challenging big data, the self-initiated development processes initiated by curiosity are likely to become key success factors in matching the requirements of the data of today and in the future.

> "In every section it's always good to be curious. Additionally, you should possess logical thinking and analytical mindset, as we're looking at a massive amount of scattered numbers and statistics that makes little to no sense, unless properly combined. You really have to come up with something that different target audience can understand. I have recently come across an interesting data metaphor in an article 'Not everyone can manufacture a car. But almost everyone can drive it'." (Junior Data Analyst, Aller Media)

Furthermore, not only is it important for the analysts to be curious, but also to obtain a business mind-set. Understanding the business' perspective of analytics and being able to identify business-related opportunities from the data analytics activities as opposed to focusing solely on the analytics is paving way for the future of data analytics. Moving from the databases into the meeting rooms, analysts with business mind-sets can identify relevant matters that may be unnoticed by other specialists. One fundamental idea of successful business is to understand the unknown needs of the customer and to match those needs. For the analysts to understand the meaning of the customer's needs, in other words to have a business mind-set, and to work upon those matters in the era of big data is important as it is arduous to browse the immense databases without any target. Therefore, having identified matters that are important to the customer, the analytics will take less time and generate more valuable results for the customer than when simply browsing the data without any specific direction.

> "Big data analytics require analysts. I'd say the most valuable analysts are those, who understand the business perspective [of analytics] and participate in the customer meetings as well. From there, they identify matters to grab onto, they take these matters with them to their office and communicate them to their team. They begin the analytics process based on those [identified matters] and start to browse open data sources and work proactively." (Account Director, Data Refinery)

In addition to challenging the personal abilities of an analyst, big data analytics also require advanced competencies regarding analytic systems and tools. The tools big data analytics require are different when compared to the ones that are mostly used and where the expertise currently is. Furthermore, as most of the tools, systems, databases and data warehouses are moved away from personal computers to cloud-based services due to the massiveness of big data, it generates additional alterations to the environment. These tools naturally require different and advanced know-how and expertise, to which the analysts and data scientists must prepare themselves for. For example, the servers and databases suitable for managing big data analytics require advanced technical expertise from the analyst or the data scientist. Fortunately, some technologies do utilise familiar commands and analytic techniques, while they are just optimised for processing more challenging data sets. Advanced technologies that are manageable with already available expertise, help in the transformation of the tools and systems from traditional data to big data. Since the primary focus can be directed to the more demanding data rather than learning entirely new systems and tools. Additional emerging trends are AI and machine learning. As the data volumes increase, the manual workflow is likely to become more arduous, which is why integrating AI and machine learning into the data managing processes will be necessary. Understanding AI and machine learning services and systems and having the ability to modify the data accordingly are key success factors in the integration process. Thus, providing new challenges for the data analysts and data scientists in the future.

> "The tools that big data analytics require are different compared to the ones that are familiar and where the strong expertise currently is. To learn to use these tools requires a lot of studying. Also, integration of AI and machine learning into the data managing processes requires that the work assignments focus on modifying the data so that it is easily managed [by machine or AI]" (VP, Tech & Development, Data Refinery)

Physical capital resources

As the case company, both Aller Media Finland and its subsidiary Data Refinery, work actively with data, one of its physical capital resources is its access to raw materials – in this case access to data. As stated before, the sources that are accessible by these two companies generate both online and offline data, that can be deemed big data. Naturally, in the case of offline data, the data does not alter nor grow rapidly, rather progressively over the decades. As for the online data, it meets all the characteristics of big data and the most notable big data source is the data generated by Aller corporation's website visitors. The magnitude of the data flowing into and managed by the case company covers almost the entire population of Finland and is undisputedly the nation's biggest consumer information database regarding magnitude. The focus of the company is not simply to capitalise on big data, rather to find innovative ways to capitalise on the data that is accessible despite its size. By enriching, combining and comparing data from different sources, either open data sources, own databases or customers' databases, to increase its value, the case company can provide innovative and efficient data-services for its customers.

All of the analytic tools utilised by the case company as well as its database are on cloud-based servers. That enables the managing of smaller data sets but also enormous data sets, owing to the deftness of cloud-based services. In case of physical resources, they might alter depending on the analyst's preferences or the means of application. A variety of analytic tools and systems, that can be used to manage big data as well, are available for the analysts to use. What guides the selection of the analytical tool is the objective of the analytics process. As some tools may be bound to specific analytic action and others may provide alternative results of the same assignment, it is important to know the differences between the tools and systems. Additionally, the data source affects the physical resources and the infrastructure they constitute. If no such restrictions are present and the data is already accessible, the analyst may choose the preferred analytic tool.

> "Suitable tools, where cloud-services play a major role, have an immense effect in maintaining effective data analytics. Through cloud-based servers we have a tool, which we can use to manage data and pose questions [to the data], that also scales well to managing even enormous sets of data." (VP, Tech & Development, Data Refinery)

Organisational capital resources

Regarding the organisational capital resources, three categories were identifiable from the research data. These categories were company-wide data processes, organisational culture and the process of converting insights into actions. The case company unarguably offers a data-driven environment and gladly, the management levels of both Data Refinery and Aller Media encourage and aim to maintain a data-driven culture. In Data Refinery, the company-wide processes are designed to focus on generating learning and self-development through efficient data analytics activities. As for the Aller Media's data unit or data team, they have established an operation model that is based on sharing information and knowledge between the team members. These kinds of established processes enhance data-driven culture that also generates learning within the employees. Having top-management level that encourages learning and efficient data operations ensures sustainable and successful data analytics practices.

> “Our unit has an operation model that is specifically based on sharing information and our objective is that after each sprint (14 days) we have learned something that hastens or facilitates the work executed during the next sprint. This is our guiding principle. In our unit the sprint is agreed to be 2 weeks period, during which we aim to set clear objectives for each team member, plan the results and focus on achieving them during the sprint.” (Lead Data Scientist, Aller Media)

Having clear processes generates also efficient workflow. When fundamental constructs of the analytic processes can be transformed into automatic codes, it allows the analyst to focus on more demanding analytical matters and on managing of the overall function. As big data is constantly altering, it is important to release resources to focus on the more demanding analytical assignments. By creating and establishing processes encompassing big data that for example update alongside with the data, the analysts can focus on identifying relevant up to date insights from data without having to focus on the data managing process itself. Additionally, by having established clear and consistent processes for development and performance-evaluation, the successful and sustainable data analytics practises are maintained.

