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

Elina Uotinen

LEVERAGING BIG DATA ANALYTICS IN MARKETING: A MULTIPLE CASE STUDY IN THE MOBILE GAMING INDUSTRY

Examiners: Assistant Professor Joel Mero
Associate Professor Anssi Tarkiainen

ABSTRACT

Author	Elina Uotinen
Title	Leveraging big data analytics in marketing: A multiple case study in the mobile gaming industry
Faculty	School of Business and Management
Master's Program	International Marketing Management (MIMM)
Year	2020
Master's Thesis	Lappeenranta-Lahti University of Technology LUT 87 Pages, 3 Figures, 1 table, 1 appendix
Examiners	Assistant Professor Joel Mero & Associate Professor Anssi Tarkiainen
Keywords	big data, data-driven marketing, knowledge management, resource-based theory, mobile gaming industry

The purpose of this research is to explore how the novel phenomenon of utilizing big data analytics in marketing can be leveraged in the mobile gaming industry. While the use of data has undeniable potential and a massive influence on reshaping decision-making and marketing management overall, there exists a gap in research concerning the topic. This study aims to develop new theory and provide a deeper understanding of the issue by combining resource-based theory and knowledge management, as well as literature on big data analytics, to identify the most crucial organizational resources required in data-driven marketing, along with its knowledge management processes and practices.

The empirical study in this thesis is a qualitative multiple case study covering five different companies in the industry. Data is collected via individual semi-structured interviews of key persons. The findings of this study provide an original theoretical model illustrating the cyclical data-driven marketing optimization process in the mobile gaming industry, while also depicting the connected relationship of a firm's organizational resources and knowledge management processes. Furthermore, the findings emphasize the importance of a proper data-driven organizational culture.

TIIVISTELMÄ

Tekijä	Elina Uotinen
Tutkielman nimi	Datan hyödyntäminen markkinoinnissa: Monitapaustutkimus mobiilipelialalla
Tiedekunta	Kauppatieteellinen tiedekunta
Pääaine	International Marketing Management (MIMM)
Vuosi	2020
Pro Gradu -tutkielma	Lappeenrannan-Lahden teknillinen yliopisto LUT 87 sivua, 3 kuviota, 1 taulukko, 1 liite
Tarkastajat	Apulaisprofessori Joel Mero & Apulaisprofessori Anssi Tarkiainen
Avainsanat	big data, dataohjattu markkinointi, tietämyksenhallinta, resurssiperusteinen teoria, mobiilipeliala

Tämän tutkimuksen tarkoituksena on selvittää, miten big data -analytiikkaa voidaan hyödyntää markkinoinnissa mobiilipelialan näkökulmasta. Huolimatta datan käytön kiistattomasta potentiaalista ja massiivisesta vaikutuksesta päätöksentekoon ja markkinointiin ylipäätään, aiheeseen liittyvässä tutkimuksessa on havaittavissa aukko. Tämän tutkimuksen tavoitteena onkin kehittää uutta teoriaa ja tarjota syvempää ymmärrystä aiheesta yhdistämällä resurssiperusteisen teorian ja tietämyksenhallinnan tieteenhaarat data-analytiikan kirjallisuuden kanssa ja sitä kautta tunnistaa tärkeimmät organisaation resurssit sekä tietämyksenhallinnan prosessit ja käytännöt, joita dataohjattu markkinointi edellyttää.

Tämän pro gradu -tutkielman empiirinen tutkimus toteutettiin laadullisena monitapaustutkimuksena, joka kattaa viisi alan yritystä. Tutkimusaineisto kerätään avainhenkilöiden yksittäisillä puolistrukturoiduilla haastatteluilla. Tutkimuksen tulokset tarjoavat alkuperäisen teoreettisen mallin, joka kuvaa dataohjatun markkinoinnin syklistä optimointiprosessia mobiilipelialalla, ja myös samalla organisaation resurssien ja tietämyksenhallinnan prosessien kytkettyä suhdetta. Lisäksi tämä tutkimuksen tulokset korostavat sopivan dataohjatun organisaatiokulttuurin merkitystä yrityksessä.

ACKNOWLEDGEMENTS

Five years ago, I started my journey at LUT, and it is hard to believe that it now comes to an end. While I am quite happy that this thesis is now finished and I am onto something new, I suspect that I will dearly miss the times I had studying here.

I would like to express gratitude to the people who participated in the interviews for this study and shared their invaluable knowledge, and therefore made the empirical research of this thesis possible.

In addition, I would like to thank my thesis supervisor Joel Mero for his valuable guidance and tips throughout this entire project.

Finally, I want to thank my closest friends and family for their support and encouragement during my entire academic journey and especially during this project.

In Jaala, June 21st, 2020
Elina Uotinen

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

This thesis explores how big data analytics are utilized in marketing within the mobile gaming industry while drawing on the disciplines of knowledge management and resource-based theory. This introduction chapter begins by presenting the background of the study, as well as the research gap and research questions along with the aims of this study. Additionally, the research methodology, definitions of key concepts, the delimitations of the study and the structure of the thesis are described to provide an outline of this thesis.

1.1 Background of the study

Data has become a new form of capital, and it is what some of today's world's wealthiest companies are built on (Sadowski, 2019). This capital is generated by the average consumer, whose personal information has been converted into a commodity in the eyes of businesses (Erevelles, Fukawa & Swayne, 2016; Cheng & Wang, 2018). It has even been said that market competition has, in fact, turned into data competition (Xie, Wu, Xiao & Hu, 2016).

Chen, Chiang & Storey (2012) define big data as "a term that primarily describes data sets that are so large, unstructured, and complex that they require advanced and unique technologies to store, manage, analyze, and visualize". The massive influence and potential of big data are undeniable, and it is reshaping markets and marketing management, as well as consumers' habits. The practice of capturing rich and plentiful, structured and unstructured, data on nearly every aspect of consumers' daily life in real-time is both understood and even expected by marketers today (Erevelles et al. 2016; O'Connor & Kelly, 2017).

However, big data by itself holds no value or solutions. In fact, big data is only raw material (Xu et al. 2016), and only when it is used to extract insightful knowledge about consumers and prospects, it becomes relevant (Amado, Cortez, Rita & Moro, 2018).

Data can be extracted from numerous different sources including mobile applications, meters and sensors and other machine-to-machine communication, the internet of things, web logs, RFID, geolocation services, and largely from social networks, that have also become a means of influencing consumer behaviour (Braganza, Brooks, Nepelski, Ali & Moro, 2017; Cheng & Wang, 2016; Moro, Rita & Vala, 2016).

The challenge of transforming raw data into insights has led to data analytics having a pivotal role in managing and leveraging big data in marketing (Amado et al. 2018). Big data analytics can be defined as extracting and exploiting hidden consumer insights through advantageous interpretation, and it has an increasingly important role in businesses (Erevelles et al. 2016) The use of these technologies is completely transforming decision-making in marketing (Erevelles et al. 2016; Sundsøy, Bjelland, Iqbal, Pentland and De Montjoye, 2014).

The video game industry, especially mobile gaming, is known for its innovative and diverse uses of data (Shields, 2018), and as Neogames states in their most recent Game Industry of Finland report (2018), data analytics are here to stay. The industry has experienced exceptionally rapid growth and risen to mainstream success in recent years, and this trend is not slowing down: in 2019, global revenues of gaming reached over $150 billion with mobile gaming being the biggest segment, a 9.6% increase from the previous year, and by 2022 they are expected to grow to $196 billion. (Newzoo, 2019; Simon, 2018). One significant driver behind the massive success is, in fact, big data (Rands, 2018). Since every single aspect of the gaming experience can be measured, companies are constantly tracking KPI's on players' activity, engagement and behaviour, which provides them with insight into, for example, customer churn, user acquisition and retention (Addepto, 2019). Data is also used in game development and personalization to improve engagement, which in turn makes players stay for longer and spend more. Furthermore, players' gameplay choices and actions can be tracked, analysed and used to build a profile about the player's personality or consumption habits (Stafford, 2019; Oliveira, Santos, Aguiar & Sousa, 2014). Data is also vital for monetization purposes. For instance, Supercell uses machine learning tools to target sales promotions at individual users (Clayton, 2018). These types of

tools are necessary, since user acquisition remains both expensive and challenging and competition in the field is fierce (Neogames, 2018; Waller, Hockin & Smith, 2017). Despite the diverse benefits and plentiful opportunities that can be gained from big data, more than half of big data projects fail (Mithas, Lee, Earley, Murugesan & Djavanshir, 2013). Erevelles et al. (2016) state that failure often stems from the unique resource requirements that big data poses. Indeed, the effective management of organizational resources is of growing significance (Braganza et al. 2017), and firms need to allocate the correct physical, human and organizational resources to big data (Erevelles et al. 2016). The failure to exploit the benefits of big data can also happen due to the low levels of knowledge, adoption and utilization of intelligent big data analytics tools in marketing management, despite their proven advantages and huge potential (Miklosik, Kuchta, Evans & Zak, 2019).

Furthermore, the need for an improvement in knowledge management in firms is also growing, since big data's great advancements have unfortunately not come with an increased information management capability (Fosso Wamba, Akter, Edwards, Chopin & Gnanzou, 2015), and managing the knowledge generated through big data analytics requires a systematic and integrated approach (Ferraris, Mazzoleni, Devalle & Couturier, 2019). Overall, despite marketing being one of the fields with experimental big data approaches (Bendle & Wang, 2016), the potential of big data in marketing remains untapped – the need for more research is apparent (Balducci & Marinova, 2018; Amado et al. 2018; Leeflang, Verhoef, Dahlström & Freudnt, 2014; Miklosik et al. 2019).

1.2 Research gap

In academic literature the amount of research on leveraging big data analytics in marketing is still too low due to the novelty of the phenomenon. However, the number of published studies on the topic has been doubling each year, indicating its growing relevance (Amado et al. 2018). Still, the utilization of big data analytics has been studied a lot more thoroughly within other contexts than marketing. For example, the

effect of leveraging big data analytics on firm performance has been studied by, among others, Dubey, Gunasekaran, Childe, Blome & Papadopoulos (2019), Ferraris, Mazzoleni, Devalle & Couturier (2019), Fosso Wamba et al. (2017), and Gupta & George (2016). Big data analytics in business has also been studied from many other viewpoints including but not limited to: market research and predictive analysis (Bendle & Wang, 2016; Mishra, Luo, Hazen, Hassini & Foropon, 2019), business processes for implementing big data initiatives (Braganza et a. 2017), projects and project-based organizations (Ekambaram, Sørensen, Bull-Berg, Olsson, 2018), and value co-creation (Xie et al. 2016) as well as customer relationship management (Zerbino, Aloini, Dulmin, Mininno, 2018).

Researches have approached the topic from different theoretical viewpoints. For instance, Chan (2014) and Del Vecchio, Secundo & Passiante (2018) discuss big data through the lens of customer knowledge management and demonstrate the theory's relevance in value creation via big data, and Chong, Ch'ng, Liu & Li (2017) integrate theory on electronic word of mouth (eWOM) into their research on predicting customer demand by utilizing big data extracted from online reviews.

Resource-based theory has been applied to various studies in marketing, for example by Day (2014) and Kozlenkova, Samaha & Palmatier (2014), but in the context of big data analytics the number of studies is fewer. It has been utilized by, for example, Dubey et al. (2019) to study predictive analysis in manufacturing activities, and by Gupta & George (2016) to identify the resources needed to build big data analytics capability. Only little research has been conducted on big data-driven marketing using resource-based theory despite its ability to offer a valuable explanation of the influence big data has on marketing (Erevelles et al. 2016).

Similarly, knowledge management has been used in several studies relating to the use of big data in businesses (Fuchs, Höpken & Lexhagen, 2014; Chierici, Mazzucchelli, Garcia-Perez & Vrontis, 2019; Ekambaram et al. 2018, O'Connor & Kelly, 2017), but when it comes to data-driven marketing, the number of studies is still small. This is despite the fact that research has shown that knowledge management has a

fundamental role in managing and applying big data analytics, and that big data has immense potential in knowledge creation (Pauleen & Wang, 2017; Sumbal, Tsui & See-to, 2016). Still, managing knowledge that has been generated by big data has only scarcely been researched, and the need for further investigation is clear (Ferraris et al. 2019; Pauleen & Wang, 2017; Sumbal et al. 2016).

Research on data-driven approaches to marketing has discussed, for example, data privacy in marketing (Cheng & Wang, 2018), predicting consumer demand via analysing data from promotional marketing and reviews (Chong et al. 2017), the selection and adoption of analytical machine learning tools in marketing (Miklosik et al. 2019), and the use of data collected from social networks and games in marketing (Oliveira et al. 2014). However, research on the applications of big data analytics in this field still remains scarce. Amado et al. (2018) conducted a literature review on the use of big data in marketing via a text mining approach, and they conclude that there is a research gap about the potential benefits and challenges that come with it.

Furthermore, in their paper Leeflang et al. (2014) focus on the challenges of digital marketing and also call for more research on the topic. A research gap also exists in the application of intelligent analytical tools within marketing: Miklosik et al. (2019) state that only little is known about marketers' knowledge about machine learning tools and their adoption and utilization. In their literature review on unstructured data in marketing Balducci & Marinova (2018) also identified many areas of research that are still to be addressed and state the need for further research.

This need is especially prominent in the context of the video game industry, since the amount of academic literature in this field is still minimal, despite the industry's phenomenal growth and success. Few examples include Mathwes & Wearn (2016), who in their paper explore the different contemporary methods that video game marketing is carried out, and Waller et al. (2017) who examine entrepreneurs' marketing strategies for mobile games.

The aim of this thesis is to fill the research gap in big data-driven marketing while using a relevant theoretical framework based on knowledge management and resource-based theory to identify the practices and opportunities of leveraging big data in marketing within the mobile video game industry. Additionally, this thesis attempts to shed light on the use of intelligent tools in big data analytics within marketing, and ultimately expand existing theory on the phenomenon.

1.3 Research questions and aims of the study

As mentioned previously, the utilization of big data in marketing has enough influence to transform marketing management, and it has been proven by prior research that big data-driven marketing has massive potential in, for example, helping predict customer demand more accurately by analyzing hidden insights about customers or prospects that have been captured in real-time. Furthermore, the use of intelligent analytical tools is vital in big data management and decision-making, as well as the development and optimization of marketing strategies, and by utilizing these tools firms can achieve superior conversion rates. Despite these numerous advances, many firms are not able to fully harness the potential of big data in marketing. This is why marketing management has a need for more efficient management of resources, better knowledge about the adoption and utilization of intelligent analytical tools, and improved knowledge management practices and capabilities.

In order to help fill the research gap in big data-driven marketing in general, and especially in the context of the mobile video game industry, this thesis aims to explore the role that big data management and big data analytics have in marketing within this field. Furthermore, this thesis identifies the different kinds of intelligent analytical tools that are utilized in firms' big data marketing processes, and the reasons why these tools are applied. Through adopting a theoretical framework built on resource-based theory and knowledge management, this thesis also aims identify the most crucial organizational resources essential to data-driven marketing, and explore the creation, analysis, and application of big data knowledge.

Therefore, the main research question of this thesis is:

"How can big data analytics be leveraged in marketing within the mobile gaming industry?"

Supporting and helping answer the main research question are three sub-questions:

SQ1: *"What are the most vital organizational resources needed to enable and facilitate data-driven marketing?"*

SQ2: *"How is big data knowledge created, analysed and applied in data-driven marketing?"*

SQ3: *"What kind of intelligent tools and techniques are employed by companies in data-driven marketing?"*

By focusing on answering these questions, this thesis will provide further insight into the role of big data analytics in marketing, its practices, its connection to knowledge creation and knowledge management, and their foundation in the relevant organizational resources that are necessary in today's marketing management. The gained understanding from this thesis will be suitable and beneficial especially for the mobile video game industry, and the results will provide insights that can help enhance firms' marketing operations, as well as optimize their knowledge management processes.

1.4 Research methodology and data collection plan

The empirical study conducted in this thesis is a multiple case study, and its aim is to gain further insights and a deeper understanding of the relatively unknown phenomenon of how big data -driven marketing is utilized in the mobile gaming

industry, along with the organizational resources essential to it, the different types of tools and techniques used in it, and its process of knowledge creation and its management. The method chosen for this research is qualitative and data is collected via semi-structured interviews of key persons in several selected organizations.

