

FACTORS AFFECTING THE SUCCESS OF AI CAMPAIGNS IN MARKETING: DATA PERSPECTIVE

**Jyväskylä University
School of Business and Economics**

Master's Thesis

2020

**Author: Eevi Varmavuo
Subject: Digital Marketing and Corporate Communication
Supervisor: Heikki Karjaluoto**



**JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ**

ABSTRACT

Author Eevi Varmavuo	
Title Factors affecting the success of AI campaigns in marketing: Data perspective	
Subject Digital Marketing and Corporate Communication	Type of work Master's Thesis
Date 28.7.2020	Number of pages 60+1
Abstract <p>Artificial intelligence (AI) has been reshaping marketing in many ways during recent years. While AI technologies and the rise of big data have enabled innovative ways to practice marketing, the industry is still rather young. The applications of AI in marketing have been of interest among researchers and industry leaders, yet prior research focus provided little knowledge on what type of data marketers should rely on to reach successful outcomes of AI campaigns.</p> <p>This study aims to offer knowledge on the data related factors that determine what type of data marketers should rely on to ensure successful delivery of AI campaigns. The empirical data was collected through thirteen expert interviews.</p> <p>This study reinforces the existing literature to great extent, but also provides new perspectives to the identified factors. The findings of this study show that data used for AI campaigns in marketing should be clean, reliable and of high quality to ensure outcomes are effective, accurate and unbiased. Internal data should be prioritized and complemented with external data such as social media data to gain deeper insights and better predictions. Combining both internal and external data sets is identified as best practice to run AI campaigns. Additionally, the business goals and wanted outcomes of AI campaigns should be placed at the center of any data collection, management and analysis process to ensure successful results of AI practices. Finally, transparency in data related matters was found important as it builds trust and ensures customers provide accurate personal data for marketing purposes.</p>	
Key words Artificial intelligence, machine learning, data-driven marketing, digital marketing	
Place of storage Jyväskylä University School of Business and Economics	

TIIVISTELMÄ

Tekijä Eevi Varmavuo		
Työn nimi Tekoälyä hyödyntävien kampanjoiden menestymiseen vaikuttavat tekijät markkinoinnissa: data näkökulma		
Oppiaine Digitaalinen markkinointi ja yritysviestintä	Työn laji Pro gradu -tutkielma	
Päivämäärä 28.7.2020	Sivumäärä 60+1	
<p>Tiivistelmä</p> <p>Tekoäly on tuonut suuria muutoksia markkinoinnin käytäntöihin viime vuosina. Vaikkakin tekoälyteknologiat sekä <i>big data</i> ovat mahdollistaneet uusia innovatiivisia ratkaisuja markkinoinnin toteuttamiseen, ala on vielä alkutekijöissään. Tekoälyä hyödyntävä markkinointi on ollut sekä tutkijoiden että ammattilaisten suurennuslasin alla viime vuosina. Aiemmat tutkimukset ovat kuitenkin tarjonneet vasta hyvin vähän ymmärrystä siihen, millaista dataa markkinoijien tulisi käyttää tekoälyä hyödyntävää markkinointia toteuttaessaan, jotta parhaat mahdolliset tulokset voidaan saavuttaa.</p> <p>Tämän tutkimuksen tavoitteena on tutkia millaista dataa markkinoijien tulisi käyttää tekoälyä hyödyntävässä markkinoinnissa onnistuakseen tämän toteutuksessa. Tavoitteena on lisäksi tarjota ymmärrystä tekijöistä, jotka vaikuttavat näihin datan ominaisuuksiin. Empiirisen datan keräämiseksi toteutettiin kolmetoista asiantuntijahaastattelua.</p> <p>Tämä tutkimus tukee pitkälti aiempaa kirjallisuutta, mutta tarjoaa lisäksi uusia näkökulmia datan tarvittaviin ominaisuuksiin. Tutkimuksen tulokset osoittavat, että datan tulisi olla puhdasta, luotettavaa ja laadukasta, jotta tulokset ovat tehokkaita, täsmällisiä ja puolueettomia. Sisäistä dataa tulisi priorisoida ja täydentää ulkoisilla tietolähteillä, kuten sosiaalisen median datalla. Näin voidaan saavuttaa syvempi <i>insight</i> ja parempi ennustus toteutetusta kampanjasta. Tulosten mukaan sisäisten ja ulkoisten tietolähteiden yhdistäminen on palkitsevin toimintatapa. Liiketoiminnan tavoitteet tulisi asettaa keskiöön datan keräämiseen, hallinnointiin ja analysointiin liittyvissä tehtävissä. Lisäksi läpinäkyvyys dataan liittyvissä asioissa tunnustetaan tärkeäksi, jotta asiakkaat luovuttavat todenmukaista henkilökohtaista dataa itsestään markkinointitarkoituksiin.</p>		
Asiasanat Tekoäly, koneoppiminen, datalähtöinen markkinointi, digitaalinen markkinointi		
Säilytyspaikka Jyväskylän	yliopiston	kauppakorkeakoulu

LIST OF FIGURES AND TABLES

FIGURE 1. Structure of the study	10
FIGURE 2. Framework.....	30
FIGURE 3. Final framework	52
TABLE 1. Data used for AI driven marketing.....	23
TABLE 2. Summary of the interviews.....	34

CONTENTS

ABSTRACT

FIGURES AND TABLES

CONTENTS

1	INTRODUCTION.....	8
1.1	Background of the study	8
1.2	Research questions and objectives	9
1.3	Structure of the research.....	10
2	AI IN MARKETING.....	12
2.1	Defining AI.....	12
2.1.1	Artificial intelligence (AI).....	12
2.1.2	Machine learning (ML)	13
2.1.3	Deep learning.....	14
2.1.4	Big Data.....	14
2.2	Current market situation.....	15
2.3	Benefits and current applications of AI in marketing.....	16
2.3.1	Targeted marketing.....	17
2.3.2	Personalization and recommendations.....	17
2.3.3	Media optimization.....	19
2.3.4	Customer experience	19
2.4	The role of data in AI-driven marketing.....	20
2.4.1	The role of data	21
2.4.2	What data is used?	22
2.4.3	Data collection and management.....	25
2.4.4	Data privacy issues.....	26
2.4.5	Data bias issues.....	29
2.5	Summary of the literature review	29
3	DATA AND METHODOLOGY	33
3.1	Qualitative research	33
3.2	Data collection methods	33
3.3	Data analysis	35
4	RESULTS AND ANALYSIS.....	37
4.1.	Data source.....	37
4.1.1	Role of internal data and external data	37
4.1.2.	Role of third-party data	38

4.2 Data collection and management.....	41
4.3 Data privacy.....	45
4.4 Data cleanness and reliability.....	47
5 DISCUSSION	51
5.1 Theoretical contributions	51
5.2 Managerial implications.....	54
5.3 Limitations of the research.....	55
5.4 Future research suggestions	56
REFERENCES	57

APPENDICES

1 INTRODUCTION

In this chapter, the background of the study, the research questions and objectives as well as the structure of the study are discussed.

1.1 Background of the study

Existing literature on artificial intelligence (AI) and its relation to marketing practices advocate further research on the topic. AI has existed as a technology and a field of study for decades, and it has recently made impressive impacts in various industries, especially within the field of marketing (Campbell et al., 2020). AI technology is not only improving existing marketing practices but is expected to reshape marketing as it is today entirely (Vishnoi et al., 2018). Furthermore, adopting AI into marketing practices has been seen as an effective way to reach competitive advantage (Chui, 2017) and the existence of a strong digital base is believed to be a factor in determining whether a company can invest in AI (Campbell et al., 2020). Therefore, having strong digital foundations in place is identified as an important actor in enabling the practices that AI offers for marketers. As a result of the recent technological advances, marketers are now phasing an era where understanding data related matters is vital (Sleep, Hulland & Gooner, 2019) and thus, research contributing to marketers' understanding of the phenomenon is identified to be important.

Data is the bedrock of AI and as there is more data available today than ever before, it offers marketers monumental opportunities (Alshura, Zabadi & Abughazaleh, 2018). Big data offers greater insights into marketing performance than have previously been possible, and thus helps marketers make successful decisions in optimizing their marketing actions and improving their return on investments (Wedell & Kannan, 2018). Marketers are leveraging AI for targeted marketing, personalization and recommendations, media optimization and improving their customer experience among other use cases. While finding data used to be the challenge for marketers in the past, today they are facing a new challenge in forms of transforming big data into something valuable (Manyika et al., 2011). As big data has attained a central role in marketing solutions, Alshura et al. (2018) suggest further research on the fundamental requirements of big data in designing these marketing solutions. Additionally, while Campbell et al. (2020) introduce various data sources that are used for AI campaigns in marketing, the existing literature does not further look into what data sources create best practices for these campaigns. Furthermore, while the value of social media data among other external sources are highlighted in existing literature, the role of third-party data as such is not touched upon in terms of its relevance in successful delivery of AI campaigns.

As significant amounts of data become available daily, marketers are leveraging this data to create more personalized experience for their customers (Ozcelik & Varnali, 2019). While personal data enables marketers to offer better customer experience through various marketing activities, collecting customer data has raised concerns about data privacy matters among consumers (Davenport et al., 2020; Martin et al., 2017). In addition, regulations such as GDPR are seen as significant game changers in data privacy regulations (Sterne, 2017) and are believed to create challenges for marketers (Kietzmann et al., 2018). While existing literature discusses the data privacy concerns of consumers as well as data privacy matters in general, the relevance of these matters in relation to defining what data is appropriate for effective AI practices is not covered. Future research related to data privacy is another area encouraged in existing literature (Davenport, Guha, Grewal & Bressgott, 2020). Another recommended area of future research provided in existing literature is data bias (Davenport et al., 2020). Algorithm bias is identified as a common output of AI systems if the data sets contain any information that can expect biased output and if they are not trained correctly (Chui, 2017; Davenport et al., 2020). Existing literature briefly discusses the role of data bias in AI-driven marketing practices but does not cover its role in determining what data should be used for these practices.