Both, Aller Media’s and Data Refinery’s, teams are small in size and therefore easily managed. The organisational culture as well as the team dynamics encourage open communication inside the team as well as between the two teams. This situation also enhances learning and encourages the employees to proactively and without any restrictions to investigate and operate. Having employees with great competencies and expertise, it is easy to communicate the insights discovered from the data or from own work to the team. Through open communication the entire team can participate and offer different perspectives to the matter which will result in offering more valuable services to the customer and enhancing the ways of operating. Open communication within the team will also generate necessary and valuable insights for the proactive analytical work. Even though, in some situations the resources are not sufficient to execute the proactive analytical assignments and to investigate the hypotheses, it is still encouraged by the company and the hypotheses are mutually created by the entire team which has a positive impact on the culture.

Another feature that depicts an efficient organisational culture is the way in which insights are turned into actions. Open communication is an antecedent in the transformation process of turning insight into actions, as according to the research data, most of the transformation processes begin by open communication. In its simplicity, the process begins by identifying the insight from the data or by demonstrating the hypothesis, which then leads to appropriate actions inside the case company. The transformation process requires team effort, as many of the ideas that direct the analytical work, from where the insight is discovered, are created mutually either inside the team or with customers. The discovered insight is communicated to the customer who then decides whether to turn that insight into practical end-result action or not. Nevertheless, it is constant communication and proactive sharing of ideas inside the team on how to increase value of the insight or to develop own performance to offer better service for the customers.

> "When the insight is discovered from the data, when the established hypothesis is demonstrated, it then leads to appropriate actions. Naturally the process and the procedures are case-related, and the forming of the hypothesis also has its own impact, yet it all begins by discovering the insight and demonstrating the hypothesis" (VP, Tech & Development, Data Refinery)

4.2 Big data and knowledge management

Creation of knowledge

Knowledge creation in the case company is the beginning of knowledge management processes that is either initiated by internal ideas or through external insights. In cases where the process is initiated internally, it begins by open communication and proactive ideation on how to add value to the current services or products. That insight is then guiding the knowledge creation process. Upon the insight, the analyst determines the tools and systems as well as the arguments which to use for harvesting the new knowledge out of the data. When the process is initiated externally by or with the customer, the process, questions and objectives are directed by the customer's need. In these situations, the team is encouraged to proactively identify and analyse matters from the data that would add value to the service or offering. One example of proactive analysis of big data and knowledge creation is as the case company analyses big data flowing in from the corporation's website and identifies emerging trends regarding service usage from there.

> “The analysis process itself usually begins by having a hypothesis or a problem, for which we try to find a solution. Basically, it is proving the hypothesis to be true or false. Here the analytic process of big data does not differ notably from the traditional data analytics. Naturally, we can use different algorithms to search differences in the data or to identify phenomena, but it is a burdensome method to begin the data operations with.” (VP, Tech & Development, Data Refinery)

The process encompassing the knowledge creation in the case company begins by, according to the research data, clarifying the data and identifying the possible new insights from the data. Since the data can be quite scattered, it is important to convert it into a form that is easily understandable. Ever more so, when the results and new insights are communicated to the customers or people who do not possibly have the necessary expertise to comprehend raw data. When these new insights are discovered the analyst discusses with the rest of team to determine the value of the discovered new knowledge. Especially in situations where the analytic work is not initiated by the customer’s needs, it is important to identify to what matter does the new insight contribute to and what are benefits it offers when the insight is turned into action.

Furthermore, the process of knowledge creation is supported by routines. Routines guide the analytic processes in different service offerings. Additionally, the objective is to integrate the new discovered insights into the routines and processes to improve the operations. Routines guide the reporting and evaluation processes where for example a common report or a code template is used to illustrate and summarise the overall project performance. That ensures consistent and clear reporting but also provides clarity for the employees in their work. Furthermore, currently there are processes in the case company to create improved documentations of the workflow. The aim is to create clear and consistent documentation that illustrates the routines appearing in the analytic processes. The documentation is accessible by the entire team and they can follow the processes step-by-step during each project. Not only does this improve the transparency of the operations but also helps in analysing and developing the performance as well as in creating and sharing knowledge inside the team.

> “We have a clear report template which is created in collaboration with the customer that we use when reporting direct mail campaigns. If some new ideas appear, we add them to there, frequency-tables or other similar matters for

> example, but for me it is a very easy and a clear routine. It is a code template where we can manually add some slight alterations, but in general that is one routine-like element in this [analytics] process, that we use the same report template." (Data Analyst, Aller Media)

Storage of knowledge

In regard of knowledge storage, two themes were identifiable from the research data. Methods for storing knowledge and the systems and tools used for knowledge storing or as knowledge repositories. Since the case company handles consumer data, it must take into consideration the GDPR legislation. GDPR determines the data that can be collected but also restricts the data that can be stored inside ad utilised by organisations. Owing to this, the case company has established impeccable processes for their data management that are in-line with the GDPR regulations. Therefore, they secure, encrypt and systematically destroy data, to ensure secure data analytics operations.

> "We make the data unidentifiable, so it does not compromise the privacy of any individual, but we still can use and utilise the data to create target groups. Target groups never provide identifiable information about an individual, rather they provide insights about groups of individuals, to which [the customer] can target appropriate marketing or advertisement. Through big data, open source databases and enrichments we can provide these insights." (Project Manager, Data Refinery)

The methods of storing knowledge vary depending on the nature of the new knowledge or insight and on the means of usage of that knowledge. Common emerging insights or new knowledge can be quickly stored and documented into easily understandable and clear form, for example to PowerPoint, where both visual and text content can be displayed. Enrichments and calculated estimates, that is insights that are drawn from the data that are meant to be utilised in different contexts to increase value, can be documented and stored in and applied to the appropriate, context related database. In other cases, the new knowledge and insights are documented into reports that are constantly updated along with the on-going process. Therefore, the means and the place of knowledge storing varies and it is context and usage related. To maintain efficient knowledge storing processes, the case company has increased focus on improving its documentation practices. The aim is to document the entire processes thoroughly and precisely, not only regarding the insights but also the technicalities such as the used commands, codes and algorithms and working methods, hence focusing on storing both tacit and explicit knowledge.