Data is analyzed by utilizing a theoretical framework that is built on combining elements from knowledge management as well as resource-based theory in the context of this study. The findings of this study will have both theoretical and managerial contributions. The research design and method are discussed in more detail later in this thesis.

1.5 Definitions and key concepts

The most relevant key concepts of this thesis are introduced and defined below.

Big data is a term that describes sets of data that are so unprecedentedly massive, complex and unstructured that they require sophisticated and advanced technologies for their storage, processing and analyzing (Chen et al. 2012; Xu et al. 2016; Chan, 2014). Big data characterized by its four key dimensions, the four V's: high volume, high velocity, high variety and high veracity (Gartner Group, 2011; Cravens & Piercy, 2014; Chan; 2014; Erevelles et al. 2016). Exploiting big data has the potential to increase profits and create competitive advantage by gaining deeper insights about consumers and prospects from the immense amounts of data available (Braganza et al. 2017; Chan, 2014; Amado et al. 2018).

Big data analytics can be described as the approach of organizations to managing, processing, and analyzing big data by extracting and exploiting hidden and valuable insights, that can be transformed into value and sustainable competitive advantage (Fosso Wamba et al. 2017; Erevelles et al. 2016; Sumbal et al. 2016; Xu et al. 2016). The different types of tools and technologies utilized within big data analytics include, for instance, social media, mobile devices, technologies enabling internet of things,

cloud-enabled platforms, predictive analysis along with various tools fueled by machine learning (Fosso Wamba et al. 2017; Dubey et al. 2019; Miklosik et al. 2019). Through big data analytics, firms can comprehend the enormous volumes of data, as well as categorize and analyze it to derive useful information, which can ultimately lead to enhanced firm performance (Chen et al. 2012; Ferraris et al. 2019; Sumbal et al. 2016).

Knowledge management is a well-established discipline developed in the early 1990's (Pauleen & Wang, 2017), that describes the process of creating, sharing, transferring and applying knowledge within an organization to capture its collective expertise and intelligence (Chan, 2014; Alavi & Leidner, 2001; Meso & Smith, 2000). According to knowledge management theory, an organization's value is limited by the amount of knowledge within it (Grant, 1996). Knowledge is an asset that is unique and difficult to imitate, which makes it a basis for ensuring sustainable long-term competitiveness and (Romano, Passiante, Vecchio & Secundo, 2014; Lusch, Vargo & O'Brien, 2007). Furthermore, the generation and dissemination of knowledge are also vital factors in improving firm performance (Sumbal et al. 2016), supporting innovation and productivity (Ekambaram et al. 2018), and developing competitive advantage (Day 1994; Grant, 1996).

Resource-based theory is described by Gupta & George (2016) as "the principal paradigm for theoretically and empirically assessing the relationship between organizational resources and firm performance". The theory depicts firms as a collection of resources (Mishra et al. 2019), that form the basis for achieving and sustaining competitive advantage, and improved firm performance, but only if these resources are valuable, rare, inimitable and not substitutable (Barney, 1991). A firm's resources can be categorized into tangible resources, such as financial and physical resources, human skills, such as employees' knowledge and competence, and intangible resources, such as organizational learning and culture (Grant, 2010).

1.6 Delimitations

The applicability of the achieved findings of this thesis is limited due to several causes. Primarily, since this the focus of this thesis is exclusively on Finnish mobile gaming companies, the results that are achieved will therefore not represent the global mobile gaming industry as a whole. In addition, since the companies that are interviewed in this thesis follow the legislative obligations of Finland and the EU, the practices and technologies they utilize are presumably not completely applicable on a global scale. Additionally, the empirical data collected in this thesis is ultimately reflective of the interviewees' personal viewpoints and experiences, and the sample size is relatively small. All of these factors lead to this thesis having limited generalizability on a larger scale.

Furthermore, it should be noted that this thesis only concentrates on examining the phenomenon of big data -driven marketing from the perspectives of knowledge management and resource-based theory while disregarding other theoretical viewpoints.

1.7 Structure of thesis

This thesis is divided into five chapters. Firstly, this introduction chapter covers the background of the study, the research gap, description of the research questions and aims of the study, a brief review of the chosen methodology, definitions of the key concepts, and finally, the delimitations of the study. In the second chapter of this thesis, the theoretical framework utilized in this study is introduced, along with a literature review of the most relevant theoretical concepts included in the framework. The chapter covers academic literature on resource-based theory, as well as knowledge management while simultaneously also relating them to practice of leveraging big data analytics in marketing.

The third chapter describes the research design and methods for data collection and analysis that are used in the empirical study conducted in this thesis. The chosen sample, along with the criteria that affected its selection and some background information on it are also discussed. In addition, the third chapter also offers a look into the study's reliability and validity.

Next, the fourth chapter examines the findings of the study in the context of the new theoretical model that this study proposes. Finally, the fifth and last chapter of this thesis includes the discussion and summary of the key findings, as well as a review of this study's theoretical contributions and managerial implications, and lastly, the limitations of this study and suggestions for further research.

2 THEORETICAL FRAMEWORK

This chapter introduces the theoretical framework of this thesis, which is an integral part of this study. The framework is utilized in this thesis in multiple ways: firstly, as a foundation to provide this study with structure, as well as to help design data collection processes and analyze achieved findings, with the ultimate goal of helping answer the research questions proposed previously. In this chapter, each element of the theoretical framework is defined and described by drawing on relevant academic literature on resource-based theory and knowledge management and linking them to the context of utilizing big data analytics within marketing management.

The theoretical framework of this thesis (Figure 1) demonstrates the combination of knowledge management and resource-based theory in the context of big data -driven marketing. It displays the following: an organizations' resources and its knowledge management processes are seen as prerequisites for the organization's competence in leveraging big data analytics in marketing.

The arrangement of organizational resources into three types, tangible, intangible and human resources is based on the classification of big data resources used in recent IT capabilities literature, for instance by Gupta & George (2016). This classification along with its justifications are illustrated in more detail subsequently in this chapter. In the framework, knowledge management is further categorized into the three phases of an organization's process of its big data knowledge management: the creation, analysis and application of its knowledge. In addition, the act of leveraging big data analytics in marketing is displayed in the framework with the tools and techniques that are available for marketing management to utilize in their big data initiatives.

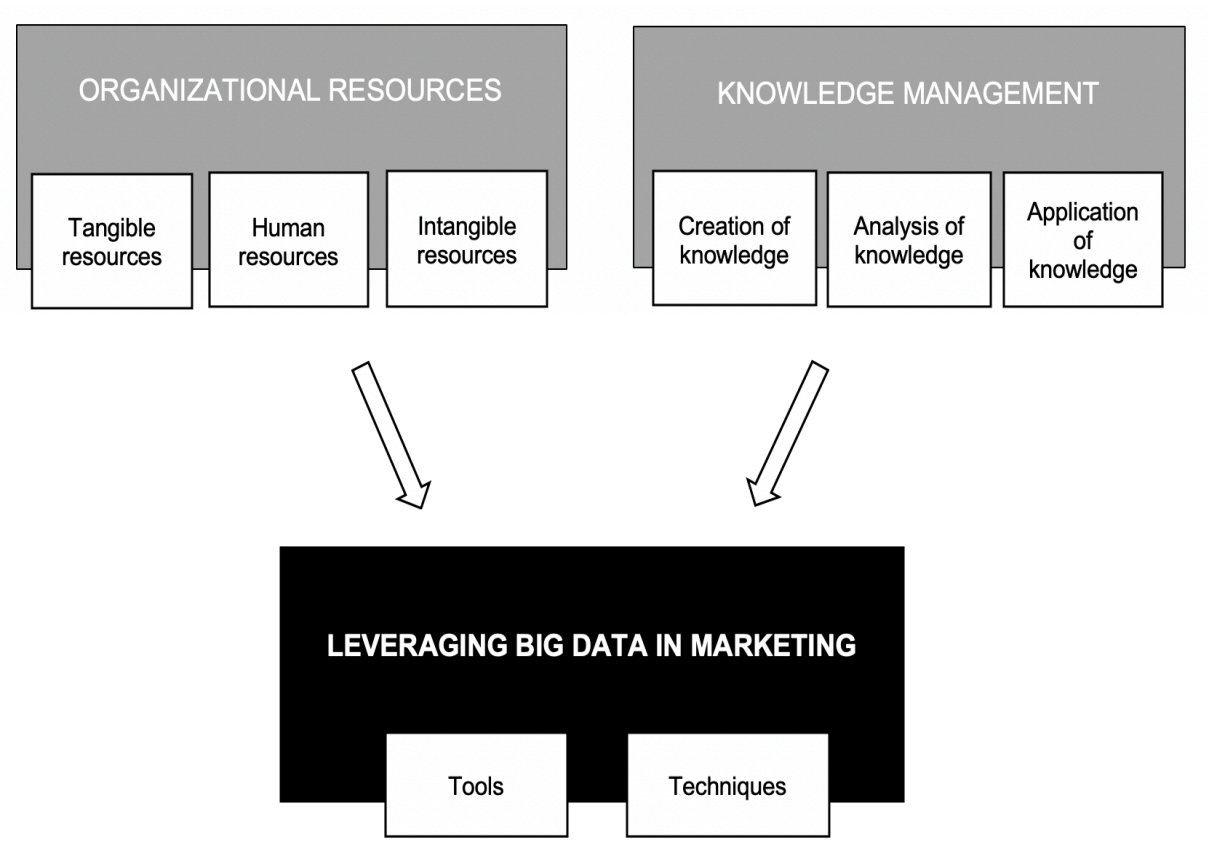

Figure 1: Theoretical framework

2.1 Resource-based theory and big data analytics

Utilized to review the relationship between organizational resources and firm performance, resource-based theory became the principal paradigm in the discipline of strategic planning during the 1990s and has since remained one of the most prominent theories in the field (Gupta & George, 2016). In fact, resource-based theory has become a well-accepted in several different business disciplines including marketing, as well as among IT scholars (Gupta & George, 2016; Erevelles et al. 2016).

Barney (1991) proposes in the article that played a pivotal role in the emergence of the theory, that strategic resources have four attributes that participate in generating sustained competitive advantage: value, rareness, imperfect imitability and non-

substitutability. When a resource gives a firm the ability to exploit market opportunities, improve its efficiency, or generate value to customers that competitors cannot achieve, it is considered valuable (Barney, 1991; Meso & Smith, 2000; Erevelles et al. 2016). A resource is deemed as rare, if it is not owned by a large number of other firms in an industry (Barney, 1991; Meso & Smith, 2000), and an imperfectly imitable resource is a resource that can be sustained for long periods of time without competitors being able to replicate it (Barney, 1991; Erevelles et al. 2016; Meso & Smith, 2000). Finally, a resource is considered non-substitutable when it has no strategic equivalents and it can be exploited in a way that others cannot (Barney, 1991; Meso & Smith, 2000; Erevelles et al. 2016).

These four conditions qualify a resource as a strategic asset (Wernerfelt, 1984; Peteraf, 1993) and they can therefore determine a firm's success in a given market (Meso & Smith, 2000). In addition, a firm's resources are able to form sources for competitive advantage, and resource-based theory is indeed also one of the most notable theories in understanding how firms can achieve and sustain competitive advantage by creating bundles of strategic resources (Barney, 1991; Mishra et al. 2019; Dubey et al. 2019).

The theory views firms as collections of resources and provides a framework for combining and deploying them to build capabilities, that can ultimately generate competitive advantage and improve organizational performance (Grant, 2010; Gupta & George, 2016). However, this is only possible if the firm's resources hold the four previously listed attributes: value, rareness, imperfect imitability and non-substitutability (VRIN) (Barney, 1991; Grant, 2010).

Resource-based theory is also suitable for understanding and illustrating the impact big data has on marketing (Erevelles et al. 2016). Indeed, big data can be considered as a resource that is necessary, but still not sufficient on its own to create big data analytics capability (Gupta & George, 2016). This is due to data not being a rare resource among firms, and because in order to build superior big data analytics capability, and ultimately competitive advantage, a firm also needs to be conscious of

the constantly evolving resources that are required for big data analytics, as well as to create an individual combination of its financial, physical, human and organizational resources (Teece, 2014; Grant, 2010; Amit & Schoemaker, 1993; Gupta & George, 2016; Erevelles et al. 2016). Moreover, the importance of effectively managing organizational resources is of growing importance especially in big data initiatives (Braganza et al. 2017).

Big data analytics capability is defined by Gupta & George (2016), in their study that identifies the required resources to build it, as “a firm’s ability to assemble, integrate, and deploy its big data -specific resources”, and it can lead to superior firm performance (Fosso Wamba et al. 2017; Ferraris et al. 2019). It is possible for firms to generate this capability to improve operational performance by creating a unique combination of the following three resources: strategic tangible resources, human skills and big data -driven culture (Gunasekaran, Papadopoulos, Dubey, Fosso Wamba, Childe, Hazen & Akter 2017; Gupta & George, 2016; Srinivasan & Swink, 2018). Furthermore, Braganza et al. (2017) propose that a firm’s capabilities enhanced by big data lead to value creation, and that big data -related organizational resources also need to meet the VRIN requirements in order to generate competitive advantage.

The different kinds of resources that a firm holds have been categorized in numerous ways in resource-based theory literature. For instance, Barney (1991) uses the classification of dividing the three resource types into physical capital resources, human capital resources, and organizational capital resources. In recent IT capability literature on big data, resources are classified into tangible, human, and intangible resources (Bharadwaj, 2000; Chae, Koh & Prybutok, 2014; Gupta & George, 2016), which is also the classification used in this thesis. These resources are illustrated in the figure below (Figure 2).

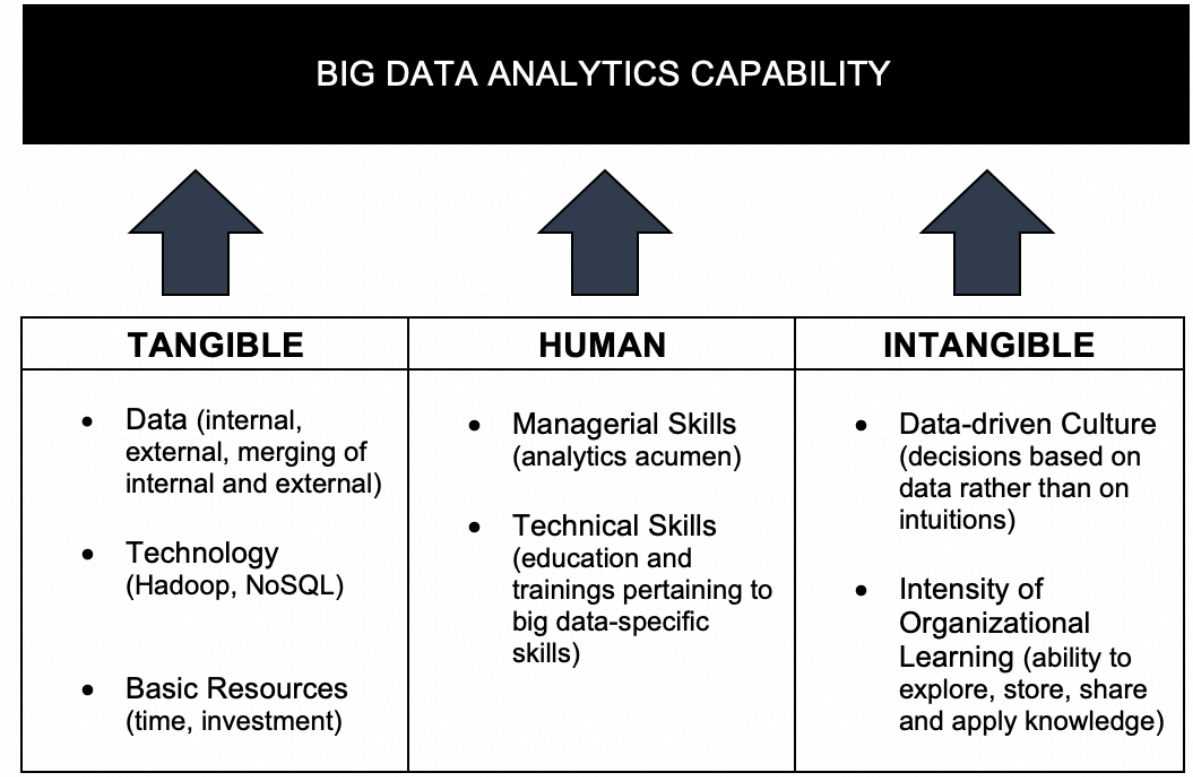

Figure 2: Classification of big data resources (Based on Gupta & George, 2016)

All three types of resources are seen to contribute to big data analytics capability, and all can be divided even further into seven subtypes. These resource types are discussed in more detail consequently.