Moreover, future research into factors that impact AI adoption is suggested in the existing literature (AlSheibani, Cheung & Messom, 2018). Thus, through existing literature and the suggested topics of research, knowing what data is required and understanding the underlying factors related to data that enables successful delivery of AI campaigns can be identified as an important research area. This research will contribute to the research gaps identified by offering marketers the understanding and guidelines to what type of data should be used with AI campaigns in marketing for them to be successful as well as providing understanding of the underlying factors involved with data. The research questions and objectives are presented in the following chapter.

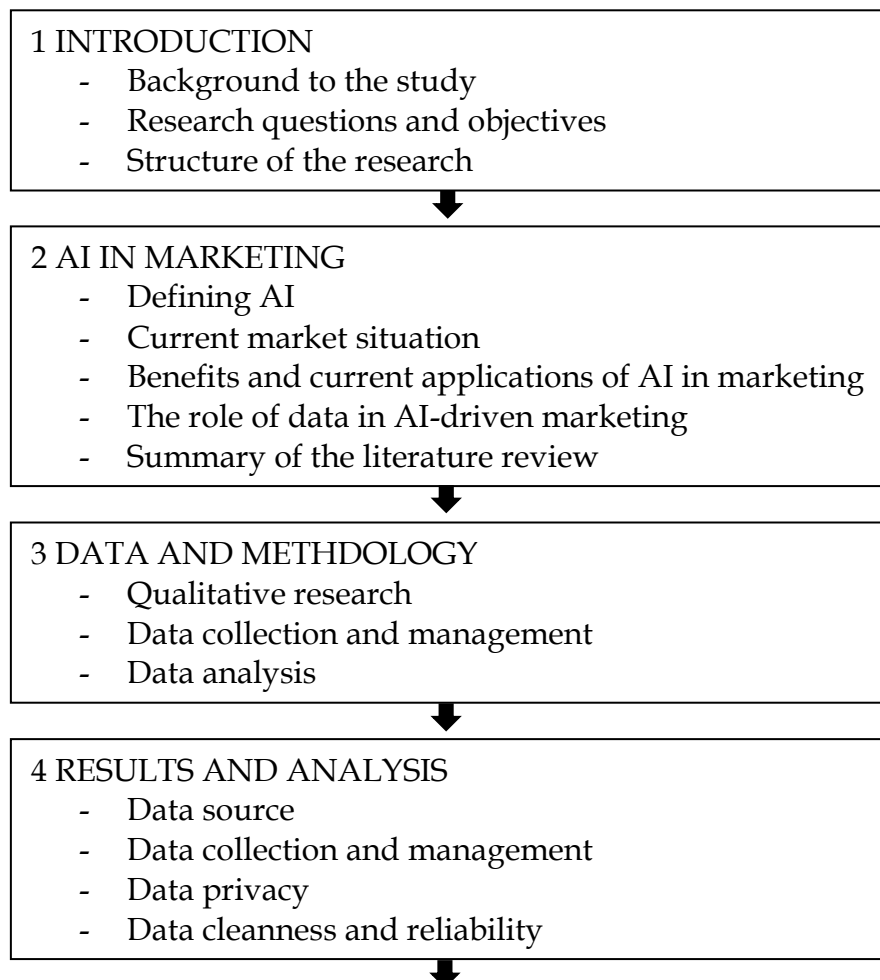
1.2 Research questions and objectives

This study aims to find out what type of data is required to run effective AI campaigns in marketing. As companies are investing in AI at an increasing rate, the data collection methods and management related aspects are also studied to ensure marketers know whether they are ready to adopt AI into their marketing practices or not. Furthermore, the study touches upon the role of third-party data in AI-driven marketing and considers the risks related to the use of data such as privacy and data bias issues. The main research question and two sub-questions are:

- What type of data is required to run effective AI campaigns in marketing?
 - How should this data be collected and managed?
 - What is the role of third-party data?

1.3 Structure of the research

This research is divided into 5 chapters as shown in Figure 1 below. Chapter 1 introduces the background to the research topic. Chapter 2 consists of a comprehensive literature review which contributes as the base for the theoretical framework. More specifically, chapter 2 defines AI, introduces the current market situation of AI in relation to business and marketing, reviews the most significant AI applications currently used in marketing, discusses the role of data in AI-driven marketing and finally presents a summary of the literature review. Chapter 3 elaborates the methodological approach to the research. Chapter 4 presents the findings of the study, which is followed by a presentation of the theoretical and managerial contributions and future research suggestions in chapter 5.



5 DISCUSSION

- Theoretical contributions
- Managerial implications
- Limitations of the research
- Future research suggestions

FIGURE 1. Structure of the study

2 AI IN MARKETING

To comprehend the use of AI in marketing, understanding the general concept of AI is crucial. Thus, this chapter first provides a definition of AI for a marketer. Next, the current market situation of AI is examined and the benefits and current applications of AI for marketing purposes are discussed. Further, the role of data is explained and finally, a summary of the literature review is presented.

2.1 Defining AI

2.1.1 Artificial intelligence (AI)

AI was first initiated at a conference at Dartmouth College in 1956 (Levesque, 2017). AI was then defined as a program with common sense (Levesque, 2017). Much like its initial definition, today AI is defined as machines, algorithms, programs or systems that demonstrate intelligence (Shankar, 2018). The intelligence that these AI machines show mimics human intelligence in forms of problem-solving, reasoning, learning, perceiving and acting (Huang & Rust, 2018). In simpler terms, AI technology makes it possible for machines to learn and to perform tasks close to human-like behavior (Campbell, Sands, Ferraro, Tsao & Mavrommatis, 2020).

AI can be thought of as a broad umbrella that has a broad range of key technologies and applications as subcategories (Jarrahi, 2018). Machine learning, deep learning, rule-based expert systems and neural networks all fall under the AI umbrella (Davenport et al., 2020). Other applications like natural language processing, robotics, visual recognition and affective computing are also subcategories of AI (Sterne, 2017). Looking at AI, there is a fundamental distinction that needs to be addressed. There are two types of AI, strong AI and weak AI (Shahid & Li, 2019). Strong AI, also known as artificial general intelligence, has intelligence in more than one domain (Shadid & Li, 2019). It is the type of AI that mimics humans to the extent that it can be comparable to the human brain and thus, allegedly, could become self-aware (Sterne, 2017). Weak AI, also known as artificial narrow intelligence, focuses on performing a specific task (Shadid & Li, 2019). Unlike strong AI, it is not capable of learning to extend into new domains (Davenport et al., 2020). According to Sterne (2017), this type of weak AI has been seen in business areas such as marketing for quite some time now. In fact, most AI technologies today are weak (AlSheibani, Cheung & Messom, 2018). A few examples of using weak AI are Pandora offering users music recommendations based on previous playlist history and Amazon offering product recommendations based on past purchases (Sterne, 2017). When referring to AI, this thesis focuses on artificial narrow intelligence.

Although AI mimics humans by demonstrating intelligence like humans do, machines can also perform many tasks impossible for humans to fulfill

(Kaplan, 2016). AI is used for example to enhance the intelligence of products, services or solutions (Shankar, 2018). In such cases, AI and humans can complement each other's abilities. AI can provide better performance through addressing complexity over human capabilities, and humans can offer a more holistic, intuitive approach than a machine can perform (Jarrahi, 2018). AI thrives to perform with intelligence and error-free decision-making (Vishnoi, Bagga, Sharma & Wani, 2018). Thus, AI minimizes human error and maximizes accuracy. Although with the use of AI, marketers can process large and complex pools of data and find valuable information in the data, it will probably not be able to determine when a certain information will become valuable. Hence, AI-powered solutions help marketers to perform better by analyzing the data faster and with more accuracy and provide better insights, but the creative and strategic thinking of a marketer still plays a key role in determining when a certain information becomes valuable and appropriate to use. (Kietzmann & Pitt, 2020.)

2.1.2 Machine learning (ML)

Machine learning (ML) is a subset of AI (Davenport et al., 2020) and is the most commonly used analytical AI application today (Huang & Rust, 2018). Instead of following a specific set of predefined rules, ML aims to learn from the available data and has thus shifted the role of algorithms used in AI (Jarek & Mazurek, 2019). In simple terms, ML operates with the data and tasks that it is given and learns from its findings (Levesque, 2017). When performing the given task, ML algorithms learn by identifying patterns in the data and making sense of them (Sterne 2017). Eventually conclusions can be made from the analyses and insights drawn by ML (Jarek & Mazurek, 2019). Recognizing patterns and trends in data makes ML an important ally for marketers (Gentsch, 2019). Campbell et al. (2020) agree by discussing ML to be a valuable tool for analysing large data sets as it can offer marketers rich and new insights to consumer behavior and enable marketers to enhance their operations based on those insights. There are multiple forms of ML, such as pattern recognition and data analytics (Jarek & Mazurek, 2019). ML approaches are often used in functions related to marketing such as web searches, content filtering, recommendation systems and speech recognition (Campbell et al., 2020).

The popularity of ML in marketing can be explained by the system's ability to collect consumer data from multiple sources, store memories, remember them, and can learn from its historic data and problem-solving experiences. Thus, it can think logically and propose the best options for consumers' needs much more efficiently than a human could do. Additionally, ML can help marketers to make predictions on for example customer lifetime value, the likelihood of conversions and consumer insights, and thus increase the efficiency of marketing practises. (Kietzmann, Paschen & Treen, 2018.)