> "New knowledge is stored, in cases of enrichments, into the original database in the form of new features or columns. Basically, new content is added to the rows [of data]. If necessary, we create new database models. Processes and other [technicalities] can be documented for later use as AdScripts, PythonScripts or as functions." (Junior Data Analyst, Data Refinery)

Sharing of knowledge

To analyse the knowledge sharing activities of the case company the following categories that emerged from the research data are inspected: internal and external knowledge sharing tendencies, culture and environment as well as the tools and systems that are used for the knowledge sharing activities. In this situation, internal teams refer to each unit, that is the Data Refinery's own unit or Aller Media's data unit. Therefore, both units are external groups for each other, although they operate in the same premises, and naturally each team's own customers form another external recipient for knowledge sharing activities.

In both units, the internal knowledge sharing, including both explicit and tacit, is mainly executed through social interaction and open communication. Both teams consist of few employees which provokes active communication inside the team. As stated, both teams have integrated processes to guide the efficient knowledge sharing and managing as well as development. Furthermore, constant knowledge sharing is maintained through internal weekly meetings where performance is analysed for initiating development. Naturally, the weekly meetings are opportunities for the employees to openly express their own ideas and insights. Both units mentioned each other's as the most important and closest external group in regard of knowledge sharing. As the two groups work within the same field and with similar matters, they do have a lot in common and similar expertise and knowledge reside in the units. Therefore, there is active sharing of tacit and explicit knowledge and collaboration between the two units. Nevertheless, the collaboration and the knowledge sharing between the units is not, according to the research data, as profound as desired. As the employees in both units are quite occupied with other pressing matters it can be the cause hindering the proactive knowledge sharing between the units. Fortunately, there are plans to enhance and to improve the collaboration and knowledge sharing of these two units. By establishing consistent knowledge sharing processes, like the two units have internally, to cover also knowledge sharing between the two units would likely generate rewarding results and improve overall performance. Such processes could also be extended to the other subsidiaries and units inside Aller Media Finland to create an organisation-wide pool of knowledge sharing and expertise.

> “A small team forces to share information and know-how quite lot, no walls or barriers exist to compromise our internal communication. In regard of learning, the most important stakeholder inside the office is Aller Media’s data analytics unit that we collaborate frequently with. The work and assignments between the teams are very similar and discussion about tools as well as sharing of information and know-how is constant and very important for both parties.” (VP, Tech & Development, Data Refinery)

As for the tools and systems, in cases of sharing documented work or explicit knowledge, interactive, cloud-based tools that can depict both visual and written content are used. The tools can be used to present technicalities such as codes as well as results from the entire analytic process. Furthermore, the entire process can be easily and clearly visualised with the tool as well as shared and saved for further use. Furthermore, in every situation it is ensured, that each employee has access to the necessary information and knowledge and that is clearly presented. In situations where the collaboration is between the two units or with other external partners, mutual operational environments are created to ensure efficient and transparent flow of work. Thus, each employee that is part of the project can easily access the necessary information and knowledge through the tools and systems.

> “In analytic work cases we have created with different teams, and for example with Data Refinery, a common operational environment, where the colleagues to whom the work relates to, are. Especially in cases of big data analytics. The environment is shared and secured in such way that the information is accessible to the colleagues when needed.” (Lead Data Scientist, Aller Media)

Application of knowledge

Knowledge application transpires through two different activities, group problem solving for the customer and through development of own operations. Both units' objective is to provide quality data-based services for their customers and therefore all of their operations are focused on achieving that objective. Data analytics are a mean of achieving and increasing the knowledge base of the units, but the main focus is to use that achieved knowledge to improve the customer's operations by providing them data-based services and products. The achieved knowledge is managed and handled by the entire team and also the entire team participates in providing the services for the customer or in solving the customer's problems through the services and products. The customers can use the received knowledge to enhance their decision-making towards data-based decision-making activities and hence to improve their performance and operations (marketing and advertisement) or to solve possible occurring problems.

> "Of course, we always have to think what we are going to do and aim for with the discovered knowledge. Our aim is to create new products that we can sell to our customers, hence our development activities always have the same goal. When we find new insights from the data, our aim is to share those insights with our customers." (Junior Data Analyst, Data Refinery)

Naturally, the achieved new knowledge is also used in the units' own decision-making and for development purposes. According to the research data, these activities are not the priority since the main focus is on serving the customers. Which is understandable since the knowledge that is harvested out of the data, is directed towards the customer work and hence, would not necessarily provide notable value to the two units. Nevertheless, if relevant new knowledge and insights for the case company would emerge, their value would be inspected and if necessary, used for improving own decision-making practices or overall operations. Most of the knowledge that is used to improve own operations is through tacit knowledge. Best practices for executing efficient work and analytics is shared both internally within the teams and externally between the units and it is actively used for development purposes. Additionally, feedback from the customers is used to develop and improve own operations.

> "The knowledge is used to both, for developing own operations and for decision-making. In regard of decision-making I'm not sure, probably it [new knowledge]

> is actually transferred to improve the customer's decision-making rather than ours. But especially in our development processes it [new knowledge] is mostly used. From the development processes we gain a lot, by working and doing we find the best ways to conduct our work, we find the right products from which we evaluate the ones we are going to offer [to our customers]." (Project Manager, Data Refinery)

4.3 Antecedents of big data analytics

In this section the research questions are analysed and findings from the research data regarding the questions are illustrated. Some of the matters are already analysed in the previous sections, but the focus of this chapter is to draw up the findings and how they contribute to the research questions of this study.