2.1.1 Tangible resources

The first type of firm resources is called tangible resources. In traditional resource-based theory literature, they typically consist of financial and physical resources, such as cash, property, inventory, equipment and facilities (Grant, 2010). They are characterized by having a physical form and being relatively simple to valuate. In the context of big data, tangible resources can include, for instance, the software or hardware that a firm utilizes to generate, store and analyze big data (Erevelles et al. 2016). Since these kinds of tangible resources are relatively easily available for a great number of firms, they are not solely sufficient in creating competitive advantage on

their own. However, they are nonetheless required in order to create big data analytics capability (Gupta & George, 2016).

Recent literature regarding big data and the resource-based theory classifies tangible resources into three subgroups: data, technology, and basic resources (Gupta & George, 2016; Dubey et al. 2019).

Firstly, data is a vital tangible resource: today firms are collecting all the data they possibly can, both structured and unstructured, and data can even be considered a crucial factor of production (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh & Byers, 2011). Data can generally be categorized into internal and external data. Internal data refers to firm-specific data collected as a result of the firm's internal operations and it is often created for a specific business purpose, while external data represents data collected from outside sources, such as the web, social media and sensors, and it has the potential to provide deeper and more novel insights about customers or rivals (Zhao, Fan & Hu, 2014; Gupta & George, 2016).

Technology as a tangible resource in big data context refers to the novel technologies that are required to handle the massive volume, variety and velocity of big data and to be able to extract profitable insights, since traditional methods are simply not adequate (Dubey et al. 2019; Gupta & George, 2016; Xu et al. 2016). An estimated 80% of an organization's stored data is in unstructured format, which demands the use of new tools and technologies that allow distributed storage and parallel processing to more efficiently store, process and visualize big data (Dubey et al. 2019; Gupta & George, 2016; Kaisler, Armour, Espinosa & Money, 2013). In fact, these kinds of advanced and intelligent analytical tools serve as the principal source of information for marketing analysis (Miklosik et al. 2019). Big data analytics solutions are able to effectively support marketers and simultaneously relieve the burden of heavy human analysis (Amado et al. 2018), which is why their correct selection is a crucial element of harnessing big data. However, when it comes to managing these big data -specific technological resources, it is challenging to keep them exclusive and hidden from

competitors, due to, for example, labor force mobility and reverse-engineering (Gupta & George, 2016; Mata, Fuerst & Barney, 1995).

Lastly, basic resources are the final type of tangible resource. In essence, they include investments and time, which are both needed to build big data analytics capability (Gupta & George, 2016; Dubey et al. 2019). Significant and persistent investments into big data initiatives are obligatory in order to fully benefit from the opportunities that big data provides, but most firms are still lacking in this capacity, due to the novelty of big data and its related technologies (Gupta & George, 2016).

2.1.2 Human resources

In resource-based theory literature, human resources and skills comprise an organization's employees' knowledge and skills, such as business acumen, problem-solving ability and leadership quality (Barney, 1991; Grant, 2010). In the context of big data, human resources are a critical dimension of resources when it comes to building big data analytics capability (Gupta & George, 2016; Dubey et al. 2019), and having big data -skilled employees can be a substantial advantage over competitors (Waller & Fawcett, 2013). In recent big data and IT capability literature, human resources are divided into managerial skills and technical skills (Gupta & George, 2016; Bharadwaj, 2000; Chae et al. 2014).

Understanding how and where to apply the insights generated by big data technologies is of imperative importance (Gupta & George, 2016). Overall, managerial skills are essential in a firm's big data initiatives, but compared to other human skills, they are not the simplest to acquire since they tend to be highly firm-specific and are only developed over the course of time (Gupta & George, 2016; Mata, Fuerst & Barney, 1995). However, this is also what makes them so valuable and critical in terms of generating competitive advantage. These fundamental skills are especially difficult to match by competitors if they are developed by having mutual trust and strong

interpersonal relationships between employees in a firm (Gupta & George, 2016; Bharadwaj, 2000; Dubey et al. 2019).

Technical skills on the other hand can be relatively easier to develop, since they can be brought into a firm by hiring new employees or training existing ones. Gupta & George (2016) define technical big data skills as “the know-how required to use new forms of technology to extract intelligence from big data”. These skills include, for instance, the insight of data scientists to capture and manage information, their competencies in machine learning technologies, data extraction and statistical analysis, as well as mathematical modelling, optimization and forecasting (Erevelles et al. 2016; Waller & Fawcett, 2013; Gupta & George, 2016). These new and advanced skills are a necessity in order to keep up with the requirements of big data -driven business (McAfee, Brynjolfsson, Davenport, Paril & Barton, 2012; Waller & Fawcett, 2013), but firms are struggling to acquire them since there still is a notable lack of professionals with sufficient big data -specific technical skills (Chen et al. 2012). Furthermore, despite their indisputable importance, technical skills cannot generate long-term sustainable competitive advantage on their own, since these skills are prone to dispersing among industry professionals, which is detrimental to the quality and rarity of this resource (Nonaka, Toyama & Konno, 2000).

2.1.3 Intangible resources

The last of the three principal types of organizational resources in resource-based theory and strategic management literature are called intangible resources: they are resources that are, for example, undocumented on a firm’s financial statements and lack a physical form, such as organizational learning and organizational culture (Grant, 2010). They are not easily tradeable, and their value is challenging to assess, but they are considered absolutely essential to a firm’s performance nonetheless (Teece, 2015; Barney, 1995; Teece, 2014). Due to their nature, they are more likely to meet the VRIN requirements of a strategic asset and therefore be greatly valuable to a firm in generating competitive advantage (Teece, 2014). In addition to organizational culture

and learning, examples of intangible resources include, for instance, trademarks, copyrights and intellectual capital such as patents (Grant, 2010).

In the context of big data, Gupta & George (2016) divide these intangible resources into two: data-driven culture and the intensity of organizational learning.

In order to transform big data insights into action, carry out successful big data initiatives, and fully realize the potential of big data, firms require a data-driven culture and organizational structure (Erevelles et al. 2016; Gupta & George, 2016). In addition, a firm's organizational culture in general can be a source for competitive advantage (Barney, 1995; Teece, 2015). Data-driven culture is defined as the extent to which a firm's employees make decisions formed on insights extracted from big data (Gupta & George, 2016; Ross, Beath & Quaadgras, 2013; McAfee et al. 2012). Despite the its importance, firms tend to overlook the fact that an organization's culture can either enhance or inhibit their ability to benefit from big data, since only a small portion of firms have been able to benefit from their investments into big data (Shamim, Zeng, Shariq & Khan, 2019; Ross et al. 2013). Indeed, the reasons why big data initiatives can fall through relate to improper organizational culture and inadequate organizational resources (LaValle, Lesser, Shockley, Hopkins & Krushwitz, 2011; Erevelles et al. 2016), and in order to improve an organization's culture Ross et al. (2013) suggest diffusing data-driven decision-making to all levels of the organization regardless of job positions.

The intensity of organizational learning is the other intangible resource needed to build big data analytics capability (Gupta & George, 2016), and it is defined as "a process through which firms explore, store, share and apply knowledge (Bhatt & Grover, 2005; Cohen & Levinthal, 1990). Moreover, the intensity of organizational learning in a firm affects the firm's ability to reconfigure their resources based on changes in market conditions, and the firms that are able to utilize this ability are inclined to possess sustainable competitive advantage (Teece, Pisano & Shuen, 1997; Grant, 1996). This ability is relevant in the context of big data, since firms with a high intensity of organizational learning are more likely to have an advantage in creating big data

capability, and therefore be able to make informed decisions based on insights extracted from big data (Gupta & George, 2016).

2.2 Knowledge management in big data analytics

Developed in the early 1990s' within strategic management literature, knowledge management is a well-established discipline (Pauleen & Wang, 2017) that depicts an organization's process of creating, sharing, transferring and applying knowledge to capture its collective expertise and intelligence (Chan, 2014; Alavi & Leidner, 2001; Meso & Smith, 2000). Today, not only is knowledge management a well-known term, but also an integral part of modern organizations (Ekambaram et al. 2018; Sumbal et al. 2016). It supports firms in integrating, building and reconfiguring their competences through knowledge practices and dealing with the changes that occur in the market environment to increase productivity and innovation (Chierici et al. 2019; Ekambaram et al. 2018).

Ultimately, a firm's competence in applying existing knowledge to create new knowledge and to take action can lead to the creation of competitive advantage, as well as improved firm performance overall, since knowledge is an asset that is relatively difficult for competitors to imitate (Alavi & Leidner, 2001; Bassi, 1997; Sumbal et al. 2016; Grant, 1996; Day, 1994; Romano et al. 2014). Furthermore, this competence is especially enhanced by new technologies (Alavi & Leidner, 2001; Sumbal et al. 2016; Ferraris et al. 2019).

According to knowledge management theory, an organization's value is limited by the amount of knowledge within it (Grant, 1996). In his seminal article, Nonaka (1991) divides the knowledge that resides in an organization into two types: tacit and explicit knowledge. Tacit knowledge refers to the type of an organization's knowledge that cannot fully be expressed in words or shared, since it resides deeply in individuals' mental models, beliefs and actions (Gore & Gore, 1999; Meso & Smith, 2000; Nonaka, 1991). Due to its nature, this kind of implicit and deeply ingrained knowledge is

relatively challenging to be codified (Gore & Gore, 1999) and can therefore be inadvertently taken for granted in organizations (Nonaka, 1991). Furthermore, since tacit knowledge is employee-specific, such as their own self-motivated creativity or will for success, it is often lost when the employees leave the firm (Meso & Smith, 2000; Nonaka, 1991).

By contrast, explicit knowledge refers to knowledge that is more structured, conveniently documented, categorized and easily shared to others (Duffy, 2000; Nonaka, 1991; Meso & Smith, 2000). Since explicit knowledge is relatively simple to be codified, most organizations have made it available to all members of the organization through, for example, organizational repositories or technical systems (Meso & Smith, 2000).

While knowledge management has been an element of firms' strategies for already a considerable amount of time, new technologies such as big data pose new challenges for firms, along with the need for improved and updated knowledge management processes and systems to keep up with quickly evolving knowledge (Sumbal et al. 2016; Gupta & George, 2016; Xu et al. 2016). In order to leverage the huge potential that big data technologies are able to bring to a firm's knowledge creation and management capabilities, firms require a systematic and integrated approach, as well as an improvement in their knowledge management processes, since big data's immense advancements have unfortunately not come with an increased organizational information management capability (Sumbal et al. 2016; Ferraris et al. 2019; Fosso Wamba et al. 2015). Furthermore, big data is tightly linked to a firm's knowledge management capability, since the valuable insights extracted from big data are knowledge that can be utilized for enhancing firm performance, as well as because of its substantial potential in knowledge creation (Sumbal et al. 2016; Pauleen & Wang, 2017).

In traditional knowledge management literature, the process of knowledge management in a firm comprises four activities: creating, storing, transferring and applying knowledge (Alavi & Leidner, 2001). In the context of big data analytics

literature, these activities are slightly different. For instance, the first activity of big data knowledge management is described as “the process of collecting and storing records of consumer activities as big data” (Erevelles et al. 2016) and also as the “aggregation of data”, which refers to data acquisition, transformation and storage (Obitade, 2019; Wang, Kung, Wang & Cegielski, 2018; Raghupathi & Raghupathi, 2014). The second activity is data analysis, which includes the process of extracting insights from big data (Obitade, 2019; Wang et al. 2018). Lastly, the third activity of big data knowledge management processes is data interpretation and application, which includes utilizing insights in decision-making to enhance firm performance (Wang et al. 2018; Erevelles et al. 2016). Following recent literature, this thesis divides the processes of big data knowledge management in a similar manner into three stages: the creation, analysis, and application of knowledge.

2.2.1 Creation of knowledge

In classical knowledge management literature, the process of knowledge creation is defined as developing new content or replacing existing content by a continuous transfer, combination and conversion of an organization’s tacit and explicit knowledge (Alavi & Leidner, 2001; Pentland, 1995; Nonaka, 1991). This ability is crucial in regard to sustaining competitive advantage and fostering innovation (Nonaka, 1991; Alavi & Leidner, 2001). Firms should also be aware of the fact that knowledge is susceptible to becoming outdated, which is why it’s essential to invest in exploring new knowledge while also exploiting existing knowledge in order to survive in uncertain market conditions (Bhatt & Grover, 2005; Teece, 2015). Based on this, Gupta & George (2016) suggest that firms with a high intensity of organizational learning have an advantage in their knowledge creation processes. Furthermore, a firm’s knowledge creation must be constant, otherwise operational performance will suffer, which is why firms are required to construct proper strategies for their knowledge creation processes (Choi & Lee, 2002).

Although big data analytics is still a relatively new and unexplored phenomenon for most firms, it has immense potential for them in knowledge creation and therefore in generating sustainable competitive advantage due to its ability to help in understanding and extracting valuable and actionable insights from huge volumes of data (Sumbal et al. 2016; Pauleen & Wang, 2017; Erevelles et al. 2016). Arguably the most influential opportunity that big data enables in marketing is the extraction of hidden insights about consumers and rivals into deep and actionable knowledge that transforms decision-making, improves performance, and provides a source for competitive advantage (Erevelles et al. 2016; Sundsøy et al. 2014; Sumbal et al. 2016; Gupta & George, 2016; Zhao et al. 2014).

Data is generated and available for collection from countless sources within multiple business processes, internal and external, including but not limited to, sales, marketing, customer feedback, reviewing competitors as well as research and development (Sumbal et al. 2016). In addition, another valuable source for hidden insights is extracting or mining data from user-generated content, such as social media or online product reviews (Trainor, Andzuliz, Rapp & Agnithori, 2014). This extensive digital footprint with a massive amount of data that every single internet user generates is available for firms to use as input for their marketing analysis to improve the understanding of customer intent (Miklosik et al. 2019). However, the enormous amount of data also poses one the biggest difficulties that marketing management has to face when transforming these unprecedented amounts of raw data into deep and useful insights to be leveraged (Xu et al. 2016; Leeflang et al. 2014). Indeed, the sheer amount of data available for firms to utilize is so massive, that information overload becomes a challenge (Zhao et al. 2014).

In addition, one the greatest advantages and sources of knowledge that big data can offer is the measurability it brings to digital marketing (Miklosik et al. 2019). In fact, one of the biggest challenges for firms is indeed the trouble of assessing the effectiveness of their digital marketing operations (Leeflang et al. 2014). In marketing, where countless potentially relevant metrics must be considered, big data is able to offer more

accurate measurements than any method before it and it is also able to provide them in real-time (Balducci & Marinova, 2018; Miklosik et al. 2019).

Finally, it should also be noted that the process of knowledge creation is not limited to the collection of data: in addition, firms must transform the disparate data into an easily readable and analyzable format (Ward, Marsolo & Froehle, 2014). Also, in order to develop effective knowledge creation capabilities, a firm can benefit from supporting knowledge fostering enablers such as people, processes, systems and organizations that are working coherently (Sumbal et al. 2016).

2.2.2 Analysis of knowledge

A central element of a firm's knowledge management process is the analysis of the knowledge that has been created. This is especially important regarding big data knowledge, since it is only after analytics and context have been applied to data that it becomes information and gains valuable meaning (Sumbal et al. 2016). Without context, data is nothing but raw material: various disconnected facts about the flow of events and activities in an organization's system or database (Xu et al. 2016; Sumbal et al. 2016).

The exponential development of big data technologies has substantially increased firms' marketing analytics capabilities and even causing them to undergo a metamorphosis (Xu et al. 2016; Van Auken, 2015). This development has a notable effect on the integration of analytical tools and marketing strategies to generate value (Xu et al. 2016; Miklosik et al 2019). This is influential, since marketing analytics in an inherent part of effectively utilizing any digital marketing tool or process, and ultimately the cornerstone of planning and executing a marketing strategy (Miklosik et al. 2019).