There are three types of ML, all of which are used for marketing purposes. The most common method of ML is supervised learning. Chatbots are a common form of AI-powered marketing solutions and an example of the supervised learning approach. With chatbots, supervised ML can learn what the most

frequent customer queries are and it can learn and train spam filters to identify spam emails. Unsupervised ML is another approach of ML. Instead of making predictions like supervised ML, it focuses on simply analyzing the data and finding structure in that data. Unsupervised ML is often used by marketers to segment customers and markets. Although AI and ML require data to run, one form of ML does not require existing data sets to work. Reinforcement ML algorithm learns from the actions it takes and creates an evaluations of the outcomes. Thus, the system learns while building its own data set. Facebook advertising is one example of such reinforcement ML. The system can modify any details of the ad such as geographic location, audience and placements by learning from its actions and success and failure results. Reinforcement ML is also used in recommendation systems. (Campbell et al., 2020.)

2.1.3 Deep learning

Deep learning is a subset of machine learning (Vishnoi et al., 2018). The algorithms used in deep learning are inspired by the structure and function of the human brain (Siau & Yang 2017) and in terms of problem-solving and decision-making it is very similar to the brain activity of a human (Jarek & Mazurek, 2019). Deep learning algorithms do not require manual management at all (Jarek & Mazurek, 2019). The algorithms rely on big data to offer new information in an instant (Jarek & Mazurek, 2019). Deep learning algorithms are often used for marketing purposes and they can, to a limited accuracy, get insight to what a person is thinking when for example going through images in an online shop (Shankar, 2018).

2.1.4 Big Data

As business processes develop and become more and more complex and widespread, the more data is collected and generated by organizations. The extensive amount of data that exceeds the traditional database technologies is referred to as big data (Vishnoi et al., 2018). Big data surpasses the potential of these traditional technologies or database software tools in collecting, storing, managing and analyzing data (Alshura et al., 2018). Big data refers to data that includes the “four V’s”; volume, velocity, variety and veracity (Martin & Murphy, 2017). Volume and velocity of data are relevant from a computing point of view whereas variety and veracity are relevant from an analytics point of view (Wedell & Kannan, 2016). With these four V’s, big data creates big opportunities to marketers.

The development of the field of big data has enabled marketers to collect and analyse greater amounts of data than ever before and transform that data into valuable insight and ultimately, a strategy they can follow (Campbell et al., 2020). Big data has a big role for organizations and is considered to be a game changer in terms of gaining competitive advantage (Wright, Robin, Stone & Aravopoulou, 2019; Manyika et al., 2011). With the help of technology, these high volumes of data can be processed and analyzed with minimal manual work

(Vishnoi et al., 2018). Using big data thus opens new possibilities for organizations in terms of analysing their current situation, predicting future outcomes of their strategies and identifying new innovative ways of doing business (Wright, Robin, Stone & Aravopoulou, 2019). Big data has created enormous amounts of opportunities and has made a huge impact on marketing practices by providing faster and cheaper insights (Alshura et al., 2018). Due to its big impact and revolutionary status in marketing, Alshura et al. (2018) call big data the “heart of data-driven marketing”.

2.2 Current market situation

Although AI has existed as a technology and a field of study already for decades, it has only recently shown rapid growth among different markets. AI is one of those technologies that disrupt a variety of industries (Campbell et al., 2020). It comes as no surprise AI has boomed in recent years as the technology has taken massive steps in its development, algorithms are much more sophisticated and there is more data generated daily than ever before (Chui, 2017). Recent surveys show a nearly 25 percent increase in the use of AI by firms each year (Cam, Chui & Hall, 2019). The growth of AI adoption and positive effects as results for businesses can be seen through estimations that predict AI to increase the global economy by 14 percent (equivalent to 15.7 trillion USD) by 2030 (AlSheibani, Cheung & Messom, 2018). By investing in AI, firms may have the possibility to increase their return on investment with great volumes, which further accelerates AI adoption from firms. This is supported by the fact that the number of AI start-ups is increasing at a rapid rate (Wirth, 2018) and companies are investing billions of dollars in AI (Levesque, 2017). During 2013 and 2016, AI companies' fundings saw an annual growth rate of over 80 percent and external investments saw an annual growth rate of nearly 40 percent (Chui, 2017). The market of big data has also been expected to reach 9.4 billion USD in 2020, compared to its 1.7 billion USD status in 2016 (Alshura, Zabadi & Abughazaleh, 2018).

The disruption potential that AI brings to new entrants is tremendous as it opens up opportunities to take over incumbents (Chui, 2017). Many companies, like Uber in the taxi industry, have already managed to do so (Chui, 2017). Big players like Google, Amazon and Apple are also racing to reach competitive advantage and market share with their use of AI (Al Sheibani et al., 2018). These tech giants are not only buying AI-powered solutions but AI startups in order to ensure they have the most qualified talent available in addition to the technology itself (Chui, 2017). This behavior shows the importance of this technology for businesses. These large companies are working with AI in various purposes such as robotics, speech recognition, virtual assistants and machine learning (Chui, 2017).

The field of marketing has seen a big change since the rise of AI adoption from firms. According to a survey by Accenture in 2017, 85 percent of executives were planning on investing heavily in AI over the next three years (Jarrahi 2018).

Although companies are investing large amounts in AI, there is still uncertainty among business leaders regarding the benefits of AI for their business, where to get these AI-powered solutions from and how to integrate them, as well as how to predict the potential return on investment from this technology (Chui, 2017). Surveys show that some managers are in fact not moving forward with AI as they are still unsure how their business can benefit from AI adoption (Bughin, Chui & McCarthy, 2017). Although, AI is considered to be an important application within marketing processes in order to keep up with the competition (Harvard Business Review, 2020). Another study by PWC supports the importance of AI adoption by indicating that 72 percent of CEOs consider AI to provide competitive advantage against competitors in today's digitized environment (Vishnoi et al., 2018).

Although companies are investing in AI more each year, the adoption of the technology is not yet as rapid as it could be (Chui, 2017). Technology adoption can be a heavy investment for organizations. This together with the uncertainty of its benefits, could also explain why some managers are still holding back with AI adoption. When adopting AI to marketing purposes, there are many aspects managers need to consider. They are expected to be able to evaluate current situations as well as the future of their business and identify potential opportunities and threats that AI adoption could bring to the business as a result (Campbell et al. 2020). Shankar (2018) states that decision makers should be familiar with how and when AI will benefit their business before adopting the technology, even if the pressure from competitors may be strong and suggest otherwise. Additionally, AI requires a lot of quality data to run effectively, thus companies need to have proper data ecosystems and digital foundations in place before they can start working with AI, which for some companies may be a longer journey (Chui, 2017). Not being able to meet these expectations might also provide an explanation as to why some companies are putting AI adoption on hold despite its potential to provide competitive advantage. Also, the uncertainty of how the technology will develop can linger the decision-making process of the technology adoption (Chui, 2017). However, use cases by companies that have adopted AI so far already demonstrate the positive outcomes of the technology (Chui, 2017). Perhaps in years to come, as more use cases become available and potentially the results show significant improvement in various areas of business, the adoption of AI will take up rapid growth.

2.3 Benefits and current applications of AI in marketing

AI provides various possibilities for firms' marketing activities. According to Vishnoi et al. (2018), AI is not only providing little enhancement to current marketing activities but thrives to reshape marketing as it is today entirely. Shankar (2018) states that AI provides the possibility of getting new insight into consumer psychology by educating us on how the human brain is thinking in any given moment. Opening a window to consumers' minds is a significant asset

to a company as they can market products and services accordingly (Shankar, 2018). At its best, AI can provide remarkable competitive advantage. However, in order to reach this, companies need to take a proactive approach into their AI efforts. (Chui, 2017.) In this chapter, few of the most common areas of current applications of AI in marketing are discussed.

2.3.1 Targeted marketing

Targeted marketing is no stranger to marketers. By practicing behavioral tracking, marketers are able to match the behavioral characteristics of the user when marketing products or services to them. There are many reasons why targeting has become popular among marketers such as increasing sales, new customer acquisition, enhancing customer loyalty, pricing, making appropriate second offers to consumers, more targeted promotions in online advertising platforms and email newsletters and so on. The benefits of targeting do not only weigh on the marketers' side but on the customers' as well. The more targeted the marketing becomes, the more relevant the promotions become for the customer. (Alreck & Settle, 2007.)

Targeting customers better is one of the key assets of using AI-powered solutions for marketing as it enhances the marketers' possibilities to market products and services at the right target audience (or individual), at the right time, at the right price and with the right message (Chui, 2017). In addition to more precise targeting, AI also helps companies to improve their customer segmentation (Campbell et al., 2020). Placing cookies to a computer or recording a computer IP address of a website visitor, companies can gain insight into the users' online behavior (Alreck & Settle, 2007). Insight gathered from customers' web browsing habits, past purchases, demographic data and so on can all be used to target the customer better (Chui, 2017).

An example of effective AI-run targeting, Harley-Davidson managed to improve its customer targeting by using the AI-powered marketing platform Albert. With Albert, the company managed to reach out to a completely new audience they had never included in their marketing strategy before and reached an increase of 2,930 percent in leads with 50 percent of those being new target audience leads. Additionally, the results showed a 40 percent increase in motorcycle sales over a period of six months. (Marr, 2018.) An online gift shop, RedBalloon, also uses Albert to identify new customer segments and has reported a nearly 750 percent increase in Facebook ad campaign conversions and about 1,500 percent return on its marketing investment (Shankar, 2018). It comes as no surprise why using AI to target existing customers and new customer segments has reached its popularity among marketers.

2.3.2 Personalization and recommendations

Personalization is a popular topic among marketers today. It is the concept of creating personalized content to individuals based on their interests (Shanahan, Tran & Taylor, 2019). Personalization builds both short-term and long-term

business by creating valuable content to consumers (Aguirre, Mahr, Grewal, Ruyter & Wetzels, 2015). Aguirre et al. (2015) also remind that personalization is no stranger to marketers, and it is known by businesses in both online and offline world. According to Shankar (2018), predicting consumer behavior and personalizing recommendations are one of the many areas of marketing where AI is making a big difference for companies' success and many companies are already using AI for these purposes. It is even argued that AI algorithms have transformed personalization into a standard procedure in today's online marketing activities (Gentsch, 2019). Thanks to the great volume of data and the AI-powered solutions available, companies can now make better predictions and recommendations targeted to an individual customer than they were able to before (Chui, 2017). Using personal data within marketing for predicting next purchase and for offering personalized recommendations has the potential of increasing effectiveness of these marketing actions as the content and communication provided reflects on the customers' characteristics and habits and thus offers a much more personal experience (Ozcelik & Varnali, 2019). In addition to recommendations, personalization is also seen as user-tailored experiences on websites, which also enhances the personal experience of the customer (Shanahan et al., 2019).