RQ 1: What resources does the exploitation of big data analytics require?

Primarily, what could be identified as fundamental constructs for conducting successful big data analytics were human and physical capital resources. Big data analytics cannot be executed without the necessary advanced cloud-based tools and systems. Nevertheless, the tools and systems themselves cannot, at least not yet, operate or execute successful big data analytics without humans. Ever more so when the analytic activities are used in the business-sense. to improve the customer company's capabilities and to bring individual, specific value to them. Humans that are competent in regard of data analytics are needed to communicate with the customers, to understand the customers' needs and desires and to direct the analytical work to fulfil those needs. The employees providing the analytics services must have a thorough understanding of the possibilities and opportunities big data analytics provide while also understanding what the requirements are for conducting and providing efficient services based on big data analytics. As the data is becoming constantly more challenging, the importance of communication between the service provider and the customer is essential. Without knowing the reason nor objectives for big data analytics, the result of big data analytics is likely to be useless. Big data is so massive and distorted that without knowing or understanding the core reason and motive for the analytic activities, the correct, relevant matters cannot be discovered from the data. Since the questions, codes and algorithms for the data are based on the primary reason and motives, it is necessary for both the customer and for the analytic service provider to have a clear and concise understanding of these matters before commencing the analytical work.

> “Successful data analytics require that the primary conversation with the company that needs analytic services, can answer the question “why”. Then the company that executes the analytical work knows what needs to be done, knows what happens after the entire process and where does it lead to. Before thriving after these things, the customer should know the reason, for example finding solution to a certain problem or future challenge, why they are asking the analytical services from us.” (Account Director, Data Refinery)

RQ 2: What knowledge management processes support the exploitation of big data analytics?

The most notable impact big data analytics have is through the changes it requires and generates to both organisational resources and knowledge management processes. As stated, big data analytics require advanced physical capital resources as opposed to the traditional data analytic tools and systems. Furthermore, big data analytics require advanced knowledge and expertise from the human capital resources and employees interacting with the data in order to utilise it efficiently. Additionally, to maintain prolific and efficient processes of big data analytics the organisational capital resources are to be altered in regard of big data analytics. In its simplicity, it means encouraging and emphasising data-driven culture in the working environment. As of the knowledge management processes, the knowledge itself, both tacit and explicit, are becoming more challenging in regard of big data. The knowledge base of the company should be constantly growing and most likely is, which is why the entire knowledge management process should be well maintained and consistent. All of these changes are possible and easily executed if the employees have enough will and are motivated. The top-level of the company should be motivated to encourage and enhance the data-driven culture and to offer an environment where successful data analytics activities and knowledge residing from them can flourish. As for the other employees, they must be motivated to proactively communicate, learn new tools, systems and methods required for big data analytics and having the will to develop themselves further in regard of big data analytics.

> “The biggest challenge regarding big data analytics is that it always in some ways changes or at least it should change the company’s operations. The desire for change is weak, most companies just want to continue to work as they do and by making only slight changes. Analytics tend to be seen as a huge investment and in reality, it necessary isn’t.” (Account Director, Data Refinery)

In order to implement sustainable big data analytics practices, the company must understand its opponent – in this case big data. Therefore, the requirements, challenges and possibilities must be acknowledged in order to implement sustainable big data analytics activities. Furthermore, by understanding the nature of big data, all of its dimensions and characteristics, thoroughly the company can answer the current requirements but also predict possible future challenges and prepare for those. By understanding all of these important matters, the company can build processes upon big data analytics rather than struggling to meet current requirements of big data.

> "We have started to modify our processes and business operations to be functional and effective regarding handling of huge data masses. How can we scale our operations to manage even bigger data masses? How can we store, maintain and design processes upon these massive data sets?" (VP, Tech & Development, Data Refinery)

Having established these processes upon big data analytics, the company can focus on maintaining and developing them and how to increase the value and competitive advantages big data analytics provide. Big data analytics enable the analysis of new sources of information. These new sources of information could not be analysed before due to lack of necessary tools and systems but also since many of the new sources that generate big data have emerged recently. Many of big data sources create current, constantly altering real-time data. This rapid change of data makes it valuable only during a short period of time. Therefore, the companies should focus on establishing processes that can capture this real-time data, draw insights from there and hence, allow the company to immediately react to the insights generated from the data. As the pace regarding the generated data grows it also shortens the time the company has to operate with the data and the insights. In this situation, a working environment where all organisational resources and knowledge management processes are orchestrated effectively upon big data analytics is essential. If the company lacks support in some of these fundamental constructs discussed in these chapters, it simply cannot maintain sustainable big data analytics practices that could effectively answer the requirements of constantly altering, massive and complex big data analytics. Nevertheless, support, development and motivation to succeed and to maintain these sustainable processes are all initiated by humans. Therefore, in order to implement sustainable big data analytics practices the company must, in its fundamental level, have competent and motivated human resources.

“By holding on to the competent people we maintain effective data analytics. There is no short-cut to this, because it is the people who do the analytical work, who teach the machines and they are the ones who can develop the operations further.” (Account Director, Data Refinery)

5 DISCUSSION

In this chapter the contributions and managerial implications of this research as well as the limitations and suggestions for future research are presented. The discussion consists of analysis of the research data as well as of the findings and their relevancy. The theoretical contributions this research provides are presented as well as managerial implications. Lastly, the limitations of this research and suggestions for future research are presented.

5.1 Theoretical contributions

When analysing the theoretical contributions this research provides, the findings regarding each discipline are well aligned with previous research. Since the case company provided a quite well-established and a prolific data environment with its successful orchestration of resources as well as knowledge management, it can be noted that the findings of this study are compatible with the challenges mentioned by other researchers regarding big data analytics. Only, in the case company these challenges are already handled but their importance could still be identified. For example, as stated by Berinato (2019) one major challenge hindering effective big data analytics processes is caused by lack of communication. Furthermore, Davenport et al. (2012) also inspected the meaning of communication in their research as the data scientists who work closely with big data must possess the ability to communicate effectively with decision-makers and hence, to ensure effective and fluent distribution of knowledge gained from big data analytics. The case company did not suffer from lack of communication nor did they lack any communications abilities, but open communication that was supported and maintained could be identified as one of the fundamental constructs in enabling and maintaining successful big data analytics. Therefore, the findings of this research emphasise the already identified significance of internal and external communication in successful and sustainable big data analytics.