Analysis of data includes organizing and structuring the data looking to discover underlying trends or patterns that can become useful information for decision-making (Sumbal et al. 2016; Ward et al. 2014). Data analysis can be further categorized by

the nature of data and the purpose of the analysis into three main types: descriptive, predictive and prescriptive analysis (Delen, 2014). In order to efficiently convert raw data into useful information, a firm must possess proper abilities to organize, coordinate, combine, integrate and distribute its knowledge from various sources (O'Dell & Grayson, 1998).

Within these processes it can be challenging to connect and integrate common data items across different internal and external sources and databases, while also selecting only the most relevant data for analysis (Zhao et al. 2014; LaValle et al. 2011). Likewise, it is also crucial to ensure the correct interpretation of data (Ekambaram et al. 2018). This is where a firm's big data capability and the intelligent analytical tools that come with it have significant meaning: they are able to alleviate tedious human analysis (Amado et al. 2018) and provide help in systemizing processes, streamlining planning and decision-making, as well as in automating work, all with increased efficiency and improved return on investment (Miklosik et al. 2019). Furthermore, in order to efficiently leverage the potential of big data in marketing analytics, marketing analytics solutions can be disseminated throughout all the levels of a firm (Laurent, 2013; Ross et al. 2013).

However, the process of transforming data into information is not enough. Firms also face the challenge of transforming this information further into fully accessible and actually valuable knowledge (Sumbal et al. 2016). This conversion from organized information to actionable knowledge is done by utilizing business intelligence tools to interpret the information, while the process is based on intuition and personal experience (Sumbal et al. 2016). Big data -fueled advanced tools that utilize revolutionary emerging technologies, such as machine learning and artificial intelligence, are not only able to analyze past events and circumstances, predict consumer behaviour and the success of strategies, but also learn entirely new skills from historical data, and optimize future activities in real-time as a continuous, ever-changing process (Heimbach, Kostyra & Hinza, 2015; Tettamanzi, Carlesi, Pannese & Santalmasi, 2007; Chen & Lin, 2014; Miklosik et al 2019). Moreover, the utilization of

these tools in marketing actually leads to superior conversion rates when compared with more traditional best-practice marketing approaches (Sundsøy et al. 2014).

However, despite the remarkable potential, Miklosik et al. (2019) find that in marketing management, there exists an alarming lack of knowledge about new intelligent analytical tools, and therefore and the adoption and utilization of them are also low. In addition, their utilization comes with limitations: for instance, tools utilizing emerging technologies such as machine learning and artificial intelligence, are not yet capable of incorporating important marketing elements such as creativity, empathy and intuition, or moral and ethical principles into their decision-making (Miklosik et al. 2019; Castelluccio, 2017; Coval, 2018). In addition, the implementation of these tools often requires considerable investments, both time and money, and even additional recruitment, and the information gained from them can still be inaccurate and lead to misrepresented decisions (Miklosik et al. 2019).

Ultimately, the entire process of transforming data into information, and information into knowledge, ultimately aims for improved decision-making, which is also a fundamental goal of the entire knowledge management practice: creating valuable knowledge to be used in decision-making through a proper analysis of the data generated via various sources (Raghupathi & Raghupathi, 2014; Sumbal et al. 2016; Lamont, 2012).

2.2.3 Application of knowledge

The applying of knowledge is recognized as the last activity of a firm's knowledge management process in classical knowledge management literature, and it is in fact the activity where the source of competitive advantage resides (Alavi & Leidner, 2001; Grant, 1996). It is the process oriented toward the actual use the accumulated knowledge, and it includes utilizing and sharing insights to enhance different organizational processes and operations, especially in prediction and decision-making (Obitade, 2019; Erevelles et al. 2016; Sumbal et al.2016).

The importance of technology in enhancing knowledge application has been well-known for a while now: for instance, Alavi & Leidner (2001) discuss how technology can improve the speed of knowledge integration and application by, among other things, codifying and automating organizational routines. Moreover, despite observing the positive influence technology has on knowledge application, they recognize that there are still challenges that firms have to face in this regard (Alavi & Leidner, 2001). Indeed, the use of big data analytics in knowledge application can pose a major challenge for organizations largely due to the novelty of the phenomenon (Sumbal et al. 2016; Leeflang et al. 2014). Firms are still critically unfamiliar with the utilization of big data analytics in their operations, and a large number of big data initiatives prove unsuccessful (Miklosik et al. 2019; Mithas et al. 2013).

Nonetheless, multiple studies find that big data analytics creates plenty of opportunities for firms in knowledge application through the potential it has in gaining an understanding of hidden information about their internal and external business processes, which can be utilized in efficient, informed, and overall improved decision-making (Sumbal et al. 2016; Gupta & George, 2016; McAfee et al. 2012; Waller & Fawcett, 2013), and ultimately creating new opportunities for generating competitive advantage (Miklosik et al. 2019). Knowledge generated with the help of big data analytics has a multitude of uses in firms. In addition to universally improving firm performance and decision-making, firm can utilize big data knowledge application to improve nearly all areas of business (Sumbal et al. 2016).

For instance: in marketing, multiple studies have found that the development and utilization of big data analytics technologies have immense potential in enhancing marketing management operations and answering the challenges firms face in today's marketing environment (Balducci & Marinova, 2018; Amado et al. 2018; Xu et al. 2016). For example, in order to optimize advertising campaigns and budgets, efficient tracking of customers is necessary, which is made possible through the utilization of big data (Leeflang et al. 2014), since big data technologies make such unprecedented amounts of consumer data available for firms to analyze and utilize.

Furthermore, the ability to manage and apply knowledge in real-time through big data, especially in marketing and sales, gives firms substantial competitive advantage over rivals (McAfee et al. 2012). The exponential development of data technologies is also transforming customer interactions (Marinova, de Ruyter, Huang, Meuter & Challagalla, 2017) and helping in customer segmentation, demand forecasting and predicting consumer behavior, creating personalized recommendations, improved targeted marketing, as well as risk management (Sumbal et al. 2016; Chen et al. 2014; Pauleen & Wang, 2017).

The knowledge mined and extracted from user-generated content can shed light on brand value attributes as well as brand threats, while providing suggestion for brand management strategies (Balducci & Marinova, 2018), along with forecasting product ratings and marketing performance (Chong et al. 2017; Moro et al. 2016), and revolutionizing market research (Bendle & Wang, 2016). Analyzing consumers' behavioral patterns is also key in improved targeting and personalization (Pauleen & Wang, 2017; Balducci & Marinova, 2018; Sumbal et al. 2016): for instance, Sundsøy et al. (2014) demonstrate how utilizing big data analytics in customer segmentation show a 13 times better conversion-rate achieved compared to best-practice marketing methods.

Moreover, the utilization of big data analytics in marketing and being able to capture consumer phenomena in real-time also brings opportunities for the optimization and automatization of advertising operations and improved effectiveness of promotion (Balducci & Marinova, 2018; Miklosik et al. 2019), as well as enhanced competitiveness and easier optimization of price levels (Chen et al. 2014). Additionally, Miklosik et al. (2019) also see the potential of big data analytics in enabling proactivity towards customers through engaging in real-time discussions, as well as in increasing the focus on key performance indicators, accelerating and automating activities such as reporting, and significantly reducing error rates.

However, when it comes to engaging customers via, for example, big data -fueled chatbots, consumers can prefer interacting with another human being instead of

software, such as automated chat bots, and automated responses can result in incorrect actions taken, which may threaten customer satisfaction (Miklosik et al. 2019; Go & Sundar, 2019).

3 METHODOLOGY

This chapter of the thesis introduces the research design and methodology that are utilized in the empirical study conducted in this thesis. In addition, the data collection process along with background information and criteria of the chosen sample, and the data analysis strategy and process are introduced. Lastly, the reliability and validity of this study are discussed in further detail.

3.1 Research design

Since the purpose of this study is to gain further insights and a deeper understanding of the relatively unknown phenomenon how data is utilized in marketing within the mobile gaming industry, the chosen research method is a qualitative approach. Indeed, a qualitative approach is an appropriate method for this study since it is suitable for discovering and comprehensively explaining complex real-life situations (Hirsjärvi, Remes & Sajavaara, 2009; Metsämuuronen, 2006), and also because, as discussed previously in this thesis, this phenomenon has not yet been researched enough. Some of the characteristics of qualitative research include the use of non-numeric data and a non-standardized method of data collection, both of which are also used in this study (Saunders, Lewis & Thornhill, 2016).

The research design chosen for this study is a multiple case study, since case studies are especially suitable for research that aims to gain an in-depth understanding of a contemporary phenomenon when contextual factors are relevant to the study (Yin, 1981). Another reason that a multiple case study was chosen for this study is the fact that existing literature on the phenomenon is still inadequate, which makes this a fitting approach (Eisenhardt, 1989). Furthermore, case studies are also convenient in generating novel theory (Eisenhardt, 1989), which is also what this study aims for.

Yin (1981) classifies case study designs into three different types: exploratory, explanatory and descriptive research designs. The design chosen for this study is explorative; a flexible research type that is often associated with qualitative research. This type of research was chosen to examine the relatively unknown phenomenon of data-driven marketing in the mobile gaming industry because, in general, exploratory research is effective in gaining a better understanding of an unfamiliar topic and its problems (Saunders et al. 2016; Brown, 2006). Exploratory research can be conducted in several different ways: for instance, by utilizing a literature review, expert interviews, or focus group interviews (Saunders et al. 2016), and the first two of these methods are applied in this study.

Furthermore, the research approach of a study can be either deductive, inductive or abductive (Saunders et al. 2016). An inductive approach was chosen for this study, since the amount of research on data-driven marketing especially in the context of the video game industry is considerably low, and an inductive approach focuses on building a richer theoretical perspective from the collected data, instead of relying on testing the collected data against existing theory (Saunders et al. 2016). An inductive approach includes identifying patterns and relationships as the primary data is collected and analyzed in order to allow meanings to emerge, and then relating these findings to existing literature (Saunders et al. 2016).

3.2 Data collection

In qualitative research, one way to collect data is through as research interviews, which can be defined as systematic conversations that have a purpose (Vilkka, 2015; Sanders et al. 2016). This is the single data collection technique that is used in this study. Interviews were chosen because this is an appropriate research method for relatively unknown phenomena that are not yet well understood (Hirsjärvi et al. 2009). Furthermore, individual interviews were chosen as the data collection method over group interviews, since individual interviews are able to provide deeper insight into an

interviewee's personal knowledge and experiences when compared to group interviews (Vilkka, 2015).

There are multiple ways to conduct research interviews, and a common way to categorize them is to divide them into three types: structured, semi-structured and unstructured interviews (Saunders et al. 2016; Vilkka, 2015). The interview type that was chosen for this particular study is semi-structured interviews, since this type is notably well-suited for this kind of exploratory research: for example, instead of using a rigid and standardized set of questions, semi-structured interviews allow for more flexibility and freedom, and the chance to collect important contextual data (Saunders et al. 2016; Hirsjärvi et al. 2009).

In general, semi-structured interviews have prearranged themes and topics to be discussed, but interview questions can be altered, added or even omitted to fit the flow of the conversation, which enables creating further discussions and attaining a rich set of data (Saunders et al. 2016). Furthermore, since this technique allows for further questions, it gives the interviewee a chance to explain more and provide profound knowledge about their reasoning (Saunders et al. 2016).

The data collection process for this study begun with finding suitable candidates to interview. The companies and experts to be contacted were chosen among Finnish video game companies that focus on mobile games and practice data-driven marketing. The companies vary by size from some of Finland's largest gaming companies to recently established start-ups. Potential candidates were contacted in April-May 2020 either via their company's email or directly via LinkedIn, and they were provided with a short summary of the topic and the themes of the potential interview questions.

Ultimately five people from different mobile gaming companies agreed to be interviewed. They all hold similar roles in their respective organizations, focusing on acquiring customers and promoting the growth of the company. The sizes and stage

of the businesses varied, but they all employed various and applicable data-driven marketing practices. Information on the interviewees is presented below (Table 1).

Table 1: Information on the conducted interviews

Interviewee	**Position in the organization**	**Approx. annual turnover of the company**	**Duration of the interview**
Interviewee 1	Performance Marketing Director	< 100 000 €	55 min
Interviewee 2	Chief Growth Officer	< 500 000 €	65 min
Interviewee 3	Head of User Acquisition	< 5 000 000 €	47 min
Interviewee 4	Chief Growth Officer	< 1 000 000 €	40 min
Interviewee 5	Performance Marketing Manager	< 5 000 000 €	37 min

A preliminary set of interviews questions was provided to each interviewee in advance to help prepare for the interview. This list of questions, which can be found in Appendix 1, is divided into three themes: data-driven marketing in general, organizational resources, and knowledge management. These questions were developed basing them on the theoretical framework of this study, and with the fundamental aim to help answer the research questions. However, these questions only served as an outline for the interviews, since open conversation and further questions were also encouraged. In addition, the respondents had the chance to add any further comments after all of the interview questions had been addressed.

The interviews were all completed during May-June 2020, and their duration varied from around 30 minutes to a little over an hour. Due to the exceptional circumstances caused by the COVID-19 situation, all of the interviews had to be conducted online via video calls instead of face-to-face meetings which had been the initial plan. The interviews were conducted in either English or Finnish depending on the interviewee's native language.

In addition, the interviews were recorded with the participant's permission to facilitate the process of analysing the research findings and to improve their fidelity (Saunders et al. 2016). Furthermore, each interviewee was briefed on the confidentiality of the interview and their responses. All of the respondents as well as their companies will remain unnamed in this thesis, since anonymity can potentially lead to responses that are more open and comprehensive, due to there being no risk of associating the companies with the confidential information presented in this thesis.

3.3 Data analysis

The analysis process and strategy in this study were formed with the intention of best answering the research questions while simultaneously following the proposed theoretical framework. The first step of the analysis process was accurately transcribing the audio recordings into written form manually after each interview. This was done to enable proper analysis of the findings. Transcribing helps with, among other things, managing the findings in a systematic manner, as well as with categorizing and classifying research findings (Vilkka, 2015; Saunders et al. 2016).

After becoming familiarized with the collected data through the transcribing process, the data was reviewed repeatedly case-by-case by employing within-case analysis. Within-case analysis involves getting closely familiar with each case on its own, and often also summarizing each case in writing (Eisenhardt, 1989). Detailed notes were compiled about each case, and the observations were organized around their relation

to the topics of this study, which is a way to help maintain focus on the actual aim of the study (Yin, 1981). Moreover, the emerging themes and patterns in each case were coded to help guide further analysis. By reviewing each case one-by-one, this method eases the burden of coping with an enormous volume of data, and also allows for unique patterns and trends to emerge from a single case before combining the data from multiple cases (Eisenhardt, 1989). Furthermore, as is common with theory-building case research, this study employed the technique of overlapping data analysis with data collection. This method included coding and analysing each interview before the next one was conducted, which allows flexibility and provides an advantage for further analysis (Eisenhardt, 1989).

Accelerated by the within-case analysis, the next step in the data analysis process was cross-case analysis, which aims to observe the data in several disparate ways (Eisenhardt, 1989). This step included comparing the cases, as well as their coded patterns and themes, amongst each other looking for applicable connections, similarities and differences, while also recognizing and ignoring irrelevant variations from case to case (Yin, 1981). These findings were initially grouped together following the theoretical framework of this study, and after a thorough examination of the data, five distinct phases of the cyclical data-driven marketing optimization process within the mobile gaming industry were identified. These findings were utilized to build an original theoretical model that illustrates the relationship between organizational resources and knowledge management processes, as well as the tools and techniques that are utilized in data-driven marketing within the context of the mobile gaming industry.

3.4 Reliability and validity

In order to ensure the trustworthiness of a study and its results, the reliability and validity of the research need to be examined. In addition, both reliability and validity also affect the quality and objectivity of research, and actions need to be taken to take them into account during every step of the research process. Firstly, a study's reliability

depicts its dependability and consistency: a study should produce similar and stable results if it is repeated regardless of the researcher's identity (Saunders et al. 2016). Essentially, this means that a study is reliable if another researcher can recognize similar observations and achieve the same results while using the same research method and process (Vilkka, 2015).