Personalized recommendations have a big impact in customer experience as well as potential increase in conversions, as it reduces the search time of consumers by presenting products and services based on their personal preferences (Zhang, Zhao & Gupta, 2018). Recommendations can be drawn for example from past purchase behavior and preferences of like minded consumers (Zhang et al, 2018). For example, L'Occitane uses AI to offer personalized product recommendations to its online users based on their previous behavior, which has resulted in a 159 percent increase in conversion rates (Maytom, 2018). Netflix also uses insight from its users' past activities for recommendation algorithms to run personalized recommendations which enables its customers to find content to watch that matches their preferences in no longer than 90 seconds (Chui, 2017). Spotify uses data from its customers' behavior on the platform by tracking their individual listening habits and similar taste to provide a Discover Weekly playlist to match the user's preferences. Each user has their own personal taste profile which is created based on various aspects of their behavior like frequency of listening to a specific song, whether the user skips specific types of songs, which types of songs the user adds to their favourites and so on. (Madathil, 2017.)

Personalized recommendations also have the potential of increasing customer loyalty (Zhang et al., 2018). By offering personalized content to customers, companies enhance engagement and thus create stronger attachment between their brand and its customers (Shanagan, Tran & Taylor, 2019). Aguirre et al. (2015) also argue that customer satisfaction and thus higher profits are central outcomes of personalization. In Netflix's case, the enhanced customer experience has led to higher customer loyalty by decreasing the amount of unsubscriptions notably, which saves the company 1 billion USD annually (Chui,

2017). AI assistants are another good example of AI-powered personalization tools. AI assistants offer information and recommendations to the user and have the potential of winning over the consumers' trust and loyalty unlike any other technology before (Dawar & Bendle, 2018). Big tech giants like Amazon and Google have already launched AI platforms with these highly skilled AI assistants (Dawar & Bendle, 2018). Smart speakers like Amazon Echo, listen to their users and record speech, run analysis on the data collected, make predictions and offer personalized recommendations (Shankar, 2018). These examples support the argument by Aguirre et al. (2015) that using personalization within marketing practices drives not only customer satisfaction and profit but enhances competitive advantage.

2.3.3 Media optimization

AI is often used for media optimization as it enables more accurate targeting of the audience and running ad campaigns accordingly in various platforms such as Facebook, Google and YouTube (Shankar, 2018). Marketers are continuously searching for ways to use AI to automate tasks and improve efficiency. Google uses AI in its advertising platform which enables marketers to automatically adjust their ads to match users searches, identify best-performing ads and offers a bidding tool to aid the optimization of ads which all results in more effective ad campaigns with less manual work. (Tran, 2018.) Additionally, Google Adwords analyzes a variety of big data to provide insights to marketers in terms of which leads are qualified and which are not, and thus helps them enhance their targeting (Kietzmann et al., 2018). Marketers can also use AI to pre-run ads and track emotions of viewers to provide insight on when their engagement decreases or drops completely in order to know when to eliminate certain ads (Kietzmann et al., 2018).

In addition to providing guidance in specific ads and automating ad optimization, AI can help companies to identify the best platforms to run their ads in. As Harley-Davidson started to use Albert, the AI-powered marketing platform, the results showed that the company's Facebook ads converted 8,5 times more than ads on other platforms, which guided them to focus their online advertising only on the platforms that work (Marr, 2018). Albert can analyse these cross-channel ad results, learn from them and adjust the ad allocations to optimize the ad success and the return on marketing investment (Shankar, 2018).

2.3.4 Customer experience

AI is not only making an impact for companies but also for customers (Shankar, 2018). According to Campbell. et al. (2020), AI-powered solutions have a big impact in how customers are managing their relationships with brands. McKinsey has forecasted that in 2020, the relationship a customer has with a brand without a human interaction in the United States will account for 85 percent of all interaction (Campbell et al., 2020). AI has also become a part of decision-making and creation of the customer experience (AlSheibani et al.,

2018). By being able to analyze the customers' behavior through the use of AI, companies can take appropriate actions to improve customer experience (Shankar, 2018). As an example, L'Occitane, in addition to their personalization efforts with the use of AI as discussed above, has also taken other actions to improve their customer experience. The company tracked the behavior of its customers online and in its mobile app to identify points within the online customer journey where the customers became frustrated and made developments according to the results, which resulted in a 15 percent increase in mobile sales (Shankar, 2018).

Companies are able to offer products and services to a consumer at the right time and in a format that is most likely to trigger consumption behavior (Levesque, 2017). According to Chui (2017), the more tailored the customer experience is, the more special customers feel which evidently leads to higher customer loyalty and thus higher revenue. As an example, a grocery store chain could offer its customers an application that stores all their purchase history which could then provide recipes based on past purchases and favoured ingredients, or could provide sales coupons directly to the application when walking near the store for mostly purchased items and so on. These types of insight-based selling methods have shown results of increasing sales by 1 to 5 percent (Chui, 2017). AI can also be used for insight-based sales through dynamic pricing (Chui, 2017). According to Chui (2017), using dynamic pricing in an online environment among the most valuable customers can lead to an increase of up to 30 percent in sales. Amazon Go has even taken the shopping experience of customers to another level by offering the possibility to walk in and out of the store without going through a separate cashier, just adding items directly to their bag (Chui, 2017).

AI assistants like Amazon's Alexa, have a great potential in terms of enhancing customer experience. These AI platforms that run the AI assistants, offer customers the possibility to receive information, products and services specifically suited for them while minimizing costs and risks and providing convenience by ensuring routine flows of purchases into their households. (Dawar & Bendle, 2018.)

In addition to more targeted and personalized experience, better predictions and forecasts can also have a great impact on customer experience. A German online retailer Otto uses AI for forecasting sales and has reached a 90 percent accuracy in a period of 30 days. The results have been so good for the company that it now builds its entire inventory based on these forecasts and can now make faster deliveries, better customer experience and has evidently lowered the amount of product returns. (Chui, 2017.)

2.4 The role of data in AI-driven marketing

Understanding the role of data is important for marketers. Below, the role of data in AI-driven marketing, the types of data that can be used, the collection and

management aspects and risks associated with the use of data are discussed. Few key regulatory affairs related to data within the EU are also touched upon.

2.4.1 The role of data

Data is a crucial part of marketing today and it can be seen as the “oil of digital economics” (Wedell & Kannan, 2018). It offers insights into marketing performance and thus assists marketers to make data-driven decisions to optimize their marketing actions to improve return on investment (Wedell & Kannan, 2018). Data has become more available than ever before and the amount of data generated daily exceeds the amount that is possible for a human to take in, analyse and make decisions based on that data (Campbell et al., 2020). This is where AI steps in to make a difference and makes the available data more valuable than ever before. AI and ML provide a much faster, richer and precise learning from the extensive amount of data than humans could ever provide (Campbell et al., 2020). AI-powered systems can also provide thorough, effective analysis of data sets, group data and recognize changes as well as react to those changes in real time (Gentsch, 2019). It comes as no surprise that data is the bedrock of AI-driven marketing. In fact, AI requires data to run and most AI-powered solutions require existing data sets to run and is thus an essential part of successful use of AI (Campbell et al., 2020; Henke et al., 2016). It is important for a marketer to understand the role of data when applying AI systems into marketing practices.

Big data brings big opportunities and big challenges for marketers (Alshura et al., 2018). Manyika et al. (2011) estimate that with big data, retailers could increase their operating margin by as much as 60 percent. When looking at the journey of computing and its development, finding the data used to be the challenge. Today however, companies and marketers are facing an information overload through the large volumes of data, and the challenge of finding data has shifted into a challenge of transforming the data into something valuable. (Kietzmann & Pitt, 2020.) Alshura et al. (2018) also stress the challenge of connecting the data from various sources as well as managing that data, in addition to and as a part of transforming it into meaningful insights. Although data collection as a phenomenon to gain competitive advantage has existed for decades, thanks to the recent developments in big data algorithms and analytics, big data has revolutionized the way companies perform their marketing (Alshura et al., 2018). This overwhelming amount of data available has created big opportunities for marketers and AI has made it even easier for marketers to find value from the data and take actions accordingly (Kietzmann & Pitt, 2020).

The benefits data provides to marketers today and especially through the use of AI-driven marketing are extensive. Jarek and Mazurek (2018) highlight the importance of data in enabling successful delivery in consumer insights, market analysis, consumer needs research and measuring the effects of various marketing activities. Alshura et al. (2018) also mention new product and service innovations, enhanced customer service, new customer acquisition and strategy development as key areas of marketing that rely on data. Wedel and Kannan

(2016) emphasize data's role in creating value for not only companies but also for its customers by increasing customer satisfaction and thus, building customer loyalty.

Thanks to AI, large volumes of data can be processed fast and with more accuracy (Kietzmann et al., 2018). Data such as customer data is used by marketers to influence the purchase decisions through personalization within their marketing practices (Aguirre et al., 2015). Using customer data for these purposes is an efficient way of increasing marketing returns (Martin, Borah & Pamatier, 2017). By using algorithms, marketers can target their customers much more effectively based on the large amounts of data it processes, and studies show that these algorithms can even provide a better prediction on human personality, characteristics and preferences than their closest friends or family (Youyou, Kosinski & Stillwell, 2015). Thus, creating a far more thorough understanding of consumers as individuals. In addition to understanding consumers better, data provides companies answers to various questions like what type of a product or service should they offer for their customers, how and where they should advertise the product or service, what the price should be and so on (Alshura et al, 2018). To analyse these amounts of data in order to reach these answers, companies often rely on AI-powered systems (Kietzmann et al., 2018).