The research data can be seemed to challenge the previous theoretical implications in four distinctive ways. Primarily, as could be identified from the research data, it is not only important for the data analyst or scientist to possess good communications skills (Berinato 2019; Davenport et al. 2012) but also to have a business mind-set and to be curious. Naturally, it can be argued whether business mind-set could be included as part of good communication skills. Nevertheless, as was indicated by the research data, the business mind-set in this sense, has a deeper meaning than just consisting of good communication practices. The data analysts and scientist should understand the business' perspective of analytics and identify business-related opportunities from the data analytics activities as opposed to focusing solely on the analytics practices.

Secondly, previous research state different responsibilities namely for the organisation's management level. The organisation's managers are responsible for identifying, developing and arranging key resources efficiently, also in regard of resources needed for big data analytics, to create value (Fahy 2000). Additionally, resistance against big data analytics or change dwelling in and originating from the organisation's managerial level has a negative effect on the success of building sustainable and efficient big data analytics practices (Mikalef et al. 2019b). The case company's managerial level did not portray any resistance towards new and innovative ways of working with data rather they encouraged it entirely. However, what could be identified from the research data, which challenges the previous theoretical insights, was that the responsibilities are not solely directed to the management level but rather in an open and well-established environment the responsibilities are mutually shared. Each employee is responsible for developing their own work further but also contributes voluntarily to identify new and innovative insights, is motivated to develop the company further and to comprehensively provide and generate additional value.

Thirdly and most significantly, previous literature focus on either identifying problems and characteristics regarding organisational resources and big data analytics or knowledge management processes and big data analytics. One notable theoretical contribution this research provides is as it combines the literature of big data analytics, resource-based theory and knowledge management into one framework and study. Based on this comprehensive framework that provided a novel and a thorough perspective on the subject, important matters regarding big data analytics could be discovered that may have been unnoticed before. By studying these three constructs together, the features that emerged from the research data can be deemed as most essential constructs and antecedents of successful and sustainable big data analytics. The meaning and the role these findings provide was only accessible by constructing the research on the two different disciplinaries combined with big data. Many of the findings do align with the previous literature on the different disciplines. However, this research provided insights that challenge the previous literature and unarguably provides more in-depth insights about the subjects at hand altogether.

Lastly, there were matters that emerged throughout the research data as the most important constructs in supporting and enabling successful and sustainable big data analytics. These matters may have been presented in the already existing literature, but the significance is underestimated when compared to the findings of this research. For example, the importance of human capital resources, their competencies and abilities were often mentioned throughout the interviews. The previous theoretical literature also implicates that the technical abilities of the data analyst or scientist regarding the required advanced analytics tools and systems are a necessity for conducting successful big data analytics (for example Davenport et al. 2012 and De Mauro et al. 2014). As could be identified from the research data the importance of human capital resources does not limit only to the technical abilities but extends to the employee's cognitive capabilities and team-player competencies. Additionally, the importance of an environment that encourages and enhances open communication can be identified as the primary source for innovation, knowledge and naturally successful big data analytics that can be used to generate value and competitive advantages. This kind of an environment is not only facilitated by the managers of an organisation rather it requires commitment and contribution equally from each individual employee. Which again, emphasises the cognitive capabilities, motivation, curiosity and commitment as important antecedents of successful and sustainable big data analytics.

5.2 Managerial implications

The substantial findings this research provides can be utilised by organisations' managers and practitioners. By taking the matters illustrated in this research into consideration and adapting them into own operations, the organisation can most likely identify antecedents that support big data analytics and establish successful big data analytics practices.

There are few fundamental matters that should be discussed in the managerial level of an organisation that desires to commence or to improve big data analytics practices. Primarily, you need to understand big data and big data analytics comprehensively. How does the organisation's resources and knowledge management practices support or enable successful big data analytics and what are the organisation's objectives for big data analytics operations? The extent to which the organisational resources and knowledge management processes must be adjusted for big data analytics is naturally affected by the organisation's operations regarding big data. For example, in the case company, they have all the necessary competencies, knowledge and capabilities to execute successful big data analytics, even though not all of the cases they work with their customers consist of big data. Therefore, their focus is to innovatively and with advanced methods to use, manage and capitalise on the amount of data available. Thus, to manage also smaller sets of data in a big data -kind of way. Therefore, relevant objectives and ways of operation can only be established when the possible opportunities provided by big data analytics are understood as well as how, and in which ways the opportunities could be leveraged to improve the organisation’s own business and what is the value they contribute. By understanding big data analytics thoroughly, the company can execute successful management of organisational resources and knowledge management practices and as a result, establish processes, principles and methods encompassing and supporting big data analytics.

Furthermore, the organisation’s managers must acknowledge the importance of employees and certain features of the working environment. As stated, in optimal situation the responsibilities for developing own and team-wide operations further would be divided equally between each individual. Nevertheless, initiating this kind of voluntary responsibility sharing requires actions from the organisation’s managerial level. Establishing constant opportunities and situations where the employees can openly share their ideas, knowledge and communicate coequally without restrictions is something that is usually primarily enabled by the organisation’s managerial level. Providing the time and situations for open communication and participating in these situations is meaningful when aiming for an efficient working environment that acts as an antecedent for successful exploitation and execution of big data analytics.

As for the human resources, the managers must understand that knowledge, which is extracted from big data analytics and which is the source of innovation and value creation for the organisation, resides in the minds of humans. Furthermore, it is the humans that operate, develop and manage and interact with the big data and the data analytics tools, systems and operations. Therefore, in order to implement successful and sustainable big data analytics practices, the managers’ must take into consideration not only the organisational resources but also the knowledge management practices. As the research data indicates and presented by this research, by efficiently orchestrating both organisational resources and knowledge management processes that support big data, the company can maintain sustainable big data analytics practices and answer the requirements of constantly altering, massive and complex data. Valuing and respecting the employees, who play a major role in big data analytics, and communicating it to them is one simple yet important factor how the organisation’s managers can initiate the creation of a healthy working environment. Providing the employees opportunities to learn, develop and innovate through encouraging them to proactively test, seek and discover new insights will result in innovative, development-oriented, committed and motivated human resources that can successfully execute big data analytics.