In the context of this study, reliability is ensured by, first off, adopting suitable research methods, and providing accurate descriptions of their aim and use. The interview questions are also put in writing, although due to the flexible and open nature of semi-structured interviews, the additional questions that were posed in the interviews cannot be documented. In addition, in order to minimize errors and bias in the research, interviewer bias should be considered, especially because there is only one researcher, and this could affect the interviews questions as well as the research as a whole. Interviewer bias is where the comments and behaviour of the interviewer can create bias in the interviewees' responses (Saunders et al. 2016).

In this thesis, this bias is combatted by recording the interviews and accurately transcribing them, as well as intentionally striving for objectivity and avoiding questions that could guide the responses. Furthermore, the respondents had the chance to add any further comments after the all of the interview questions had been addressed, which improves the reliability and adds to the trustworthiness of the interview (Saunders et al. 2016). In addition, the anonymity of the respondents and their companies in this thesis contributes positively to the reliability of this study, since this approach can promote more open and extensive answers.

Validity on the other hand refers to a study's ability to measure what it is actually supposed to measure: for instance, in order to gain validity, a study has to have appropriate methods for accurate data collection and analysis, and generalizability of its results (Vilkka, 2015; Saunders et al. 2016). To ensure validity and avoid systematic errors in the research, validity should be considered during the process of forming the research questions, as well when planning the interview questions for data collection (Vilkka, 2015).

The validity of this study is enhanced by systematically following the proposed theoretical framework throughout the entire empirical research process, which includes, for instance, selecting the most suitable methods for data collection and analysis, as well as the careful preparation of the interview questions which was also done based on the theoretical framework. However, ensuring validity during the interviews was compromised due to having to conduct the interviews online via video calls, instead of face-to-face which had been the original plan. It is possible that this change may have led to some lost information and therefore decreased validity.

Furthermore, as discussed in the introduction chapter, the generalizability of this study is limited. This is due to a relatively small sample size, and only focusing on Finnish gaming companies that produce mobile games. In addition, it should be noted that in this type of research, the responses of the interviewees are ultimately representative of their individual experiences and opinions.

4 FINDINGS

In this chapter, the findings of the empirical study conducted in this thesis are presented and reviewed. The discussion of these findings follows a cyclical model (Figure 3) that was developed based on the research data.

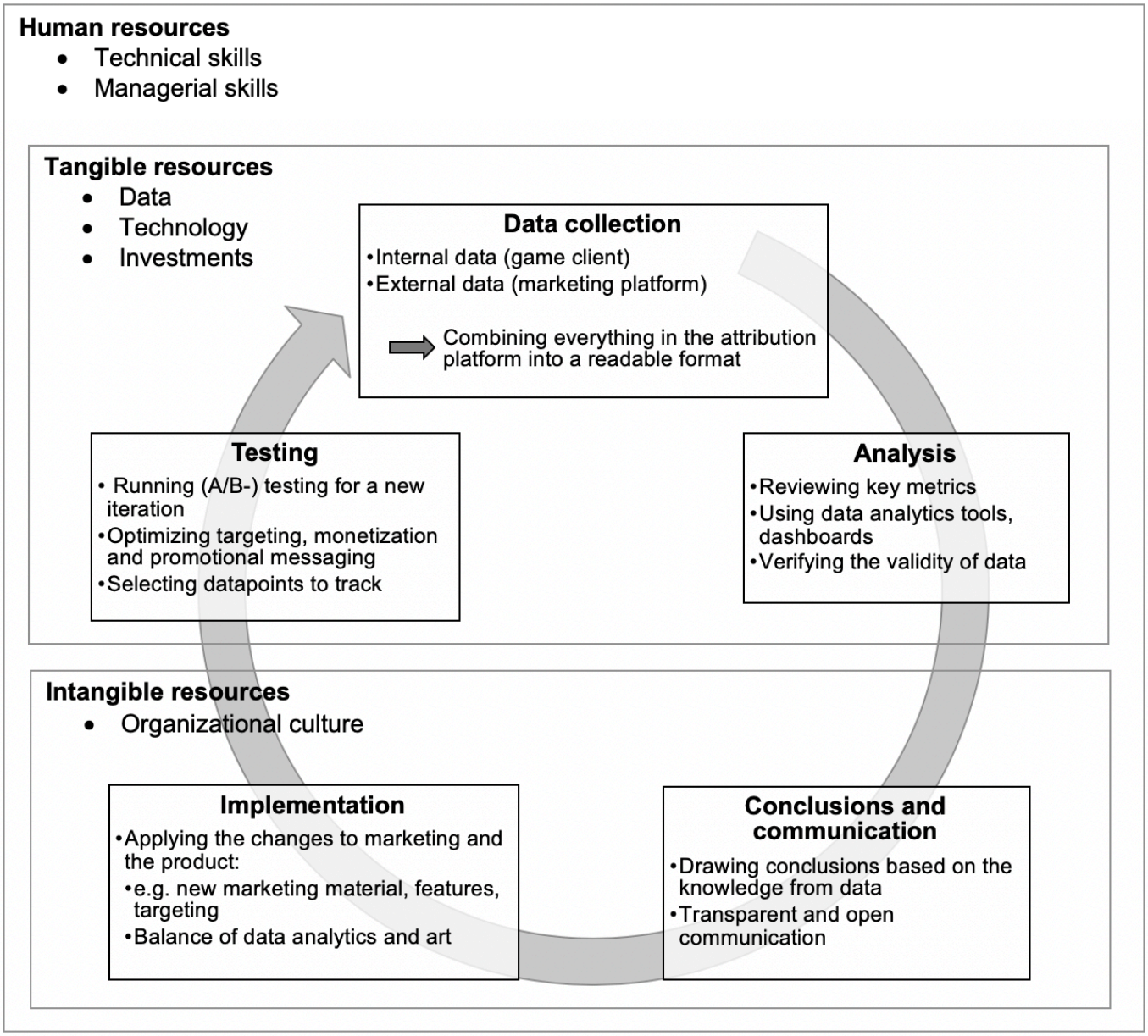

Figure 3: The data-driven marketing optimization process in the mobile gaming industry

This model presents five distinct phases of a typical approach to optimizing data-driven marketing in mobile gaming companies. These tightly linked and to some extent overlapping phases, that also display a data-driven knowledge management process, are enabled and affected by several crucial organizational resources, with the most crucial ones presented in the figure. As seen in the model, human resources are vital in each phase of the marketing optimization process. The importance of tangible and intangible resources on the other hand has been divided into the phases that are the most affected by them. In addition to illustrating the relationship between organizational resources and knowledge management processes, the model presents a summary of data-driven marketing practices in the context of the mobile gaming industry.

Each phase of the model along with its key elements is examined in more detail followingly. Thus, reviewing the findings begins with discussing the phase of data collection and knowledge creation, and finishes with examining how the applied knowledge and its effect on marketing performance are tested.

4.1 Data collection and knowledge creation

The first phase of the proposed data-driven marketing optimization model to be presented is the data collection and knowledge creation process, since without data, the entire practice of data-driven marketing could simply not function. This phase can be described as gathering and compiling hidden insights about customers from data that is collected from various different sources, storing it, and putting it all together in a consistent way to form usable and readable insights, for instance, in the form of key metrics to be tracked.

Data, which is also recognized as an undoubtedly essential tangible resource, is collected from numerous different sources and datapoints across a potential customer’s journey. Firmly linked with data is the other crucial tangible resource needed in this process: the technology required to collect and manage it. In the

interviews it became clear that it is essential that data comes from both internal sources as well as external sources.

The most significant internal source for knowledge creation, which according to interviewee 2 is also the absolute bare minimum needed for data-driven marketing in the mobile gaming industry, is an SDK, a software development kit, that is used to collect and interpret data from the game. The game client allows marketers, as well as product developers, to know exactly how users behave in the game, as every single one of their actions can be tracked. For instance, it is common to measure how long players stay in the game and how often they come back, how many sessions they have per day, and how far are they advancing in the game. Interviewee 4 mentioned how it is essential that this in-game data infrastructure is able to produce the right data at the right time. Furthermore, interviewee 2 pointed out that it is necessary to take into account the stability of these data systems: server infrastructure can break and the dependence on third-party vendors can pose a risk. Indeed, as mentioned by interviewees 4 and 5, it is vital that a company's internal technology resources, such as their back-end infrastructure and analytical resources are in a good condition and are able to handle all of the data.

Internal data can also be gathered from in-depth questionnaires that are sent to the players of the game, as explained by interviewee 5. This more of a traditional approach allows more profound insights to be gained on customers' motivations and preferences, which is useful for product development as well as marketing.

Regarding external sources of data, the most important one is the marketing and advertising platform, which allows data to be collected from every step of a potential customer's journey. For instance, it is possible to see exactly how, when, and on what device a customer reacts to an advertisement on a marketing platform, and how they proceed after that, as well as detailed information on the progression of ongoing and past paid user acquisition campaigns. In addition, it is vital to keep a customer's interest through each of these datapoints. The advancement of these campaigns and advertisements is tracked routinely through several key metrics that illustrate their

performance. Another source for external data is third-party services that provide detailed insights on market intelligence. These platforms provide information on, for instance, revenue per-download -rates and other key metrics across operators on the market.

It is important to note that the technology and procedures used to collect, store, and manage data vary by the size and the available resources of a company. In smaller gaming companies and start-ups, it is common to only use paid third-party software in their data-driven marketing processes. This is mostly due to the financial constraints and the high price of developing one's own tools and systems for data management. In addition, several convenient commercial tools are already easily available. However, there are large game studios that conduct all of their data collection and management all on their own without the use of third-party software. This requires significant investments to both personnel and technology, which makes this approach only really available for large companies with extensive tangible and human resources. For instance, interviewee 2 explained that in order to collect and store data without relying fully on commercial third-party software, the skills of a data engineer are necessary for building and maintaining infrastructure, scripting and converting data into organized tables to allow visualization. Furthermore, larger companies can also employ a specialized data scientist and even a machine learning expert.

A very common practice is to collect all of the data onto an attribution platform, also called a mobile measurement provider, which was regarded as the most used, and easily most important tool in most of the interviews, as well as the primary source of knowledge creation. This tool is able to collect accurate data on all actions of the customer from numerous different platforms and do so almost in real-time, and bring it all together onto one platform. Several interviewees stated how the attribution tool is truly central to everything, since it provides key metrics for marketers to examine from which channel the customers are coming from, and at what price, as well as how long they stay in the game and how much money they spend in it.

As pointed out by interviewee 1, all of this information allows marketers to know how to use their marketing budget: for instance, how to spread it across different channels, what kind of targeting to use, how much does a single user cost, and ultimately, what is the maximum user acquisition price that should be paid in order to still be profitable. Essentially, the attribution platform provides a detailed and almost real-time account of customer flow and behaviour, which makes it fundamental to data-driven marketing.

According to interviewee 3, there are two approaches to how the raw data collected by the attribution platform can be visualized, which is a vital step in the process of transforming data into actionable knowledge to be used in decision-making. Firstly, the more advanced and resource-intensive approach is to upload the raw data onto a data warehouse and manually create some code there in order to visualize it in a visualization and reporting tool. This approach requires substantial technical skills in the company, which why some firms tend to favour the second approach: uploading data onto paid third-party analytics tools to create dashboards for easy manageability and readability. Still, both of these approaches require the marketer to know what to look for and what questions to ask the data in order to gain any insight from it.

4.2 Analysis of knowledge

The second phase of model, the analysis of knowledge, is a key to leveraging data in marketing. Without analysis, raw data has no value. As interviewee 1 put it: “Data by itself is not enough. One really important thing is that whoever is managing it needs to be able to make certain decisions based on the it. Analysing data is central to that, since data is flowing from all directions these days and you have to do something to put it all together.” Furthermore, interviewee 1 continued:

> *“If there is nobody who is able to turn the data into business intelligence, then it’s pointless, then you’re just collecting data but not basing your tactical decisions on it.”*
> *– Interviewee 1*

Tracking the key metrics that show the progress of marketing and analysing their data week by week is a central part of data-driven marketing in the mobile gaming industry. It is important to know what to ask the data and what exactly to look for, and as interviewee 3 pointed out, the right questions help guide the time spent on looking for insights. For example, in order to find the precise moment a player becomes a loyal customer, there are specific patterns and reoccurring trends in the data to search for. In addition to looking for problems that need solving, analysis includes looking for opportunities. Interviewee 2 summarized the process of analysing followingly:

> *"Assuming the data is as accurate as possible, we look for trends, patterns, repetitions in it, we look for sudden movements and things that make you think that something is wrong."*
> *– Interviewee 2*

Several tangible resources were found essential to this phase of the marketing optimization process. Firstly, the basic resources, time and investments, that a company has access to largely define its available data analysis practices. For instance, a large factor in a company's data analysis capability is the number of employees dedicated to this area, as well as their variety. Larger companies have a higher feasibility of hiring a multifaceted team of specialists, focusing on, for example, creating advanced prediction models based on artificial intelligence technologies, while smaller companies with smaller budgets are more likely to have to make do with fewer people.

Overall, human resources and technical skills were considered vital in data analysis. Arguably the most important technical skills in data-driven marketing are those of a data analyst: their responsibility is analysing the data, uncovering and identifying patterns and meanings to warn about problems and opportunities in it. Interviewee 2 mentioned that a human is still much better than a machine for this task, since there are nuances and trends in data that require a human eye to observe and identify.

In addition to knowing what to ask the data and what to look for, a marketer should also be aware of the statistical aspect of analysing data, as pointed out by interviewees 3 and 5. Interviewee 3 described how it is crucial understand even a little about basic statistical concepts such as the confidence interval, adequate sample size, and statistical significance. In fact, interviewee 3 mentioned how an unintended disregard for statistical significance in the past had led to inaccurate estimations and therefore making misguided decisions in marketing and product management.

The tools used in data analysis also depend heavily on a company's available resources. As mentioned, smaller gaming companies and start-ups tend to only utilize commercial third-party software and tools in their marketing operations due to financial constraints. Developing and harnessing analytics in-house requires significant investments. Nonetheless, there are large game studios that are able to conduct all of their data analysis without using any third-party software. Furthermore, larger game studios have superior access to intelligent analytical tools that utilize, for instance, machine learning and artificial intelligence. These advanced tools can be used to, among other things, create detailed sales funnels and prediction models that can, for example, reveal exactly how much the company will earn in a certain time frame from a certain group of players.

The significance of using dashboarding tools in data analysis also came up in several interviews. Interviewee 3 explained how these tools allow the data collected on users' behaviour to be visualized and presented in a user-friendly manner that makes it simpler to find what you need while tracking multiple key marketing and product metrics. However, as interviewee 2 mentioned, these dashboards need to be crafted to be insightful in order to be useful: there is no need for vanity data, instead all of the visible data should actually be relevant. They also summarized this issue followingly: "Better to see problems in the data than perfection".

Another important issue that came up in multiple interviews is the validity and integrity of data. Interviewee 2 saw this as a challenge and explained how nobody really has completely and truly accurate data, and data can sometimes be inaccurate in general.

The proposed solution to this is to choose and decide one interpretation of data and stick to that in decision-making. Interviewee 4 shared similar views and explained how one of the biggest challenges in data-driven marketing is learning to trust the data and analytics tools. This requires that the company's data infrastructure is in a solid condition. In addition, as mentioned previously, statistical significance, adequate sample sizes and confidence intervals should be taken into account when analysing data, as pointed out by interviewee 3. Overall, a governance system to verify the validity of data is required in order to make the right decisions.

4.3 Conclusions and communication

The fourth phase of the data-driven marketing optimization process involves drawing conclusions based on the data and making the correct decisions, and perhaps even more importantly, communicating openly about the decision-making throughout all levels of the company. This phase of leveraging data-driven marketing relies heavily on human skills and intangible resources. In fact, these resources turned out to be exceptionally important in the interviews. For instance, the importance of a data-driven culture in the company became evident in all of the interviews. This was even recognized as a fundamental element of the entire practice of utilizing data in marketing and even product development.

The insights gained by analysing data, from, for example, reviewing key metrics on marketing as well as product performance, or examining detailed sales funnels that depict a customer's entire journey via several key conversion points, are utilized in decision-making in both marketing and product development. While optimizing marketing, for instance, if the data indicates that a certain marketing campaign is performing exceedingly well in a certain target group or area, it should be noticed so that the user acquisition budget can be shifted towards it. There exist intelligent machine learning and AI-fuelled tools that are able to help in spotting these opportunities and problems, as well as in making the best possible decisions based on them. However, the adoption and utilization of these intelligent analytical tools is limited

especially in smaller gaming companies, since they currently still require significant investments that tend to be possible for larger companies only.