2.4.2 What data is used?

The volume of data available today is substantial, and it is more rich and diverse by nature than ever before (Henke et al., 2016). The amount of data generated daily in today's digital world is so significant, that all data generated during the past few years potentially exceeds the amount of data generated throughout history (Ashura et al., 2018). Recent innovations in technology such as the wide environment of social media has made the data collection faster and easier than ever before (Fan, Lau & Zhao, 2015). The amount of user-generated data alone reaches up to 2.5 billion gigabytes per day (Kietzmann et al., 2018). Facebook's users share millions of pieces of content per minute while Google receives billions of search queries per minute; all of which are stored and then analyzed (Wedel & Kannan, 2016). Jarek and Mazurek (2019) highlight the role of this overwhelming amount of consumer data in creating powerful marketing through AI. Marketers rely on AI systems to make sense of and to transform this large pool of data available into valuable insights (Kietzmann et al., 2018).

In order to analyse the significant amount of big data, AI looks into two different types of data; structured data and unstructured data. Customer demographics, transaction records and web-browsing history are forms of structured data. Unstructured data can be in the form of text, speech and images and accounts for approximately 80 percent of the 2.5 billion gigabytes of user-generated data that is created daily. The more this type of unstructured data is run by an AI system, the more detailed and insightful its results are for marketers. (Kietzmann et al., 2018.)

The types of data that can be used to run AI systems are diverse. AI can process a variety of data from numerical data to sounds and images (Jarek & Mazurek, 2019). ML algorithms, which are the most common AI approaches used for marketing purposes, use not only the companies' internal data like transaction data but also external data such as local weather forecasts (Chui, 2017). Campbell et al. (2020) discuss the potential that AI can offer for marketing purposes and provide insight into the types of data required to run those specific tasks. As Chui (2017) states, companies use both internal data and external data to run AI systems. The various actions enabled by AI in different stages of marketing planning and data required for these actions according to Campbell et al. (2020) are shown in table 1 below. The external data used for marketing purposes such as recommendation systems, product demand predictions, improved targeting of ads, identifying potential buyers and so on can be anything from demographic data and social media discussions to sales data of competitors, weather data, and any third-party data. Internal data which is often used either by itself or as a combination of other data, can be anything from sales and customer data to brand perception and historical data. As seen in table 1, there are various types of internal and external data which can be either historical data or real-time data. (Campbell et al., 2020.)

TABLE 1 Data used for AI-driven marketing (Campbell et al., 2020, 232-235)

What AI can offer	Data requirements
<i>Analysing the current situation</i>	
<ul style="list-style-type: none"> • Analysis, simplification, provision, and understanding of large unstructured data sets • Identification of events in the market • Recommender systems to identify likely future events • Sentiment analysis 	<ul style="list-style-type: none"> • External data, including census data, demographics, consumer confidence, macro-market trends, third-party data, social media discussions
<i>Understanding markets and customers</i>	
<ul style="list-style-type: none"> • Identification of changes in competition behavior (e.g. as pricing) • Estimation of product demand • Assessment of customer sentiment (e.g. customer satisfaction, social media sentiment analysis) 	<ul style="list-style-type: none"> • Internal data, including sales (current and historical, sales of own products), customer data (satisfaction, attitudes, demographics, etc.), market research (e.g. ad/promotion testing) • External data, including market share, scanner data, sales (sales of competitor's brands, seasonality, weather, holidays), social media comments, competitor's pricing and product availability
<i>Segmenting, targeting and positioning</i>	
<ul style="list-style-type: none"> • Classification and clustering of customers into distinct segments • Estimation of the probability of response to promotions • Improved targeting of ads 	<ul style="list-style-type: none"> • Internal data, including loyalty and sales information, customer willingness to purchase, and brand perceptions • External data, including demographics, census data, and location

<ul style="list-style-type: none"> Product and brand recommendations 	
<i>Planning direction, objectives, and marketing support</i>	
<ul style="list-style-type: none"> Provision of digital customer service (e.g. chatbots) Estimation of the responsiveness of consumers to price changes and promotions Combinations of information from the macro- and microenvironments to better inform marketing objectives Identification of those likely to purchase 	<ul style="list-style-type: none"> Internal data, including historical data on areas of marketing (e.g. ads/sales support) and associated outcomes (e.g. site traffic, leads, sales) External data, including census data, demographics, consumer confidence, macro-market trends, and third-party data
<i>Developing product strategy</i>	
<ul style="list-style-type: none"> Identification of gaps in the market for new product development Creation of more customized and boutique products Awareness of what is in style or trendy and thus worth producing and selling Assistance with designing and producing and selling Assistance with designing and producing products customized to individual consumers 	<ul style="list-style-type: none"> Historical data on customers, their purchases, and associated outcomes (e.g. satisfaction, returns) in order to create recommendations Databases of consumer profiles from which to estimate new customers' sizes/profiles depending on inputs Information on trending products, topics, and styles from social media, press articles, etc.
<i>Developing pricing strategy</i>	
<ul style="list-style-type: none"> Estimation of consumer price elasticity at both individual and collective levels Provision of dynamic pricing (e.g. surge pricing) and price discrimination Detection of anomalies (e.g. errors in pricing, fraud, or non-profitable customers) 	<ul style="list-style-type: none"> Both historical and real-time sales, search, and price data on firm and competitor products
<i>Developing channels and logistics strategy</i>	
<ul style="list-style-type: none"> Prediction and optimization of distribution, inventory, store displays, and store layouts (both brick-and-mortar and online) Enabling voice and visual search 	<ul style="list-style-type: none"> Data at the store level (historical and real-time sales, real-time inventory, in-store and web traffic data) and location level (local competitors, demographics of local catchment) Data on individual customers (historical sales, search history, any other customer-level data useful for making product recommendations) Historical customer service queries, responses, and satisfaction scores

<i>Developing marketing communication and influence strategy</i>	
<ul style="list-style-type: none"> • Creation of different ads depending on permutations of content, and on related words • Development of individual promotional offers and ads • Running of AI-driven A/B testing • Optimization of ad placement • Reduction in cart abandonment • Contextual ad targeting • Optimization of ad retargeting • Keyword bidding and cost reduction • Automation and personalization of content creation 	<ul style="list-style-type: none"> • Both historical and real-time data on ads, including their content (both text and images), placement, and performance • Information on potential ad placements (e.g. costs, audience characteristics) • Real-time data on customer behavior at all points along the consumer journey
<i>Planning metrics and implementation control</i>	
<ul style="list-style-type: none"> • Better prediction of expected revenues and profits, as well as their variability • Identification of metrics linked to key outcomes • Prediction of the effect of correct actions and, in some cases, automatically taking steps to diagnose, correct, and improve on poor results 	<ul style="list-style-type: none"> • Both historical and real-time sales and marketing performance data • Real-time data facilitates diagnosing problems, while historical data enables prediction of corrective actions

The data used for the AI systems can be anything from numbers and text to images, audio and video (Alshura et al., 2018). These forms of data can be gathered from the web, social media, payment systems, cameras and wearable devices (Henke et al., 2016). In addition to data generated by individuals' online behavior and other data with a digital footprint, data can also be generated from human activities such as facial expressions, body gestures, voice, speech, eye movement, heart-rate and so on. This type of data can be used to identify a person's emotional state. (Campbell et al., 2020.) Furthermore, this data could then, potentially, be analyzed and linked to other data of the individual to offer more personalized advertising and thus, create higher profit.

2.4.3 Data collection and management

There are multiple ways data can be collected, stored and managed. Henke et al. (2016) introduce three key categories of data ecosystem to help understand the journey of data from collection to usage. These categories are data generation and collection, data aggregation and data analysis (Henke et al., 2016). As discussed in chapter 2.4.2, recent advancements in technology such as social media have made data collection faster and easier than ever before (Fan, Lau & Zhao, 2015). Kietzmann et al. (2018) stress that the daily amount of user-generated data alone reaches a figure as high as 2.5 billion gigabytes. In fact, the amount of data available today is so significant that storing all of it is even considered impossible (Manyika et al., 2011). With new and diverse data sources emerging and with the

increasing volumes of data gathered, aggregating the data to one source can be challenging (Sleep, Hulland & Gooner, 2019; Henke et al., 2016). However, as AI requires large volumes of quality data to run, proper data ecosystems need to be in place for companies to run effective AI campaigns (Chui, 2017). Additionally, making data available for marketing managers from a single source is important to smoothen the operations and make it easy for the marketing team to gather insights from the data and rely on it within their decision-making processes (Sleep, Hulland & Gooner, 2019).

Data analytics creates extensive opportunities for marketing by fueling innovation. The complexity of data and analytics requires experienced translation as well as substantial computational power and infrastructure. (Henke et al., 2016.) Marketing managers are now also phasing an era where understanding modern technology and acquiring related skills has become important for enhancing the operations between data, technology and insights (Sleep, Hulland & Gooner, 2019). Organizations often overlook this fact by searching for data scientists to run their analytics and ignoring the link between them and those with practical knowledge of business (Henke et al., 2016). While technology innovations are a part of everyday operations not only within the IT department but marketing as well, the marketing teams might, and probably do, still require support from IT when delivering projects that require these technological skills. Thus, making close cooperation between the departments important. (Redhat, 2015.) Sleep and Hulland (2019) also emphasize the importance of cooperation between marketing and IT departments in terms of data management and data analysis activities to ensure successful operations and to avoid conflicts between the operations.