By being competent in executing big data analytics activities, the case company can enhance its competitiveness. At the same time, its data analytics activities are used to enhance its customers' competitiveness and to provide them a competitive advantage and value. In this sense, the value generated by big data analytics in this case is twofold. The case company does not primarily use the insights derived from big data analytics for discovering and creating opportunities through which the case company's dynamic capabilities could be improved. The case company uses big data analytics to discover and to create opportunities through which its customers can improve their dynamic capabilities. The ways in which the case company can enhance its dynamic capabilities is by maintaining its expertise and innovative competencies regarding big data analytics. This is also something that the managerial level of organisations' must take into consideration. To maintain competencies regarding big data analytics, the pace and scope of change regarding data must be acknowledged. Already some of the characteristics of big data can be identified from the "traditional" data as well. Many of the sources generating data nowadays have emerged during the recent years, robot vacuums and smart televisions for example, and thus it allows organisations to leverage data that was not available earlier due to advancements in technology. Additionally, the data that has been available and has been stored over decades, has progressively grown into massive databases. The managerial level of organisations must analyse the meaning of different and new sources of data and their value for the company in order to maintain its competencies and improve competitiveness. Therefore, the company must constantly develop its processes to meet the requirements of the current data while also predicting and preparing for the future challenges – the company must focus on being the best and most innovative and agile data analytics service provider or operator in its own market.

To conclude, the managers of an organisation must comprehend important constructs regarding successful big data analytics. Understanding big data analytics comprehensively in order to manage the organisational resources and knowledge management processes accordingly and to prepare for future challenges. Additionally, focusing on establishing a working environment where open communication and therefore learning and innovation is encouraged and facilitated as well as on providing its human capital resources an environment where they can express themselves, learn and develop themselves further to bring additional value to the working environment.

5.3 Limitations and suggestions for future research

As the findings of this research are based on a case company study, although combining the insights of two separate units, it still implies findings based on one company. Furthermore, as this study does not focus on the technicalities of big data analytics, which tend to be quite universal, rather to organisational resources and knowledge management which both may possess different characteristics in different companies, environments, cultures, personnel sizes and industries. Therefore, findings and contributions based on this research, although valid and reliable in regard of the case company, are not likely to apply universally to all companies.

Since this research studied big data analytics in regard of a case company that primarily offers data-based services to its customers, it would be interesting to study whether the findings presented in this research are similar to cases where the company utilises data analytics solely for own purposes. Furthermore, as the case company’s two data units were rather small in size, future studies could investigate organisations where the personnel are bigger and hence provide insights how these findings relate to such environments. Additionally, since this research focused on interviewing individuals that constantly work, on some level, with data, future research could also include other employees within the same company that do not necessarily work with data to provide a wider perspective on the subject.

Additionally, as it could be identified from the research data, even though the case company possesses all the capabilities and competencies to execute big data analytics many of the customer cases they work with do not necessarily contain big data. Therefore, the reasons hindering exploitation and utilisation of big data and big data analytics in organisations could be studied. The reasons may be due to resistance against change residing in organisations, either managerial level or employees, but the motives initiating this resistance could be a relevant topic for research. Furthermore, the reasons can be caused by impartial data sources which would then require the analysis and thorough studying of the data production’s condition.

REFERENCES

Abbasi, A. Sarker, S. & Chiang, R. H. L. (2016). Big data research in information systems: Toward an inclusive research agenda. Journal of the Association for Information Systems, 17(2). [www-document] [Accessed 14th June 2019] Available: https://ezproxy.cc.lut.fi/docview/1773772634?accountid=27292

Alavi, M. & Leidner, D. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly, 25*(1), 107-136. [www-document] [Accessed 24th May 2019] Available: https://ezproxy.cc.lut.fi/docview/218141939?accountid=27292

Aller Media Finland (2019a). Online-Mainonta. [www-document] [Accessed 24th August 2019] Available: https://www.aller.fi/media-2/onlinemainonta/

Aller Media Finland (2019b) Seiska on Suomen tavoittavin aikakausimedia. March 1, 2019. [www-document] [Accessed 23rd August 2019] Available: https://www.aller.fi/seiska-on-suomen-tavoittavin-aikakausmedia/

Aller Media Finland (2019c) Tietoa Allerista. [www-document] [Accessed 24th August 2019] Available: https://www.aller.fi/tietoa-allerista/

Argote, L. & Ingram, P. (2000). Knowledge Transfer: A Basis for Competitive Advantage in Firms. Organizational Behavior and Human Decision Processes, 82(1), pp. 150-169. 8[www-document] [Accessed 17th May 2019] Available: doi:10.1006/obhd.2000.2893

Balducci, B. & Marinova, D. (2018). Unstructured data in marketing. Journal of the Academy of Marketing Science, 46(4), pp. 557-590. [www-document] [Accessed 3rd March 2019] Available: http://dx.doi.org.ezproxy.cc.lut.fi/10.1007/s11747-018-0581-x

Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99. [www document] [Accessed 5th March 2019] Available: https://ezproxy.cc.lut.fi/docview/215258436?accountid=27292

Beach, C. & Schiefelbein, W. (2014). Unstructured data. Journal of Accountancy, 217(1), pp. 46-51,11. [www-document] [Accessed 15th May 2019] Available: https://ezproxy.cc.lut.fi/docview/1475080280?accountid=27292

Cepeda-Carrion, I. Martelo-Landroguez, S. Leal-Rodríguez, A. & Leal-Millán, A. (2017). Critical processes of knowledge management: An approach toward the creation of customer value. European research on management and business economics, 23(1), pp. 1-7. [www-document] [Accessed 15th May 2019] Available: https://doi.org/10.1016/j.iedeen.2016.03.001.