Decision-making in product development also relies heavily on the data obtained by marketing. In addition to revealing unprecedentedly accurate information on users' in-game behavior and reactions to changes or updates in the game, data can even provide the insight needed to either begin or discontinue the development of a game, and consequently indicate how to reconfigure the company's resources according to the new market situation. By having this knowledge and having access to all the required datapoints, interviewee 1 noted that companies are able to save considerable amounts of money, for example, when deciding to discontinue a game instead of spending excessive amounts of resources for its development and marketing, when the data is able to indicate beforehand that these efforts will not be profitable.

Furthermore, the interviews revealed that due to this heavy focus on tracking and understanding in-game customer behaviour, the skills required of a marketer in the current world of data-driven marketing are evolving. Several participants talked about how marketing management in this field is transforming into product management, and therefore the role of a marketer is becoming increasingly like that of a product marketer. Moreover, data-driven mobile game marketing is unlike traditional marketing in other ways as well, possibly the most distinct way being the ability to receive immediate reactions by tracking potential customers and their behaviour in real-time.

Data also allows for conclusions to be drawn from past experiences by reanalysing or combining old data with current data, since data can be stored accurately and indefinitely. Interviewee 1 presented an example of utilizing data in decision-making from a previous product launch by comparing its key metrics to the performance of a current project. This method can be used to create valuable insights that are unprecedented in their accuracy when comparing with traditional marketing methods.

Communicating openly to the rest of the game studio about the conclusions and decisions that are made was regarded important by many interviewees. This requires

appropriate human resources in the form of managerial skills, which are a fundamental element of a company's ability to leverage data in marketing. Multiple participants talked about the critical need to demonstrate the value and importance of data in decision-making to all members of the game studio, or at least to everyone who is involved in decision-making, and not just to the marketing team. Furthermore, interviewee 5 stated:

> *"You need to have trust in other people's expertise and that they are doing things right. Transparency and a flat organization are important cornerstones for this."*
> *– Interviewee 5*

Indeed, it is extremely important that there is an adequate level of transparency in data-driven decision-making across the entire company, and members of the studio have to understand how it works, as pointed out by interviewee 3. For example, people should be aware of the ongoing advertising campaigns, their target audiences, their costs, as well as the number of acquired customers and how well they stay in the game. Interviewee 1 described the requirements for communication:

> *"You have to be able to communicate to the studio what should be done, what changes should be made and why. And based on the data, you need to tell in layman's terms what the data is telling us, and how we should react to it as a studio."*
> *– Interviewee 1*

In essence, it is imperative that data is presented in a format that is readable for all so that everyone in the company understands how and why data is used, and ultimately, as interviewee 1 put it: "I think that the main thing is that, in a way, the entire studio has to be built on being data-driven."

4.4 The implementation of knowledge

In the next phase of the proposed model, the knowledge that has been gained is applied: the decided changes and updates to marketing, as well as the product, are implemented.

A company's ability to apply the knowledge gained from data and use it in decision-making can determine their success. The huge importance of applying data-based insights in marketing and in product management became clear in the research data. Ultimately, the utilization of data helps to set budgets and guide business growth, and even model and build a future roadmap of the company. Interviewee 2 even pointed out that "Data is evolving marketing itself". Furthermore, interviewee 1 stated:

> *"Data produced by marketing steers everything that is done in a game studio, everything is ultimately based on what kind of data is collected from the market."*
> *– Interviewee 1*

In marketing, the knowledge gained from data is used in a plenty of ways. Interviewee 3 discussed how measuring different parts of the customer flow provides insight on how to improve marketing operations and paid user acquisition, as well as how to position a product and how to develop messaging for it. Indeed, data serves to optimize marketing and it supports everything in it in a rapid, almost real-time, pace. Knowledge from data also helps to spread risk, avoid fraud, conduct market analysis, and provide significant cost savings by maximizing return on ad spend. According to interviewee 5, some the most significant benefits in data-driven marketing are the ability to determine your exact target group of customers, and the predictability that utilizing data brings. Furthermore, interviewee 4 considered the notable costs savings as the most important benefit that data-driven marketing is able to offer, since it contributes to greatly increased marketing efficiency as well as reduced personnel costs.

Applying the insights gained from data in practice starts from everyone in the company standing behind the data collected from the masses, and the decisions that are made based on it, which was an issue brought up by many participants. As mentioned, having a data-driven organizational culture was found to be an absolutely essential intangible resource in data-driven marketing. Furthermore, interviewee 4 emphasized the importance of people being open to receiving the quantitative and qualitative feedback gained from data. This should be the case even if one's own personal opinions are different.

For instance, when data indicates that the development should be discontinued on a game the team has been passionately working on for months, it poses a challenge to keep everyone motivated, as pointed out by interviewee 1. This is why it is important that everyone in the studio has a data-driven, and business-driven, mindset. Many interviewees mention a possible lack of a business mentality in some of the more artistic-driven departments of a game studio, which is concerning, since data know-how is needed for successful companies and fast growth, as mentioned by interviewee 2. Moreover, interviewee 2 also suggested that the data-driven mentality needs to have its boundaries pushed as a phenomenon in general as well, not just in this context.

Furthermore, the inadequate understanding of data can also cause conflict when it comes to marketing and product decision-making. An example was even brought up of when the lack of a data-driven mindset ultimately caused the company to make misplaced decisions. Furthermore, a game designer and a marketer can have very disparate views on product development and what changes to implement, which is a considerable challenge. This is why a balance is required. Interviewee 3 explained:

> *"Games should be data-driven, but we can't forget about the players. It's difficult to make a game enjoyable only using a data-driven approach. I mean being data-driven should be your help to make the game better, but not be the key point."*
> *– Interviewee 3*

Furthermore, interviewee 4 summarized:

> *"After all, creativity and artistry are the nature of this industry, they are what advances this field, but analytics and analytics-based design are inevitably present in successful games."*
> *– Interviewee 4*

Indeed, in most interviews it was brought up that focusing heavily on data and profitable monetization does not signify a successful product. Even though data is extremely important, it cannot be everything, and product development should always include considering the quality of a game from the customer's point of view. Interviewee 4 mentioned how focusing too much on data and analytics poses the risk of the company ending up in a situation where absolutely no decisions or advancements are made without data-driven testing first, which is not an efficient approach. A balance is required, and as interviewee 3 explained: "We are here to make it profitable and enjoyable for users at the same time." Furthermore, interviewee 2 stated that an ideal game studio CEO is able to balance this creative conflict between data and art. They also mentioned how this conflict is easier to manage in smaller teams, which is why some game studios tend to divide their employees into small-scale units.

4.5 Testing and optimization

The last phase of the data-driven marketing optimization model to be presented is the process of testing a new iteration to learn how potential customers and users react. This process includes utilizing different kinds of tools and platforms to select and review the most relevant key metrics that are used to disclose crucial information on marketing performance, as well as insights on how to optimize it. The interviews revealed that adequate human skills and tangible technological resources are necessary in this process.

This phase if crucial, since the possibility to optimize targeting and promotional messaging, both in advertisements and in application store pages, iteration by iteration and gaining systematic knowledge on their performance is one of key aspects that make data-driven marketing so efficient. This allows profitable optimization of marketing and its budget, which in turn can ultimately lead to accelerated growth for the company. Sustaining growth at a healthy level is also helped by utilizing data, as pointed out by interviewee 2.

When it comes to testing and optimization in practice, marketing platforms are crucial. They are used to run paid user acquisition campaigns, as well as the testing of, for example, new targeting for the campaigns, and new or updated marketing material. The best platform for a game depends on its target audience and a few other aspects, but Facebook was generally regarded by the interviewees as the superior marketing platform, especially for testing. This was mainly due to Facebook's expertise in data and how they are able to provide persona-based attribution, which is more insightful and advanced than the regular cookie-based attribution that other advertising platforms use.

The most efficient utilization of marketing platforms requires technical skills, since these sills are helpful in tackling some of the issues that come with paid user acquisition. For instance, nearly every participant mentioned how user acquisition marketing platforms are becoming increasingly like a 'black box', meaning how marketers have less and less control over adjusting certain elements of their advertising campaigns, and how these campaigns are run and optimized by unknown and automated machine learning algorithms that can provide inconsistent results. Essentially, once a marketer feeds marketing material onto the platform, the platform shows it to different groups of people, and based on what kind of people react to it and how, it starts to distribute the advertisement to the most profitable people. While a marketer is still able to adjust the targeting and some other aspects of a campaign, it is entirely unknown how the platform functions and optimizes your campaigns, hence the term black box. Interviewee 3 summarized the issue:

> *"You can't be sure what you buy, what kind of traffic you buy. It's difficult because some week could be super profitable, and another week could be a lot worse, and you're always stressed about it."*
> *– Interviewee 3*

While this issue poses a challenge, the fact is that no human simply could match the agility and efficiency that a machine has when testing the profitability of different advertisements within different target groups. Interviewee 2 proposed that the solution to this issue is that instead of fighting an uphill battle, you should have your own weapons: innovative advertisements and a good ability to use data. Furthermore, in addition to knowing the marketing platforms along with their functionality and best practices, its beneficial to develop your own understanding of how the black box of user acquisition works and use that notion to your advantage in order to stand out from competitors. In addition, interviewees 1 and 5 emphasized the importance of constantly producing an extensive and diverse flow of marketing material: pictures, videos and playable advertisements, since a broad variety of advertisements helps optimize marketing and reach the largest number of potential customers possible.

When it comes to tracking the performance of user acquisition campaigns and testing, data is able to provide unprecedentedly accurate insights into a customer's entire journey. Interviewee 1 described this followingly:

> *"Data tells absolutely everything. We follow a sort of a sales funnel about when we show people advertisements: we know how many people see it, how many click on it, how many of those come to the store page, how many download the game, and how many of those start playing, how far they play, how many come back after an X number of days and so on. And in a way, refining this sales funnel is what marketing is about nowadays, you just have to take care of the funnel's functionality day in, day out."*
> *– Interviewee 1*

Many interviewees highlighted the importance of knowing exactly which datapoints to track and measure and select to be included in the funnel, since there practically countless points to choose from. These metrics are seen as the key to everything in data-driven marketing, and they provide valuable insight into the progression of marketing as well as the game itself. Interviewee 3 stated that: “It’s really difficult to overvalue to importance of analytics and measuring these things.” Furthermore, interviewee 4 stated:

> “*For us, it’s essential that every single euro we invest in marketing can be tracked in some way to find out if it works or if it doesn’t”*
> *- Interviewee 4*

The most important metrics were noticeably similar between all of the interviewees. Perhaps the most important one is retention, which illustrates how well a customer stays in the game by measuring how often they are still playing after, for example, one day, a week, or a month. Retention is one of the main metrics to track, since it is not enough to just get customers to download the game, but the goal is to keep them playing, as pointed out by interviewee 2. Another vital metric is the lifetime value, LTV, of players, since it shows how much time and money a user is spending in the game and therefore how much money they are earning for the company. Other key metrics include, among others, the click-through-rates on advertisements, the conversion rates from clicking to installing, and numerous different kinds of rates that measure the player’s advancement in the game. In addition, interviewee 5 mentioned how data is also able to reveal and even predict the point of saturation for advertisements. In fact, data can accurately indicate the need for new marketing material, which brings predictability into the optimal renewal cycle of advertisements.

By tracking these metrics and paid acquisition costs, it is possible to visualize an equation that shows exactly how much a single user costs and how much they produce. The aim of data-driven marketing and paid user acquisition is to optimize this equation: to operate as low-cost as possible to get the users with the highest returns, or in other words, to obtain users cheaper than what they produce. However,

interviewee 2 pointed out that the most expensive users also tend to turn out to be the best spenders. Furthermore, getting this equation to work in a profitable enough manner is one the biggest challenges in data-driven marketing, as stated by many interviewees. Ultimately, the return on ad spend, ROAS, is what all user acquisition activities are focused on: the better a user monetizes, the higher ROAS is, so the more can be spent on user acquisition.

Since marketing in mobile gaming is essentially based knowing the exact price it takes to bring a single user into the game and therefore knowing the user acquisition price precisely, as well as knowing the lifetime value of players, it is also possible to know the exact number of players needed to keep a game running profitably. In addition, data-driven marketing is also able to provide the tools and the know-how needed to reach this necessary group of players and retain them as customers. However, in order to get noticed and reach these people, mobile gaming companies face the challenge of beating highly intense competition, which was a challenge mentioned by almost all interviewees.

A significant aspect of data-driven marketing in the mobile gaming industry is also using the knowledge gained from data in product optimization. A key element is working on fine-tuning the monetization of a game and finding a balance between striving for profitability and for the enjoyability of a game. Moreover, marketing data also helps guide product development by noticing issues to be fixed in the game, as well as new opportunities. By analysing data on users' behaviour in the game, it is possible to see exactly how any changes that are applied affect key product metrics such as retention and conversion rates, which offers extremely valuable insights for the company.

Furthermore, the knowledge gained from marketing data affects even the very first steps of product development. Interviewee 1 explained how potential game concepts, along with its genre, world, monetization structure, and even its characters, are tested on the market, for example with A/B-testing, to determine what works and what type of product should be developed. In addition to testing the marketing performance of one's own game concept and how it resonates with potential customers, data is also

gathered to observe how competitors and their games are performing in order to comprehensively assess the market situation, as pointed out by interviewee 5. After the preliminary testing and development, a game is soft-launched in a few select markets, and at this stage marketing plays a central role. Knowledge gained from marketing data is able to tell if changes are necessary and point the developers and marketers in the right direction after carefully testing each iteration.

Essentially, key metrics about the product and marketing performance can show how the product, along with its marketing, should be optimized. Ultimately, the insights obtained from data reveal if a product has a future: interviewee 1 even stated that the biggest advantage of data-driven marketing is indeed how the numbers are able to tell explicitly if it is worth continuing to develop a game for all markets, or if it should be discontinued or, in other words, ‘killed’.

5 DISCUSSION AND CONCLUSIONS

The final chapter of this thesis provides a summary of the findings, which are discussed in relation to relevant existing literature. Subsequently, the theoretical contributions of this study are identified along with the most notable theoretical contributions. Next, this study provides managerial implications and suggestions for practitioners. Lastly, the limitations of this study and ideas for possible future research are discussed.

5.1 Summary of the findings

The aim of this study was to provide deeper insights into the role of big data analytics in marketing within the mobile gaming industry, as well into the fundamental organizational resources and knowledge management processes that play a key role in the phenomenon. Furthermore, the goal of this research was also to help fill the research gap that exists in data-driven marketing overall, but especially in combining resource-based theory and knowledge management in this context. In addition, the existing literature concerning this particular industry is still inadequate despite its relevance, which is why this study also aimed to expand existing theory on data-driven marketing in this exact field.

In order to conclude the findings of this research, the answers to the research questions of this study are discussed in relation to existing literature. This study had one main research problem:

> *"How can big data analytics be leveraged in marketing within the mobile video game industry?"*

As recognized in existing theory, the use of novel data analytics technologies is thoroughly transforming decision-making in marketing (Erevelles et al. 2016; Sundsøy et al. 2014) The findings of this study strongly support this claim: in fact, it was made

clear in the interviews that in addition to evolving marketing itself, data also truly steers everything that is decided in a game studio, and it can even be leveraged to build a future roadmap of the company. This study identified five distinct phases of the data-driven marketing optimization process within the mobile gaming industry, and combined them into a cyclical model, that is reflective of a firm's knowledge management processes as well as the organizational resources needed for it. This original theoretical model is able to illustrate how big data analytics can typically be leveraged in marketing within this specific industry.

The main research problem was supported by three sub-questions, with the first of them being:

SQ1: *"What are the most vital organizational resources needed to enable and facilitate data-driven marketing?"*

In order to enable and manage data-driven marketing operations, companies need to either possess or have access to the correct organizational resources, which is also supported by existing literature (Erevelles et al. 2016; Gupta & George, 2016; Gunasekaran et al. 2017). The importance of a few resources stood out from others in the findings of this study: firstly, every single participant highlighted the significance of a fitting, data-driven organizational culture present on all levels of a company. Indeed, it was seen as an absolutely essential intangible resource that everyone in the company involved with decision-making, through all levels of the organization, possesses a data-driven mindset, which is also recognized in existing theory (Erevelles et al. 2016; Gupta & George, 2016; Ross et al. 2013).