2.4.4 Data privacy issues

An important aspect for marketers to consider are risks associated with data. Thanks to AI and ML, companies have more insight about their customers than ever before due to the extensive amounts of data generated by an individual's online behavior (Davenport et al., 2020). While companies know more about their customers and can thus provide a more personalized experience for them throughout the customer journey, customers are becoming more aware of the data collected of them which raises concerns regarding their privacy (Davenport et al., 2020). Martin et al. (2017) agree by arguing that the growing efforts of data collection and usage increase the customers' concern about their privacy. They might feel uncomfortable receiving personalized advertising and content as they realize how much data of them is actually being collected and analyzed (Aguirre et al., 2015). Gentsch (2019) calls this type of use of deep insights from personal information to create a more personalized content "overkill marketing". A recent study shows that a majority of consumers feel uncomfortable with the amount of data that is collected from their online behavior by advertisers, and about half of the consumers assume websites do not comply with the current privacy regulations (Wedel & Kannan, 2016). According to Martin and Murphy (2017), the more worried customers are of their data privacy, the more negative their

response towards the brand becomes. Gentsch (2019) also believes “overkill marketing” to have a negative impact on performing successful marketing.

Ensuring privacy of customer data is a complex issue. As data collection has become easier and storage has become less expensive, the amount of data stored by one company on one individual can be substantial, and as the storage is easy and affordable, this data may be saved for longer than it was first intended to (Davenport et al., 2020). This data can also be collected from various different sources and then put together for analysis to reach a better understanding of customers and create customer profiles of them (Wedel & Kannan, 2016). All this data collected from customers whether it is combined with data from various sources or it is kept untouched but stored for a long time, can later be re-used and even for purposes it was not intended to be used in the first place (Davenport et al., 2020). Tucker (2018) supports these arguments by pointing out that there is no limit to the amount of times a piece of any data can be used and the number of times a certain piece of data is used is often expanding due to the lowered costs of using AI systems. Also, data collected of an individual customer may contain information about another individual (Davenport et al., 2020). Wedel and Kannan (2016) also raise the issue of information that should be private becoming revealed as a side effect of combining various datasets. From a company’s perspective, this is challenging as companies need to find the correct way of complying with privacy regulations and address the concern of customers but at the same time ensure these actions will not stand in the way of new innovations by making the data collection process too difficult for the company and thus limit the data insights and potential actions taken based on those insights (Davenport et al., 2020).

Davenport et al. (2020) discuss the concerns consumers have on their data privacy and Martin and Murphy (2017) bring up that some consumers may falsify information about themselves if possible, to feel more in control of their data. Consumers have a habit of analyzing what the tradeoff for them is when giving out personal information about themselves and are more likely to provide this information when they believe the benefits they gain overcome the potential costs of giving that information. Negative effects that data privacy issues may raise among customers can be addressed by being transparent about the data collected and how it is used. Also, giving the customer some control over data enhances the trust between the brand and its customers, which is seen as an important factor in customer relationships and the likelihood of consumers providing data for a certain company. (Martin & Murphy, 2017.) According to Martin et al. (2017) however, when a company has access to sensitive personal data of a customer, it already has a negative impact on the customer’s trust towards the company. Addressing these concerns and making the right strategic decisions in terms of data collection and usage is a complicated yet a necessary process for companies (Davenport et al., 2020).

Martin and Murphy (2017) also raise the question of whether the personal data of a consumer is provided to a company willingly or not. The use of data for marketing purposes often raises questions about these privacy issues and

marketers must comply with the current privacy standards when practising marketing (Kietzmann et al., 2018). There are some key regulatory affairs linked to the use of data privacy that marketers should be aware of. The EU General Data Protection Regulation (GDPR) is the most significant change in data privacy regulations in 20 years (Sterne, 2017). GDPR defines personal data as *"any information relating to an identified or identifiable natural person"* (GDPR, 2016). Thus, all data collected of an individual that is and can be identified should comply with the GDPR. GDPR also requires the data to be used only for the purpose that the individual has given consent to (GDPR, 2016). As discussed above, the amount of data stored of an individual customer can be extensive and marketers should take these issues into consideration. The penalty for not complying with the GDPR can be as high as 4 percent of a company's global turnover (GDPR, 2016).

In Europe, there are few other regulations to be considered that reflect on an individual person's right to privacy. The European Convention on Human Rights, which regulates human rights as an international instrument, states that everyone has a *"right to respect for his private and family life, his home and correspondence"* (Council of Europe, 1950). Also, the European Union Charter of Fundamental Rights, which regulates fundamental rights as an EU instrument, ensures *"Everyone has the right to respect for his or her private and family life"* and touches upon the protection of personal data by stating *"Everyone has the right to the protection of personal data concerning him or her"* (European Union, 2012). In addition to the possibility to access the data, the data also must be processed *"fairly for specified purposes"* and on the consent of the person or other legitimate basis laid down by law, such as the GDPR (European Union, 2012). A person is also required to have access to the data and correct it if needed (European Union, 2012). Consumers in the EU also have a *"right to be forgotten"* and they are able to request a removal of web links (Martin & Murphy, 2017). These regulations are by no doubt to be taken seriously by marketers and taken into consideration when collecting, storing, managing and using data. As these regulations demonstrate, the data privacy issues are a serious case for marketers to address.

Although marketers are aware of these regulations, some challenges are inevitable. As the GDPR (GDPR, 2016) states, a consumer should have the right to know what the data collected of them is used for when they give permission to access their personal data. This can be challenging for marketers as companies have been storing data for years, and that data might have been recombined and for obvious reasons, the ones responsible for creating that data in past, potentially a few decades ago, could not have been able to consider this when making their decisions regarding the data at that time (Tucker, 2018). Wedel and Kannan (2016) also raised the concern about private data being revealed as a result of dataset combinations throughout a company's history. They also argue that the privacy laws have not been able to keep up with the technology developments and this causes private consumer data coming revealed (Wedel & Kannan, 2016). Thus, there are many aspects to consider in terms of regulations when using

customer data in AI-driven marketing, as historic data should also be accounted for.

Although the privacy regulations create risks for marketers, they also create opportunities for more effective marketing practises through better customer understanding and more precise targeting throughout the customer journey (Kietzmann et al., 2018). However, marketers need to take care of their data management carefully to ensure negative effects do not occur (Martin et al., 2017). How to find the perfect balance in complying with privacy requirements, building trust between customers, having proper data management and being able to put new innovations into use to increase marketing return is a challenge every marketer can be expected to face.

2.4.5 Data bias issues

Another aspect of data related risks involve data bias. Algorithm bias can be a common outcome of an AI-run system if the data set it works with contains information that can expect a bias output (Davenport et al., 2020). As systems use specific data sets, they can be biased if users do not know how to train them with comprehensive data sets (Chui, 2017). Tucker (2018) also explains that earlier data history can reflect the predictions made by AI for example in cases where the data set used by an AI system has data that is endogenous. An example of such a situation is Amazon's recruitment system that favored male applicants and was then withdrawn from use due to the bias outcomes during the recruitment process (Davenport et al., 2020).

Bias can also be seen for example in targeted advertising. Algorithms can detect the ethnicity of an individual based on their online behavior such as liking a certain page, celebrity, music, movie or other on a social media site. The algorithm can then make predictions on individuals such as a consumer who engages more by expressing interest in various TV shows and music are likely to have less income than those who engage in discussions about politics and trending news. This type prediction can be biased if it for example, automatically offers certain products and services based on the predicted income. From a consumers' perspective, they might not even consider that liking a certain something could affect how the algorithm on social media sites like Facebook will then predict their ethnic group, income level and adjust the advertising according to those predictions. (Tucker, 2018.)

2.5 Summary of the literature review

The purpose of this study is to discover the factors that define what type of data is required when aiming to deliver successful AI-driven approaches for marketing purposes. With the research questions of this study in mind, the literature review covered the issues related to what type of data is used in AI for marketing purposes from the source of data to data collection and management as well as the aspects marketers should consider regarding data privacy and bias issues that were raised in existing literature.

Figure 2 below presents a summary of the literature review and presents a framework for this study. Based on the existing literature covered in this thesis, it illustrates the relationship between data and AI applications in marketing, and highlights the aspects of data that can be identified relevant to determine what type of data is required for AI-driven marketing.

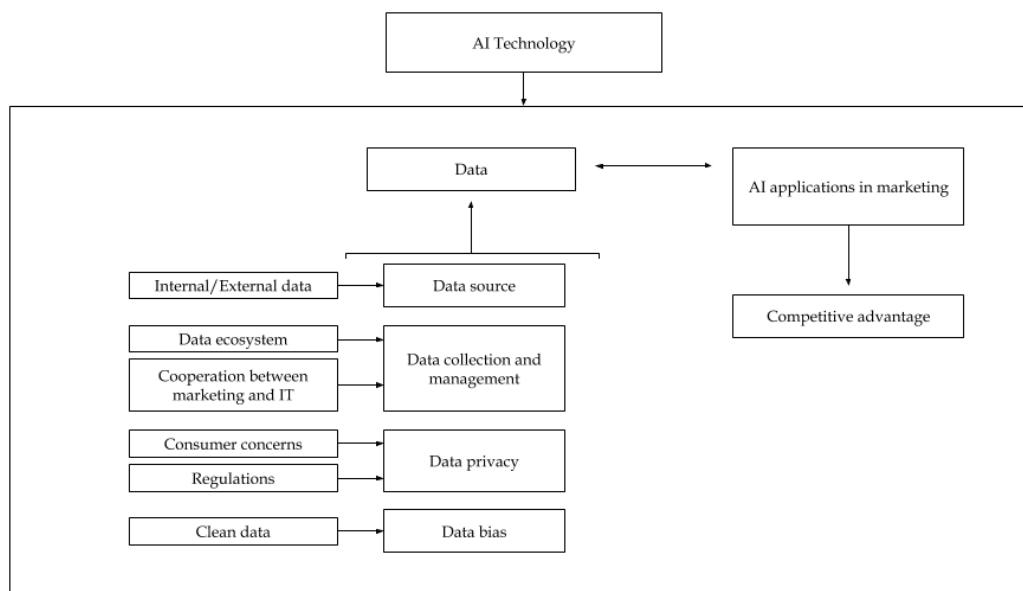


FIGURE 2. Framework

AI has the potential to transform marketing to a completely new level and thus competitive advantage is seen as the outcome of AI applications in marketing. Chui (2017) highlights targeted marketing, customer insights, personalization and customer loyalty as profitable performances and outcomes of AI-powered marketing. Campbell et al. (2020) add customer segmentation and enhanced customer experience as key takeaways from AI-driven marketing. Through the marketing practises led by AI, companies can achieve competitive advantage over peers (Aguirre et al., 2015; AlSheibani et al., 2018; Chui, 2017). However, to be able to carry out these activities with AI, marketers need to rely on data. Without data, AI-powered systems can not provide the potential benefits for marketers. Hence, it is important for a marketer to understand what AI-technologies enable for marketers and what type of data is required for these AI-driven approaches to work successfully.