Chen, H. Chiang, R. & Storey, V. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly, 36(4), 1165–1188. [www-document] [Accessed 27th May 2019] Available: https://doi-org.ezproxy.cc.lut.fi/10.2307/41703503

Cohen, W. & Levinthal, D. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly, 35(1), pp. 128-52. [www-document] [Accessed 17th May 2019] Available: https://doi-org.ezproxy.cc.lut.fi/10.2307/2393553

Côrte-Real, N. Ruivo, P. Oliveira, T. & Popovič, A. (2019). Unlocking the drivers of big data analytics value in firms. Journal of Business Research, 97, pp. 160-173. Available: https://doi.org/10.1016/j.jbusres.2018.12.072.

Côrte-Real, N. Oliveira, T. & Ruivo, P. (2017). Assessing business value of Big Data Analytics in European firms. Journal of Business Research, 70, pp. 379-390. [www-document] [Accessed 27th May 2019] Available: https://doi.org/10.1016/j.jbusres.2016.08.011.

Davenport, T. H. Barth, P. & Bean, R. (2012). How big data is different. MIT Sloan Management Review, 54(1), 43-46

Dayan, R. & Evans, S. (2006). KM your way to CMMI. Journal of Knowledge Management, 10(1), 69-80. [www-document] [Accessed 17th May 2019] Available: http://dx.doi.org.ezproxy.cc.lut.fi/10.1108/13673270610650111

Data Refinery (2019a) About us. [www-document] [Accessed 23rd August 2019] Available: https://www.datarefinery.io/en/about-us/

Data Refinery (2019b) Customer insight and analytics – Enrichments. [www-document] [Accessed 23rd August 2019] Available: https://www.datarefinery.io/en/customer-insight-and-analytics/enrichments/

Data Refinery (2019c) Target groups and campaigns. [www-document] [Accessed 23rd August 2019] Available: https://www.datarefinery.io/en/target-groups-and-campaigns/

De Mauro, A. Greco, M. & Grimaldi, M. (2014). What is Big Data? A Consensual Definition and a Review of Key Research Topics. [www-document] [Accessed 13th June 2019] Available: https://www.researchgate.net/publication/265775800_What_is_Big_Data_A_Consensual_Definition_and_a_Review_of_Key_Research_Topics

Durst, S. & Edvardsson, I. (2012) "Knowledge management in SMEs: a literature review", Journal of Knowledge Management, Vol. 16 Issue: 6, pp. 879-903. [www-document] [Accessed 8th May 2019] Available: https://doi.org/10.1108/13673271211276173

Erevelles, S. Fukawa, N. & Swayne, L. (2016). Big data consumer analytics and the transformation of marketing. Journal of Business Research, 69(2), 897.

Eskola, J. & Suoranta, J. (1998). *Johdatus laadulliseen tutkimukseen.* Tampere: Vastapaino.

Fahy, J. (2000). The resource-based view of the firm: Some stumbling-blocks on the road to understanding sustainable competitive advantage. Journal of European Industrial Training, 24(2), 94-104. [www-document] [Accessed 10th May 2019] Available: https://ezproxy.cc.lut.fi/docview/215396502?accountid=27292

Ferraris, A. Mazzoleni, A. Devalle, A. & Couturier, J. (2018). Big data analytics capabilities and knowledge management: Impact on firm performance. Management Decision. [www-document] [Accessed 19th June 2019] Available: https://doi-org.ezproxy.cc.lut.fi/10.1108/MD-07-2018-0825

Fosso, S. Akter, S. Edwards, A. Chopin, G. & Gnanzou, D. (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics, 165, pp. 234-246. [www-document] [Accessed 27th May 2019] Available: https://doi.org/10.1016/j.ijpe.2014.12.031.

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

Gao, T. Chai, Y. & Liu, Y. (2018) "A review of knowledge management about theoreticalconception and designing approaches", International Journal of Crowd Science, Vol. 2 Issue: 1, pp. 42-51 [www-document] [Accessed 8th May 2019] Available: https://doi.org/10.1108/IJCS-08-2017-0023

Gonzalez, R. & Martins, M. (2017). Knowledge Management Process: a theoretical-conceptual research. Gestão & Produção, 24(2), 248-265. Epub January 26, 2017. [www-document] [Accessed 8th May 2019] Available: https://dx.doi.org/10.1590/0104-530x0893-15

Grant, R. (1996a). Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration. *Organization Science, 7*(4), pp. 375-387.

Grant, R. (1996b). TOWARD A KNOWLEDGE-BASED THEORY OF THE FIRM. Strategic Management Journal (1986-1998), 17, 109. [www-document] [Accessed 24th May 2019] Available: https://ezproxy.cc.lut.fi/docview/231085587?accountid=27292

Gupta, M. & George, J. (2016). Toward the development of a big data analytics capability. Information & Management, 53(8), pp. 1049-1064. [www-document] [Accesses 9th June] Available: https://doi.org/10.1016/j.im.2016.07.004.

Labrinidis, A. & Jagadish, H. (2012). Challenges and opportunities with big data. Proceedings of the VLDB Endowment, 5(12), pp. 2032-2033. [www-document] [Accessed 27th May 2019] Available: doi:10.14778/2367502.2367572

Laney, D. (2001). 3-D data management: Controlling data volume, velocity and variety. Application Delivery Strategies by META Group Inc. [www-document] [Accessed 5th March 2019] Available: https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Marr, B. (2018). How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read. Forbes Innovation. [www-document] [Accessed 13th March 2019] Available: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#241a48c860ba

Merendino, A. Dibb, S. Meadows, M. Quinn, L. Wilson, D. Simkin, L. & Canhoto, A. (2018). Big data, big decisions: The impact of big data on board level decision-making. Journal of Business Research, 93, pp. 67-78. [www-document] [Accessed 13th March 2019] Available at: https://doi.org/10.1016/j.jbusres.2018.08.029.