In value creation and in generating competitive advantage, organizational culture is especially relevant, since it often meets the VRIN requirements of a strategic asset: value, rareness, imperfect imitability and non-substitutability (Teece, 2015; Braganza et al. 2017). In addition, an inadequate understanding and mentality towards data in the organization can lead to misguided decision-making in both marketing as well as product development. The struggles in big data initiatives associated with improper

organizational culture are also reflected existing literature (LaValle et al. 2011; Erevelles et al. 2016; Shamim et al. 2019; Ross et al. 2013).

Having employees with valuable skills to handle big data can be a considerable advantage over competitors (Waller & Fawcett, 2013), which also became evident in the research data. It is vital that a company has adequate human resources at all stages of the data-driven marketing process (McAfee et al. 2012; Waller & Fawcett 2013), and it was recognized that acquiring more technical data-related skills in the company leads to improved capability and performance. However, while these skills are relatively simple to acquire via recruitment, solely focusing on data and technical skills is not favorable either and can lead to failure. Indeed, as noted in existing literature, technical skills alone are not enough to generate long-term sustainable competitive advantage (Nonaka et al. 2000).

Furthermore, in regard to human resources, it should also be noted that communication skills were found to be particularly important, since the importance and usage of data in decision-making should be conveyed to all members of the organization in terms that comprehensible to all. Existing literature remarks the significance of managerial skills and their development in leveraging big data analytics (Gupta & George, 2016; Dubey et al. 2019), but communication is seldom acknowledged. Furthermore, it was noted that a marketer should in some cases also possess expertise in product management, since the importance of utilizing marketing data in product development and optimization was found to be crucial. Overall, there exists the need for an exact appropriate amount of focus on utilizing data that should exist in decision-making. Fixating excessively on data in marketing, as well as in product development, does not result in the best possible outcome: instead, a balance is required.

Several undoubtedly essential tangible resources were identified. While data-driven marketing could simply not function without data itself and the technology needed to manage it, these resources are in no way sufficient to successfully leverage big data analytics in marketing. This aligns with existing literature, since it is recognized that the

tangible resources required in data-driven marketing often do not meet the VRIN requirements, since they are relatively simple to acquire and imitate, and they do not solely suffice in generating competitive advantage (Gupta & George, 2016).

The second sub-question was:

SQ2: *"How is big data knowledge created, analysed and applied in data-driven marketing?"*

The knowledge management processes in data-driven marketing employed by mobile gaming companies involve a similar process of knowledge creation, analysis and application that can be seen in existing literature.

While data itself represents explicit knowledge, a great deal of the crucial knowledge needed to manage it is in fact tacit. Indeed, the significance of tacit knowledge in data-driven marketing aligns well with the huge influence of a data-driven organizational culture. In addition, it became evident in the research data that there exists a need for intuition and a human eye, both of which can be considered tacit knowledge, in data analysis as well as in decision-making.

As noted in existing literature, novel big data technologies pose new challenges for firms, since processes and systems need to keep up with quickly evolving technologies to manage knowledge (Sumbal et al. 2016; Gupta & George, 2016; Xu et al. 2016). This study recognized some of these challenges as well as their potential solutions. For instance, despite the huge potential that leveraging big data analytics has in knowledge creation, companies struggle with the overwhelming number of potential factors, actions and elements available to choose from when selecting specific datapoints to track and turn into key metrics, which is a struggle that is also recognized in existing literature (Xu et al. 2016; Leeflang et al. 2014; Zhao et al. 2014).

However, while Leeflang et al. (2014) noted that one of biggest challenges for firms is assessing the effectiveness of their digital marketing, this study discovered the

opposite: many interviewees discussed how data-driven marketing allows companies to know exactly how much it costs to acquire a single customer, as well as how much they produce. In fact, utilizing data in marketing is what enables marketers to know precisely how potential customers react and how effective their marketing actually is, since every step of a potential customer's journey to a loyal player can be tracked and turned into key metrics.

The importance of selecting and examining these key metrics is apparent in both existing literature and the research data. For instance, Miklosik et al. (2019) and Balducci & Marinova (2018) both mentioned how one of the most valuable advances that big data brings to marketing is the access to real-time measurements that are more accurate than any other method before it. Indeed, many participants brought up measurability and tracking key metrics as some of the most fundamental aspects that data-driven marketing is built on, and how it is nearly impossible to overestimate their value.

Furthermore, it became evident in this study that without analysis, data has no use, which is a claim also supported by existing literature (Sumbal et al. 2016; Xu et al. 2016). Companies must possess the resources and know-how needed to turn data into business intelligence and make decisions based on it for the data to have any value. A key element in the process of analysing the knowledge gained through data, recognized by both this study and existing literature, is having the expertise to know what exactly to search for among the vast amounts of data coming from multiple different sources (Zhao et al. 2014; LaValle et al. 2011).

Another crucial element in the process of analysing data is ensuring its validity (Ekambaram et al. 2018). The findings of this study indicate that in order to confirm the correct interpretation of data, it is vital that the statistical significance of it as well as the integrity of the infrastructure that collects, stores and visualizes data are considered. Essentially, a governance system to verify the validity of data is necessary.

Knowledge gained from data can be applied in a myriad of ways in data-driven marketing. Furthermore, a company's success can be determined by their ability to apply this knowledge. This study recognized many practices of applying knowledge that align with the findings of previous literature. For instance, some of the most important ways to apply data-driven knowledge included accurately tracking customers' reactions and behavior to optimize advertising campaigns (Leeflang et al. 2014), identifying the correct target audience (Sumbal et al. 2016), and profitably optimizing marketing operations through continuous and precise testing (Balducci & Marinova, 2018; Miklosik et al. 2019).

Overall, utilizing data in marketing provides extensive advantages to decision-making and business operations as a whole (Sumbal et al. 2016; Gupta & George, 2016; McAfee et al. 2012; Waller & Fawcett, 2013). In addition, this study recognized the important role that marketing data has in product development and optimization in this industry, and how this data is vital through a product's entire lifecycle all the way from preliminary game concepts to scaling the product globally.

Finally, the last sub-question of this study was the most practical one:

SQ3: *"What kind of intelligent tools and techniques are employed by companies in data-driven marketing?"*

The tools and techniques that companies employ in data-driven marketing depend heavily on the company's available resources. However, this study did find that there are common practices and similar invaluable tools that every organization utilizes. For instance, the internal and external sources for data as well as the methods to collect it were extremely similar among respondents, as were the practices of running paid user acquisition campaigns and optimizing them.

The most notable differences in the used techniques were recognized in the process of visualizing raw data, and in analysing it. Some companies relied on commercial third-party tools for their data management, while others employed the more resource-

heavy approach of manually visualizing the data. Furthermore, while intelligent tools that utilize artificial intelligence or machine learning technologies were seen as valuable, their adoption was still low, especially in smaller companies. This issue is also noted by Miklosik et al. (2019), and it is largely due to the substantial financial resources needed in to integrate and employ these tools. Nonetheless, the level of knowledge about these tools and their worth is relatively high among marketers in this industry, contrasting the findings of Miklosik et al. (2019). Similarly to their findings however, the level of utilization of these intelligent tools is still relatively low.

5.2 Theoretical contributions

This study contributes to literature on data-driven marketing in several ways. In addition to helping alleviate the research gap that exists on this novel phenomenon, particularly in this industry, this research provides verification for the findings of multiple previous studies. In addition to supporting previous research and theories, this study introduces new insights on data-driven marketing, and even proposes a new theoretical model. Furthermore, besides identifying the most crucial organizational resources needed for data-driven marketing along with their value, based on the categorization by Gupta & George (2016), this research examines the knowledge management processes and practices utilized by firms. Furthermore, this study provides novel insights into the tools and technologies employed in data-driven marketing, as well as the challenges that companies in this industry face when leveraging big data analytics in marketing, with some of these challenges contradicting ones noted in existing literature.

This study provides three main theoretical contributions. First and foremost, in order to expand existing theory on the issue, the most notable theoretical contribution this study proposes is an original theoretical model that combines the theoretical concepts of multiple disciplines to illustrate how big data analytics can be leveraged in marketing, particularly in the mobile gaming industry. Five distinct data-driven marketing optimization phases that are reflective of a firm’s knowledge management processes

are recognized in the model along with their key elements. In addition, each phase is displayed with the most important organizational resources that influence it. Ultimately, this cyclical model manages to combine literature on leveraging big data analytics, resource-based theory as well as knowledge management in the context of this phenomenon, along with displaying their connected relationship, which makes it a noteworthy theoretical contribution.

Secondly, this research confirms and even further underlines the crucial importance of one intangible resource in a firm's big data analytics capability: a data-driven organizational culture through all levels of decision-making in the organization (Gupta & George, 2016; Erevelles et al. 2016; Ross et al. 2013). Furthermore, this study notes that the importance of a proper data-driven culture is made even more relevant by its influence in value creation and building sustainable competitive advantage, since organizational culture often meets the criteria of a strategic asset (Teece, 2015; Braganza et al. 2017). Moreover, this study confirms that improper organizational culture can indeed lead to ill-advised decisions and therefore diminish organizational performance (LaValle et al. 2011; Erevelles et al. 2016; Shamim et al. 2019; Ross et al. 2013).

Lastly, this study provides a contextual theoretical contribution. Previous research on the utilization of data in marketing has examined multiple different industries and fields, but this industry has not yet been represented. In addition, the scope and volume of research on the mobile gaming industry overall, as well as on the video game industry in general, are notably low. This study is able to reveal compelling novel observations about this inventive field that has pioneered in their utilization of data in business operations, both in marketing and product development.

5.3 Managerial implications

By connecting abstract theoretical concepts and practical empirical findings, this study contributes to a deeper understanding of the phenomenon that is leveraging big data

analytics in marketing. The findings of this study provide managers and practitioners within data-driven marketing, particularly in the mobile gaming industry, relevant insights into the crucial organizational resources needed for data-driven marketing, as well as its knowledge management processes, and the tools and techniques employed in them.

Combining all of these elements, this study presents a cyclical model that represents a typical data-driven marketing optimization process in five distinct phases within this specific industry. This model can be examined in relation to one's own organization to give perspective on data-driven marketing practices.

Out of the important organizational resources identified in this study, arguably the most valuable in terms of building sustainable competitive advantage is a proper data-driven organizational culture in the company. The importance of organizational culture as an essential intangible resource was recognized both in the empirical findings of this study as well as in existing literature. Everyone in the organization, or at least everyone involved in decision-making in any way, should possess a data-driven mindset, and have an adequate understanding of leveraging data along with its significance, and the advantages it brings. In addition, companies should also acknowledge the importance of suitably communicating this information to others in the organization in a readable and understandable format.

Furthermore, in addition to the relevant technical skills that are needed to collect, manage and visualize data, and run paid user acquisition operations overall, this study acknowledges how skills related to product management are also valuable in this context, at least in this specific industry. In fact, the utilization of marketing data in product development and optimization leads to a marketer having to consider the perspective of product management in, for instance, collecting data and selecting the most suitable datapoints and key metrics to track.

Moreover, this study identifies and provides insightful examples of the applications and benefits of data in marketing, as well as the motivations behind its usage and the most

influential challenges that come with it. Managers and practitioners can reflect on these issues in relation to their own data-driven marketing operations, and gain guidance and general advice on how to overcome challenges and ultimately enhance their marketing performance.

5.4 Limitations and suggestions future research

While this study made both theoretical and managerial contributions, it nonetheless comes with certain limitations. First of all, this research only covered mobile gaming companies from Finland, which limits the generalizability of the findings. This is due to cultural issues, and because the participating companies are under the obligation to follow the legislative rules of Finland and the EU, and thus these results are not representative of the global mobile gaming industry as a whole. However, it should be noted that many interviewees had an international background as well as experience in the field from other countries than Finland, which contributes to a more global viewpoint in this study. Still, in order to gain a more universal view of data-driven marketing in this industry, companies from multiple countries should be included in future research.

Furthermore, the context of this study was limited exclusively to the mobile gaming industry, and therefore is does not represent the video game industry collectively. There are major differences in how games are marketed and developed depending on their platform, and thus the findings of this study cannot be generalized to apply to the entire video game industry. In addition, it should be noted that the empirical data collected in this study is ultimately reflective of the participants' own experiences and opinions on the topic, which makes them not unquestionably generalizable. Also, in order to gain a comprehensive and universal understanding of data-driven marketing as a phenomenon, research should not focus only on one industry, but instead include diverse perspectives from a wide variety of fields.

Moreover, another factor limiting the generalizability of this research is that the sample size of the empirical study was relatively small and comprised mostly of smaller companies. The participating companies varied in their ages, sizes and stages of growth, but this study could have benefitted from having even larger companies represented. This is because company size largely determines the resources available for data-driven marketing, and as seen in this study, some of the more advanced and more insightful tools and techniques, as well as a larger number of personnel, are only available for companies with vast budgets. Having bigger companies participate in the study could have provided even deeper insight about, for example, advanced analytical tools that are fuelled by artificial intelligence and machine learning technologies and their utilization in data-driven marketing.

While this study identified some of the challenges related to the utilization of data in marketing, a deeper look into their causes and probable solutions could provide further insight on the topic. For instance, future research topics could cover the importance of ensuring data's validity in marketing, or the crucial influence of proper organizational culture. Furthermore, in order to assist in the employment of data-driven marketing, it could be useful to conduct research that focuses more on the actual process of building and developing the organizational resources and knowledge management processes needed to enable it. Also, another possible angle for future research could be the evolution that data technologies are bringing to marketing, and how marketing management, at least in this field, is becoming increasingly like product management. It could also be valuable to research how data-driven marketing, and novel data technologies in general, are affecting product development and even innovation in companies.

Lastly, it should also be noted that the tools, techniques and practices of data-driven marketing presented in the study reflect the phenomenon at the time of writing this thesis, but due to the fast-paced development and emergence of new technologies and trends, the findings of this study might not be entirely relevant or applicable in the future. In addition, the perspective of this study is relatively broad, and the phenomenon is still rather novel, therefore in order to examine some of the more

technical and profound aspects of data-driven marketing as well as its practices in more detail, further research is needed.

REFERENCES

Addepto. (2019). Benefits of Big Data Analytics in the Mobile Gaming Industry. *Medium.* [online document] [Accessed 12th February 2020] Available: https://medium.com/datadriveninvestor/benefits-of-big-data-analytics-in-the-mobile-gaming-industry-2b4747b90878

Alavi, M. & Leidner, D. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *Mis Quarterly, 25*(1), pp. 107-136.

Amado, A., Cortez, P., Rita, P. & Moro, S. (2018). Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis. *European Research on Management and Business Economics, 24*(1), pp. 1-7.

Amit, R. & Schoemaker, P. J. H. (1993). Strategic assets and organizational rent. *Strategic Management Journal, 14*(1), pp. 33-46.

Balducci, B. & Marinova, D. (2018). Unstructured data in marketing. *Journal of the Academy of Marketing Science, 46*(4), pp. 557-590.

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management, 17*(1), pp. 99-120.

Barney, J. (1995). Looking inside for competitive advantage. *The Academy of Management Executive, 9*(4), p. 49.

Bassi, L. (1997). Harnessing the power of intellectual capital. *Training & Development, 51*(12), pp. 25-30.

Bendle, N. T. & Wang, X. (2016). Uncovering the message from the mess of big data. *Business Horizons, 59*(1), pp. 115-124.

Bharadwaj, A. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *Mis Quarterly, 24*(1), pp. 169-196.

Bharadwaj, A. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *Mis Quarterly, 24*(1), pp. 169-196.

Bhatt, G. & Grover, V. (2005). Types of Information Technology Capabilities and Their Role in Competitive Advantage: An Empirical Study. *Journal of Management Information Systems, 22*(2), pp. 253-277.

Braganza, A., Brooks, L., Nepelski, D., Ali, M. & Moro, R. (2017). Resource management in big data initiatives: Processes and dynamic capabilities. *Journal of Business Research, 70*(C), pp. 328-337.

Brown, R. B. (2006). *Doing your dissertation in business and management: The reality of researching and writing*. London: SAGE.