As seen in figure 2, the relationship between data and AI applications is a two-way relationship. The more AI-ran campaigns are performed in marketing activities of a company, the more data is generated. In simpler terms, the existing data generates more data through AI run activities (Henke et al., 2016). However, when this data is generated through various marketing efforts, and when it is collected from various sources, marketers should understand what the data should be like to perform successful marketing. Thus, by understanding the underlying factors of data for AI-driven marketing, marketers can take these factors into consideration when planning their marketing activities in terms of

what data can or should be collected, how it should be collected and managed, what should be taken into consideration in terms of that data such as customers' privacy concerns and data bias and how to address those issues. A thorough understanding of the phenomena can provide a meaningful advantage when designing strategic decisions for marketing practises.

Looking at the factors presented under data in figure 2, the relevant aspects of data that rise from existing literature can be seen. The amount of data available to today is extensive as recent developments in technology have made data collection faster than ever before (Fan, Lau & Zhao, 2015). The types of data marketers can use reaches beyond the internal data of a company such as sales and customer data. External data such as demographic data, data from social media channels, weather data and any form of third-party are also used and is often combined with internal data. (Campbell et al., 2020.) The forms in which this data comes available can be numbers, text, images, audio and video (Alshura et al., 2018). The data can be gathered from websites, social media, cameras and payment systems to name a few (Henke et al., 2016). As the literature shows there are multiple forms of data, both internal and external that can be used for AI-driven marketing, it can be expected that any data can be used. However, while existing literature covers these sources and presents examples of data that are often used, the role of third-party data was left unaddressed as such.

Chui (2017) argues that companies are required to have proper data ecosystems and digital foundations in place to incorporate AI to even begin their marketing applications with AI. Furthermore, Henke et al. (2016) stress the challenge the increasing volumes and complexities of data creates for combining all the data to a single source. Sleep, Hulland and Gooner (2019) support this argument by highlighting the existence of a single data source and access to the data insights is crucial for marketers as they require this data in their decision-making processes. According to Sleep, Hulland and Gooner (2019), mutual understanding and smooth communication between the marketing and IT department is necessary for successful operations. From the existing literature, it can be expected that as marketing is becoming more data-driven and as AI technology is becoming a part of the daily operations of marketing, the understanding of data collection and management matters can not be left to the IT department alone, but an understanding of such matters is also of relevance to marketing managers. Henke et al. (2016) points out companies often rely on data scientists alone in terms of running analytics, and the relationship between these experts and those with practical knowledge of business is forgotten. Hence, it can be expected that the matters of data collection and management should be of mutual responsibility between these two departments so that successful and new innovative ways of performing marketing can be identified.

With the growing amounts of data available of an individual and the help of AI systems, marketers are able to provide more personalized marketing towards consumers (Davenport et al., 2020). However, while offering new effective ways of running marketing, this also raises privacy concerns among consumers (Davenport et al., 2020; Martin et al., 2017). For a marketer, it is

important to recognize these issues of privacy concerns when determining how they prefer to use data to run marketing with AI. Regulations such as the GDPR guide marketers towards the right direction when practicing marketing in terms of data privacy.

Data bias is another concern that rises from the existing literature. Algorithm bias can be a common outcome of an AI-run system if the data sets used contain any information with expected bias (Davenport et al., 2020). Thus, it can be concluded that data sets used for AI-driven marketing should be “clean” and clear of such information to avoid data bias. It is clear marketers should take these factors into consideration when practising marketing with AI. According to Tucker (2018), targeted advertising has already seen such cases where algorithms have delivered biased outcomes.

To conclude, the framework indicates that the role of data is central when applying AI to successful marketing applications that can eventually provide competitive advantage. Furthermore, it demonstrates the key aspects of data; source (internal/external), collection and management (ecosystem and cooperation between marketing and IT) as well as privacy (consumer concern and regulations) and bias issues (clean data). These factors can hereby be defined to be relevant for basing decisions on what type of data to use for these AI purposes. The empirical data will further indicate if these aspects of data are in fact relevant and whether this framework can be applied in practice. Before these findings and conclusions are discussed, the methodological approach to this research is presented in the next chapter.

3 DATA AND METHODOLOGY

This chapter introduces the methodological approach to the research. First, the type of the research, qualitative approach, is presented. Furthermore, the data collection and analysis are discussed.

3.1 Qualitative research

The purpose of a research approach is to define a strategy that enables the research questions to be answered and thus, the research question guides the decision of which approach is appropriate (Eriksson & Kovalainen, 2008). As the goal of qualitative research is to discover new insight to a phenomenon instead of testing something that has already been done, a qualitative approach is chosen as the research method for this study (Flick, 2009; Patton, 2014). The objective of this study is to discover what type of data should be used for AI-driven marketing for it to be successful. Qualitative research is also used to discover social relations and reality as described by the interviewees (Adams, Khan & Raeside, 2014). Thus, this type of approach provides a more comprehensive understanding of the underlying factors related to use of data as well as discovering the factors that make data appropriate. By identifying these factors, such as those related to data privacy and bias, provide an understanding of what type of data is required for marketers to run effective AI campaigns. It is common for qualitative research to enhance knowledge of those who are interested in understanding a certain phenomenon or problem (Patton, 2014). As AI-driven marketing has become a phenomenon of interest for researchers, policymakers and professionals in the field, providing further knowledge of the phenomenon is useful for various audiences. The results of qualitative research are in non-quantitative form and do not go through a laborious quantitative analysis (Kothari, 2004; Flick, 2009). Thus, a more in-depth understanding of the underlying factors can be presented by following a qualitative research approach. The justification for this approach is rather clear, as the research aims to discover what type of data is required to run successful AI campaigns in marketing through the experience and knowledge of experts and offer knowledge to those interested in the phenomena.

3.2 Data collection methods

There are various data collection methods for qualitative research such as interviews, observation and case studies (Eriksson & Kovalainen, 2016). To reach a holistic understanding of what data should be used and what should be taken into account by marketers, semi-structured interviews were chosen as the data collection method for this research. Interviews are an effective way to study a

phenomenon that is complex and may have sensitive aspects involved (Hair, Wolfenbarger, Money, Samouel & Page, 2015). Additionally, semi-structured interviews offer the possibility for unexpected information to appear during the interview, although the questions have been designed prior to the interviews (Flick, 2009; Kothari, 2004). The interviewer can, however, use initiative by asking related questions that were not included in the pre-designed list of questions (Hair et al., 2015). By providing flexibility, semi-structured interviews allow more in-depth knowledge of the phenomenon to appear that can enrich the results. For the research to be more holistic and trustworthy, experts working with AI technology who have knowledge, understanding and/or experience in marketing were interviewed. As the research problem of this study is rather complex and the understanding of the phenomenon studied may vary between the interviewees depending on their expertise, it is important to offer the possibility for new information to emerge while conducting the interviews. Table 2 below shows the titles of the interviewees in their profession, the country in which they work in as experts of their field, the type of the interview (face-to-face or over the phone) and the duration of the interview. The questions used for all expert interviews are presented in Appendix 1.

TABLE 2 Summary of the interviews

Interviewee	Country	Title	Type	Duration
I1	Finland	Head of Data	Face-to-face	28 min
I2	Finland	Chief Growth Officer	Phone	52 min
I3	Finland	Lead Data Scientist	Phone (with I4)	46 min
I4	Finland	Business Lead AI	Phone (with I3)	46 min
I5	Finland	CEO	Phone	37 min
I6	Finland	CEO and Co-Founder	Phone	49 min
I7	Finland	Director, Analytics	Phone	36 min
I8	Finland	CEO	Phone	24 min
I9	Finland	Founder and Chairman	Phone	38 min
I10	Bulgaria	CEO and Co-Founder	Phone	51 min
I11	Switzerland	AI expert	Phone	55 min
I12	United Kingdom	CEO	Phone	39 min
I13	United States	Previous Digital Director/ Adjunct Professor	Phone	1 h 2 min

The sampling method chosen for this research was judgement sampling, also known as purposive sampling. This sampling method guides the researcher to use his or her own judgement to choose the sample. (Hair et al., 2015.) As seen in table 2 from the titles of the interviewees, they vary in their level of expertise. However, they were all selected based on their expertise in AI and understanding

of its relations to marketing. In order to solve the research problem, it is not enough to compare information provided only from data scientists but also some understanding of marketing is required. Thus, with the chosen versatile group of interviewees, a holistic understanding of what it means to apply AI to marketing activities of a company, what data can be used and what risks are involved with data sets used for those purposes can be discovered.

There were a total of thirteen interviews and they were all recorded with a permission of each interviewee. Interviewees 3 and 4 were interviewed together. The interviews were then transcribed and analyzed. All interviews were conducted in English as the interviewees participating in the interviews were from various different countries. Translating gathered information of some interviews to match the language of others for the purpose of data analysis could potentially harm the initial meaning of the answer given. Thus, by conducting all interviews in English, false translations can be avoided.