Mikalef, P. Boura, M. Lekakos, G. & Krogstie, J. (2019a). Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. British Journal of Management, 30(2), pp. 272-298. [www-document] [Accessed 10th June 2019] Available: https://doi-org.ezproxy.cc.lut.fi/10.1111/1467-8551.12343

Mikalef, P. Boura, M. Lekakos, G. & Krogstie, J. (2019b). Big data analytics and firm performance: Findings from a mixed-method approach. Journal of Business Research, Volume 98, May 2019, pp. 261-276. [www-document] [Accessed 12th June 2019] Available: https://doi.org/10.1016/j.jbusres.2019.01.044

Morgan, N. (2012). Marketing and business performance. Journal of the Academy of Marketing Science, 40(1), 102-119. [www-document] [Accessed 15th May 2019] Available: http://dx.doi.org.ezproxy.cc.lut.fi/10.1007/s11747-011-0279-9

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. Organization Science, 5(1), 14–37. [www-document] [Accessed 17th May 2019] Available: https://doi-org.ezproxy.cc.lut.fi/10.1287/orsc.5.1.14

Nonaka, I. & Konno, N. (1998). The concept of "ba": Building a foundation for knowledge creation. California Management Review, 40(3), 40-54. [www-document] [Accessed 17th May 2019] Available: https://ezproxy.cc.lut.fi/docview/216149525?accountid=27292

Hilbert, M. (2016). Big Data for Development: A Review of Promises and Challenges. Development Policy Review, 34(1), pp. 135-174. [www-document] [Accessed 12th June 2019] Available: https://doi-org.ezproxy.cc.lut.fi/10.1111/dpr.12142

Hirsjärvi, S. Remes, P. & Sajavaara, P. (2007). *Tutki ja kirjoita* (13.-14. chapters updated edition). Helsinki: Tammi.

Hota, C. Upadhyaya, S. & Al-karaki, J. (2015) Advances in secure knowledge management in the big data era. Information Systems Frontiers, 17(5), 983-986. [www-document] [Accessed 9th March 2019] Available: http://dx.doi.org.ezproxy.cc.lut.fi/10.1007/s10796-015-9593-y

Intezari, A. & Gressel, S. (2017). Information and reformation in KM systems: Big data and strategic decision-making. Journal of Knowledge Management, 21(1), 71-91. [www-document] [Accessed 3rd March 2019] Available: http://dx.doi.org.ezproxy.cc.lut.fi/10.1108/JKM-07-2015-0293

Jasimuddin, S. Klein, J. & Connell, C. (2005). The paradox of using tacit and explicit knowledge: Strategies to face dilemmas. Management Decision, 43(1), 102-112. [www-document] [Accessed 17th May 2019] Available: https://ezproxy.cc.lut.fi/docview/212085587?accountid=27292

Kozlenkova, I. Samaha, S. & Palmatier, R. (2014). Resource-based theory in marketing. Journal of the Academy of Marketing Science, 42(1), 1-21. [www-document] [Accessed 9th May 2019] Available: http://dx.doi.org.ezproxy.cc.lut.fi/10.1007/s11747-013-0336-7

Pauleen, D. J., & Wang, W. Y. C. (2017). Does big data mean big knowledge? KM perspectives on big data and analytics. Journal of Knowledge Management, 21(1), 1-6. [www-document] [Accessed 21st April 2019] Available: http://dx.doi.org.ezproxy.cc.lut.fi/10.1108/JKM-08-2016-0339

Rajaraman, V. (2016). Big data analytics. Resonance, 21(8), pp. 695-716. [www-document] [Accessed 12th June 2019] Available: https://doi-org.ezproxy.cc.lut.fi/10.1007/s12045-016-0376-7

Roberts, J. (2000). From know-how to show-how? questioning the role of information and communication technologies in knowledge transfer. Technology Analysis & Strategic Management, 12(4), 429-443. [www-document] [Accessed 17th May 2019] Available: https://ezproxy.cc.lut.fi/docview/226916149?accountid=27292

Scarpellini, S. Marín-Vinuesa, L. Portillo-Tarragona, P. & Moneva, J. (2018). Defining and measuring different dimensions of financial resources for business eco-innovation and the influence of the firms' capabilities. Journal of Cleaner Production, 204, pp. 258-269. [www-document] [Accessed 9th May 2019] Available: https://doi.org/10.1016/j.jclepro.2018.08.320.

Teece, D. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal, 28(13), pp. 1319-1350. [www-document] [Accessed 10th May 2019] Available: doi:10.1002/smj.640

Vidgen, R. Shaw, S. & Grant, D. (2017). Management challenges in creating value from business analytics. European Journal of Operational Research, 261(2), pp. 626-639. [www-document] [Accessed 14th June 2019] Available: https://doi.org/10.1016/j.ejor.2017.02.023.

Yin, R. (2014). Case study research: Design and methods (5th edition.). Los Angeles: SAGE.

Zollo, M. & Winter, S. (2002). Deliberate learning and the evolution of dynamic capabilities. Organization Science, 13(3), 339-351. [www-document] [Accessed 18th May 2019] Available: https://ezproxy.cc.lut.fi/docview/213833893?accountid=27292

APPENDIX

Appendix 1 – Interview questions

1. Describe briefly your job.

Big data analytics

2. In what ways does your work relate to data and analytics?
3. What is the role of (big) data in your business?
4. How and to what extent has big data changed your company's business activities and processes?
5. What benefits it provides to the company?
6. What are the notable challenges regarding the exploitation of big data?
7. What kind of data do you collect and from what sources?
8. What happens to the collected data?
9. How do you maintain effective data analytics?

Organisational resources

10. What kind of human resources and knowledge big data analytics require?
11. What kind of tools/systems are used to collect, store and manage the data?
12. How are insights from data transformed into action and implemented in your organisation?

Knowledge management

13. How do you analyse big data in your company? Describe the process.
14. What do you do with the new knowledge extracted from data analytics?
15. Where/how do you store the insights gained from analytics activities?
16. How do you handle/process the knowledge extracted from data?
17. How is the knowledge extracted from data used?

Data analytics

18. What do you feel is required/essential for conducting successful data analytics? -
19. What kind of challenges/difficulties have you encountered (recently) while working with data analytics?