Castelluccio, M. (2017). The far limits of artificial intelligence. *Strategic Finance, 99*(6), p. 55.

Chae, H., Koh, C. & Prybutok, V. (2014). Information Technology Capability and Firm Performance: Contradictory Findings and Their Possible Causes. *MIS Quarterly, 38*(1), p. 305.

Chan, J. (2014). Big Data Customer Knowledge Management. *Communications of the IIMA, 14*(3/4), pp. 45-55.

Chen, H., Chiang, R. H. L. & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly, 36*(4), pp. 1165-1188.

Chen, X. & Lin, X. (2014). Big Data Deep Learning: Challenges and Perspectives. *IEEE Access, 2*, pp. 514-525.

Cheng, F. & Wang, Y. (2018). The Do Not Track mechanism for digital footprint privacy protection in marketing applications. *Journal of Business Economics and Management, 19*(2), pp. 253-267.

Chierici, R., Mazzucchelli, A., Garcia-Perez, A. & Vrontis, D. (2019). Transforming big data into knowledge: The role of knowledge management practice. *Management Decision, 57*(8), pp. 1902-1922.

Choi, B. & Lee, H. (2002). Knowledge management strategy and its link to knowledge creation process. *Expert Systems With Applications, 23*(3), pp. 173-187.

Chong, A. Y. L., Ch'ng, E., Liu, M. J. & Li, B. (2017). Predicting consumer product demands via Big Data: The roles of online promotional marketing and online reviews. *International Journal of Production Research, 55*(17), pp. 5142-5156.

Clayton, N. (2018). How Supercell uses machine learning to automate monetisation in Clash Royale. *Pocket Gamer.* [online document] [Accessed 10th February 2020] Available: https://www.pocketgamer.biz/news/68576/supercells-jarno-seppnen-on-how-clash-royale-uses-machine-learning-automate-monetization/

Cohen, W. M. & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly, 35*(1), p. 128.

Coval, T. (2018). Artificial Intelligence, Bias & Your Business. *Journal of Property Management, 83*(2), pp. 6-9.

Cravens, D. W., & Piercy, N. (2013). Strategic marketing. 10th edition. New York: McGraw-Hill.

Day, G. (1994). The capabilities of market-driven organizations. *Journal of Marketing 58(4)*, pp. 37–52. *The Journal of Product Innovation Management, 12*(3), pp. 257-258.

Day, G. S. (2014). An outside-in approach to resource-based theories. Journal of the Academy of Marketing Science, 42(1), 27–28.

Del Vecchio, P., Secundo, G. & Passiante, G. (2018). Analyzing Big Data through the lens of customer knowledge management. *Kybernetes, 47*(7), pp. 1348-1362.

Delen M. (2014). Real-world data mining: applied business analytics and decision making. Upper Saddle River: FT Press.

Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C. & Papadopoulos, T. (2019). Big Data and Predictive Analytics and Manufacturing Performance: Integrating Institutional Theory, Resource-Based View and Big Data Culture. *British Journal of Management, 30*(2), pp. 341-361.

Duffy, J. (2000). Something funny is happening on the way to knowledge management. *Information Management Journal, 34*(4), pp. 64-67.

Eisenhardt, K. (1989). Building Theories From Case Study Research. *Academy of Management. The Academy of Management Review, 14*(4), p. 532.

Ekambaram, A., Sørensen, A. Ø., Bull-Berg, H. & Olsson, N. O. (2018). The role of big data and knowledge management in improving projects and project-based organizations. *Procedia Computer Science, 138*, pp. 851-858.

Erevelles, S., Fukawa, N. & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research, 69*(2), pp. 897-904.

Ferraris, A., Mazzoleni, A., Devalle, A. & Couturier, J. (2019). Big data analytics capabilities and knowledge management: Impact on firm performance. *Management Decision, 57*(8), pp. 1923-1936.

Fosso Wamba, 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*(C), pp. 234-246.

Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R. & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research, 70*(C), pp. 356-365.

Fuchs, M., Höpken, W. & Lexhagen, M. (2014). Big data analytics for knowledge generation in tourism destinations – A case from Sweden. *Journal of Destination Marketing & Management, 3*(4), pp. 198-209.

Go, E. & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior, 97*, pp. 304-316.

Gore, C. & Gore, E. (1999). Knowledge management: The way forward. *Total Quality Management, 10*(4-5), pp. S554-S560.

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

Grant, R. M. (2010) Contemporary Strategy Analysis: Text & Cases. 7th edition. John Wiley & Sons.

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B. & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research, 70*(C), pp. 308-317.

Gupta, M. & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management, 53*(8), pp. 1049-1064.
Heimbach, I., Kostyra, D. & Hinz, O. (2015). Marketing Automation. *Business & Information Systems Engineering, 57*(2), pp. 129-133.

Hirsjärvi, S., Remes, P., Sajavaara, P. & Sinivuori, E. (2009). *Tutki ja kirjoita* (15. uud. p.). Helsinki: Tammi.

Kaisler, S., Armour, F., Espinosa, J. A. & Money, W. (2013). *Big Data: Issues and Challenges Moving Forward*. System Sciences (HICSS). 46th Hawaii International Conference On, IEEE (2013) 995–1004.

Kotler, P., & Keller, K. L. (2015). Marketing Management. Upper Saddle River: Prentice Hall.

Kozlenkova, I. V., Samaha, S. A., & Palmatier, R. W. (2014). Resource-based theory in marketing. Journal of the Academy of Marketing Science, 42(1), 1–21.

Lamont, J. (2012). Big data has big implications for knowledge management. *KM World, 21*(4), pp. 8-11.

Laurent, G. (2013). EMAC Distinguished Marketing Scholar 2012: Respect the data! *International Journal of Research in Marketing, 30*(4), pp. 323-334.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. & Kruschwitz, N. (2011). Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Management Review, 52*(2), pp. 21-32.

Leeflang, P. S., Verhoef, P. C., Dahlström, P. & Freundt, T. (2014). Challenges and solutions for marketing in a digital era. *European Management Journal, 32*(1), pp. 1-12.

Lusch, R. F., Vargo, S. L. & O'Brien, M. (2007). Competing through service: Insights from service-dominant logic. *Journal of Retailing, 83*(1), pp. 5-18.

Manyika, M., Chui, B ., Brown, J., Bughin, R ., Dobbs, C. & Roxburgh, A. H. (2011) Big Data: The Next Frontier for Innovation, Competition, and Productivity, M.G. Institute.

Marinova, D., de Ruyter, K., Huang, M., Meuter, M. L. & Challagalla, G. (2017). Getting Smart: Learning From Technology-Empowered Frontline Interactions. *Journal of Service Research, 20*(1), pp. 29-42.

Mata, F., Fuerst, W. & Barney, J. (1995). Information technology and sustained competitive advantage: A resource-based analysis. *MIS Quarterly, 19*(4), p. 487.

Mata, F., Fuerst, W. & Barney, J. (1995). Information technology and sustained competitive advantage: A resource-based analysis. *MIS Quarterly, 19*(4), p. 487.

Mathews, C. & Wearn, N. (2016). How Are Modern Video Games Marketed? *The Computer Games Journal, 5*(1), pp. 23-37.

McAfee, A., Brynjolfsson, E., Davenport, T. & Patil, D. J. (2012). Strategy & Competition Big Data: The Management Revolution. *Harvard Business Review, 90*(10), pp. 60

Meso, P. & Smith, R. (2000). A resource-based view of organizational knowledge management systems. *Journal of Knowledge Management, 4*(3), pp. 224-234.

Metsämuuronen, J. (2006). *Laadullisen tutkimuksen käsikirja*. Helsinki: International Methelp.

Miklosik, A., Kuchta, M., Evans, N. & Zak, S. (2019). Towards the Adoption of Machine Learning-Based Analytical Tools in Digital Marketing. *IEEE Access, 7*, pp. 85705-85718.

Mishra, D., Luo, Z., Hazen, B., Hassini, E. & Foropon, C. (2019). Organizational capabilities that enable big data and predictive analytics diffusion and organizational performance. *Management Decision, 57*(8), pp. 1734-1755.

Mithas, S., Lee, M. R., Earley, S., Murugesan, S. & Djavanshir, R. (2013). Leveraging Big Data and Business Analytics [Guest editors' introduction]. *IT Professional, 15*(6), pp. 18-20.

Moro, S., Rita, P. & Vala, B. (2016). Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. *Journal of Business Research, 69*(9), pp. 3341-3351.

Neogames. (2018). The Game Industry of Finland Report 2018. [online document] [Accessed 12th February 2020] Available: http://www.neogames.fi/wp-content/uploads/2019/04/FGIR-2018-Report.pdf

Newzoo. (2019). Global games market report 2019. [online document] [Accessed 12th February 2020] Available: https://resources.newzoo.com/hubfs/2019_Free_Global_Game_Market_Report.pdf?utm_campaign=Games%20Market%20Report&utm_source=hs_automation&utm_medium=email&utm_content=76474808&_hsenc=p2ANqtz-8R78qpVNcMMR4HsaRGDfZt5JgZ6xhDzFrIhePQr3S-2bzDYcBnSqdX-HQBMyIj-8cQTgIrVLEfcchKrJtm9qrETn95kQ&_hsmi=76474808

Nonaka, I. (1991). The Knowledge-Creating Company. *Harvard Business Review, 69*(6), p. 96.

Nonaka, I., Toyama, R. & Konno, N. (2000). SECI, Ba and Leadership: A Unified Model of Dynamic Knowledge Creation. *Long Range Planning, 33*(1), pp. 5-34.

O'Dell, C. & Grayson, C. (1998). If only we knew what we know: Identification and transfer of internal best practices. *California Management Review, 40*(3), pp. 154-174.

O'Connor, C. & Kelly, S. (2017). Facilitating knowledge management through filtered big data: SME competitiveness in an agri-food sector. *Journal of Knowledge Management, 21*(1), pp. 156-179.

Obitade, P. (2019). Big data analytics: A link between knowledge management capabilities and superior cyber protection. *Journal of Big Data, 6*(1), pp. 1-28.

Oliveira, F., Santos, A., Aguiar, B. & Sousa, J. (2014). GameFoundry: Social Gaming Platform for Digital Marketing, User Profiling and Collective Behavior. *Procedia - Social and Behavioral Sciences, 148*(C), pp. 58-66.

Pauleen, D. J. & Wang, W. Y. (2017). Does big data mean big knowledge? KM perspectives on big data and analytics. *Journal of Knowledge Management, 21*(1), pp. 1-6.

Pentland, B. T. (1995). Information systems and organizational learning: The social epistemology of organizational knowledge systems. Accounting, Management and Information Technologies, 5(1), pp. 1-21.

Peteraf, M. (1993). The cornerstones of competitive advantage - a resource-based view. *Strategic Management Journal, 14*(3), pp. 179-191.

Raghupathi, W. & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health information science and systems, 2*(1), p. 3.

Rands, K. (2018). How big data is disrupting the gaming industry. *CIO.* [online document] [Accessed 12th February 2020] Available: https://www.cio.com/article/3251172/how-big-data-is-disrupting-the-gaming-industry.html

Romano, A., Passiante, G., Vecchio, P. & Secundo, G. (2014). The innovation ecosystem as booster for the innovative entrepreneurship in the smart specialisation strategy. *International Journal of Knowledge-Based Development, 5*(3), pp. 271-288.

Ross, J., Beath, C. & Quaadgras, A. (2013). You May Not Need Big Data After All. *Harvard Business Review, 91*(12), pp. 90-98.

Sadowski, J., (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society, 6*(1)

Shamim, S., Zeng, J., Shariq, S. M. & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information & Management, 56*(6)

Saunders, M., Lewis, P. & Thornhill, A. (2016). *Research methods for business students* (Seventh edition.). Harlow, Essex: Pearson Education.

Shields, M. (2015). Mobile Game Companies Are Starting to Run Ads on Instagram. *The Wall Street Journal.* [online document] [Accessed 10th February 2020] Available: https://www.wsj.com/articles/mobile-game-companies-are-starting-to-run-ads-on-instagram-1444903202

Simon, J. P. (2018). Triggering the emergence of digital ecosystems: The role of mobile and video games in emerging economies. *Digital Policy, Regulation and Governance, 20*(5), pp. 449-478.

Stafford, P. (2019). The dangers of in-game data collection. *Polygon.* [online document] [Accessed 10th February 2020] Available: https://www.polygon.com/features/2019/5/9/18522937/video-game-privacy-player-data-collection

Sumbal, M. S., Tsui, E. & See-to, E. W. (2017). Interrelationship between big data and knowledge management: An exploratory study in the oil and gas sector. *Journal of Knowledge Management, 21*(1), pp. 180-196.

Sundsøy, P., Bjelland, J., Iqbal, A., Pentland, A. & De Montjoye, Y. (2014). Big data-driven marketing: How machine learning outperforms marketers' gut-feeling. *Social Computing, Behavioral-Cultural Modeling and Prediction, 8393*, pp. 367-374.

Teece, D. (2014). The foundations of enterprise performance: dynamic and ordinary capabilities in an (economic) theory of firms. *The Academy of Management Perspectives, 28*(4), p. 328.

Teece, D. (2015). *Intangible assets and a theory of heterogeneous firms*. Springer 217-239.

Teece, D. J., Pisano, G. & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), pp. 509-533.

Tettamanzi, A., Carlesi, M., Pannese, L. & Santalmasi, M. (2007). Business intelligence for strategic marketing: Predictive modelling of customer behaviour using fuzzy logic and evolutionary algorithms. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 4448*, pp. 233-240.

Trainor, K. J., Andzulis, J. M., Rapp, A. & Agnihotri, R. (2014). Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM. *Journal of Business Research, 67*(6), pp. 1201-1208.

Vilkka, H. (2015). *Tutki ja kehitä* (4., uudistettu painos.). Jyväskylä: PS-kustannus.

Waller, M. A. & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics, 34*(2), pp. 77-84.

Waller, T., Hockin, R. & Smith, G. (2017). Marketing Strategies of Mobile Game Application Entrepreneurs. *International Journal of Applied Management and Technology, 16*(1).

Wang, Y., Kung, L., Wang, W. Y. C. & Cegielski, C. G. (2018). An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management, 55*(1), pp. 64-79.

Ward, M. J., Marsolo, K. A. & Froehle, C. M. (2014). Applications of business analytics in healthcare. *Business Horizons, 57*(5), pp. 571-582.

Wernerfelt, B. (1984). A Resource-based View of the Firm. *Strategic Management Journal (pre-1986), 5*(2), p. 171.

Xie, K., Wu, Y., Xiao, J. & Hu, Q. (2016). Value co-creation between firms and customers: The role of big data-based cooperative assets. *Information & Management, 53*(8), pp. 1034-1048.

Xu, Z., Frankwick, G. L. & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research, 69*(5), pp. 1562-1566.

Yin, R. (1981). The Case Study Crisis: Some Answers. *Administrative Science Quarterly, 26*(1), p. 58.

Zerbino, P., Aloini, D., Dulmin, R. & Mininno, V. (2018). Big Data-enabled Customer Relationship Management: A holistic approach. *Information Processing and Management, 54*(5), pp. 818-846.

Zhao, J., Fan, S. & Hu, D. (2014). Business challenges and research directions of management analytics in the big data era. *Journal of Management Analytics, 1*(3), pp. 169-174.

APPENDICES

Appendix 1: Interview questions

General / big data analytics

- Briefly describe your role in the company
- How would describe your company's approach to data-driven marketing?
- What about the role of big data in marketing?
    - How is data used in marketing? Why is it used?
    - What kind of benefits or challenges does your company face with data-driven marketing?

Organizational resources

- What do you think are the most important organizational resources that enable/facilitate data-driven marketing?
    - What kind of tangible resources are needed/used?
    - What kind of human resources are needed/used?
    - What kind of intangible resources are needed/used?
- Why are these resources important, what makes them valuable?

Knowledge management

- How would you describe the process of turning raw data into marketing insights?
- How is data generated?
    - What are its sources? How is data captured?
- How do you manage the knowledge gained from data?
- How is the data analyzed?
    - Do you utilize any kinds of intelligent analytical tools in your data analysis?
        - If so, how are they used? How do you see their role in data-driven marketing?
- How are the insights utilized?
- How do the insights gained from big data affect marketing / business?