3.3 Data analysis

The aim of data analysis is *“to identify, examine, compare, and interpret patterns and themes”* (Hair et al., 2015, p.301). In simple terms, the data collected should be reduced from its original form into a simple and comprehensible form without losing the initial value of the information gathered (Adams, Khan & Raeside, 2014). The data analysis of this study was conducted by following thematic analysis method. Thematic analysis in qualitative research is especially recommended for young researchers as it offers flexibility together with rich, comprehensive and complex data (Braun & Clarke, 2006).

Braun and Clarke (2006) have defined six phases of thematic analysis that were used as a guideline for the data analysis of this study. First, Braun and Clark (2006) stress the researcher should familiarize himself or herself with the data. During the transcribing process of this study, attention was paid so that no information of value was left out from the transcriptions. The interview recordings were first transcribed by writing down each interview in its entirety. Filler words such as “so”, “um” and “like” were left out. Both the interviews as well as the transcriptions were done in English. Second phase of thematic analysis is generating initial codes (Braun & Clarke, 2006). Coding is defined as *“the process of assigning meaningful numerical values or names that reduce data from a large amount of undifferentiated text”* (Hair et al., 2015, p.302). Thus, coding aids the researcher to focus on valuable key characteristics of the data. During this step, the entire data set was read and coded with precision. Third step for the researcher is to search for themes within the data (Braun & Clarke, 2006). In this step of the analysis, similar codes were put together in potential themes. Codes that demonstrated patterns throughout the data set were considered as themes. The themes were then color-coded to help analyze the data. The two following steps outlined by Braun and Clarke (2006) are to review the themes and name them. Here, the themes were reviewed while emerging similar themes to avoid

duplicates and named to provide a clear visual map of the data. Finally, Braun and Clarke (2006) guide the researcher to provide a report of the data. The results and analysis are presented in the following chapter.

4 RESULTS AND ANALYSIS

This chapter presents the research findings based on the empirical data as well as the analysis of these findings. The analysis consists of themes that emerged from the semi-structured interviews regarding what type of data is required to run effective AI campaigns in marketing. First, the role of internal data and external data as well as the role of third-party data is discussed. Next, the suggested data collection and management methods are explained. Finally, data privacy matters as well as data cleanness and reliability are presented. Conclusions drawn from the interviews are further discussed in chapter 5.

4.1. Data source

4.1.1 Role of internal data and external data

The importance of various data sources was recognized as a main prerequisite to run effective AI-driven marketing by the interviewees. Over half of the interviewees believed any data to be useful for these AI purposes. Correspondingly to Chui (2017) arguing companies use both internal and external data to run their AI systems, a combination of internal data and external data was specified as the most effective way of running AI campaigns.

Internal data was highlighted to be priority over external data by the interviewees, and external data sources were mentioned to enrich the existing internal data sets. However, two of the interviewees raised the aspect of external third-party data being extremely useful especially for companies that do not own any data, or for younger and smaller companies that require more data to get meaningful insights. These findings related to third-party data are discussed in more detail the following chapter 4.1.2. In terms of both internal and external data, one of the interviewees (I2) highlighted that the more data sources are used and synchronized, the better the results will be, as presented below. Another interviewee (4) also emphasized the potential of combining data from internal and external sources, and even combining data sets from various internal units, to see lucrative results.

“It is not just the question of what data, it is a question of how much data can we have cost-effectively, because the more data sources we connect and synchronize, the better we understand the problem that we are looking at and we are able to take action. Because I am not interested in insights that are left into PDFs and reports, I am interested in decision-windows so that companies and individuals within companies are able to make better decisions for business based on data.” (I2, Chief Growth Officer)

“Of course, all the CRM, ERP data should be utilized. All the data which is generated when the company is doing business ... so all those general data sources

but not forgetting the public external data sources to see what happens if you combine the data. It might even happen internally if you have different units that do not talk to each other, what can happen if you combine the data. That could be fruitful to see.” (I5, CEO)

It became clear that various sources of data, both internal and external, are relevant and useful for running AI for marketing purposes. One of the interviewees (I9) raised an important aspect for the process of deciding what data to use; the question that is wanted to be solved should always be understood first in order to make decisions regarding what data is needed. Thus, although any data could potentially be used, the purpose of what is wanted as a result and output of any AI campaign should be identified beforehand. Another interviewee (I11) stressed that when data is used at consumer level, for example to enhance or create new products or services, quality of data always exceeds quantity of data.

The forms and specific types of internal data and external data that were raised by the interviewees were diverse. Data from web services, customer data, competitor data, demographic data, weather data, CRM and ERP data, conversational data, transaction history and data of people's behaviour were all identified as important forms of data companies can use to run AI campaigns. As discussed in chapter 2.4.2. , Campbel et al. (2020) also distinguish many of these as data used in AI-driven marketing. In addition to the diverse list of different data types and forms mentioned, one of the interviewees (I4) stressed that there is no single right answer as to what data is important for running AI campaigns in marketing and that it depends highly on the individual situation of a company.

To conclude, the interviewees highlighted a variety of data sources to be important when choosing what data is used and a majority believes any data can be used for running AI campaigns. Understanding what the desired output of an AI run campaign is should determine the actual data sets that are required and used. Internal data is seen as more important than external data, and for smaller and younger companies, external data has more value than for bigger companies with more internal data sets available. Combining both internal and external data sets are seen as the best practice for running successful AI campaigns in marketing.

4.1.2. Role of third-party data

The use of both internal and external data were covered in the existing literature, but the nature of third-party data as such was not discussed in more detail. Thus, the interviewees were asked two questions regarding third-party data. First question was “what are the biggest sources for third-party data, can organizations trust third-party data or should they depend upon their internal data collection methods?”. Second question was “what are the advantages of using third-party data?”.

The interviewees’ views varied in terms of the biggest sources for third-party data. Two of the interviewees mentioned national institutes of statistics

such as the statistics bureau in Finland as well as Statista and European central agency in Europe as biggest and most trusted sources for third-party data. Open source data was also mentioned by two interviewees. Other sources that were specified were weather data, trend data, scientific paper archives, credit bill data and government data (given it would be provided to companies). In unity with Campbell et al. (2020), social media was brought up as one of the biggest sources of third-party data.

"... it would be social media. I mean the majority of people are on it and I would say that would be the biggest element of that." (I12, CEO)

As discussed in chapter 2.4.2., on Facebook alone, users share millions of pieces of content per minute and all this data is then stored and analyzed (Wedel & Kannan, 2016). Furthermore, the advantages of social media as a source of third-party data or data that provides information of the social networks of an individual was explained by two of the interviewees as follows:

"So the advantages of using third-party data is that as a company you are still missing a lot of information about your user, about its updates, about its social experiences and so on. So of course, if you use third-party data that enables you to discover for example the social network of a consumer, so how does he work inside the society, what type of influences emake for him to see, for example ... social platform data, this will enable a company to act more intelligent to the individual data." (I11, AI expert)

"... say I was a mortgage company and you were applying for a mortgage loan for your house from me. If my AI system is not just incorporated to the standard corporate environment checks that they do today but they also checked social media and interaction context outside of that, based on your spending profiles and how you live, I would have a much more accurate picture of whether I should lend you that money." (I12, CEO)

Kietzmann et al. (2018) argue the more unstructured data is run by an AI system, the more detailed and insightful the results are for marketers. The responses of the interviewees reinforce this argument and also refer back to the finding that a combination of both internal and external data is best practice to reach best results when performing marketing with AI. Also, these arguments demonstrate the advantages of using third-party data, as it offers a more thorough understanding of an individual for the purpose of marketing.

However, regardless of what the third-party data is and how much information it could provide of an individual, the trust factor of the third-party data should also be addressed. According to the interviewees, when using data from a third-party data source, it is important to understand what data it actually is, how trustworthy the third-party is and how the data was collected. When these factors are clear, the data can be trusted. While one of the interviewees (I6) raised a concern of unreliability risk of third-party data due to companies not

having control over it nor visibility to see how clean it is, another one trusts its quality as seen through the following argument:

"... normally there is somebody who takes care of the quality there because the intention is that many companies would utilize these open data sources. And the requirement is that they need to be able to trust the data quality. So, if it does not have good quality it will die and disappear later." (I5, CEO)

In addition, the same interviewee (I6) who questioned the reliability of third-party data further stressed that if the data sets can be accessed or the third-party data is a separate data set itself, it can be trustworthy as one can have visibility to the data. Yet another interviewee (I10) lists third-party data as the least trustworthy source and names public open data more trustworthy and internal data as the most trustworthy source of data. Based on the opinions and knowledge of the interviewees, it can be concluded that having a clear understanding of the data in terms of what it is and where it comes from is relevant in order to be sure the data is reliable enough to include in any AI campaigns.

Some other concerns were disclosed by the interviewees regarding organizations and their relationships with using third-party data. One of the interviewees (I2) mentioned that organizations are very lazy in collecting third-party data, and another interviewee (I6) said he believes third-party data is always required to perform successful campaigns.

"... it has always been so that they need third-party data sets. Because we have to combine a lot of data and if we think about certain algorithms or certain use cases it is quite evident that existing inhouse data is not enough, so third-party data sets will be required." (I6, CEO and Co-Founder)

One of the interviewees (I3) believed that when there is not enough data available, third-party data is one choice for companies to use but did not give it much weight if internal data sources are versatile and big enough. Thus, the necessity of third-party data was not completely unanimous among the interviewees. Although, as the findings discussed earlier showed, the more data sources that are used, the better the performance of an AI system will be and the combination of internal data and external data is the preferred way of running AI campaigns effectively as described by the interviewees.

The advantages of third-party data were more consistent between the interviewees. Over half of the interviewees emphasized that third-party data offers wider aspects and better predictions for running and enhancing companies' businesses when added together with the companies' internal data. They highlighted that if a company looks at their internal, very narrow data, they can analyze and optimize that, but they will miss out on the bigger picture. Also, when a third-party data is added as a data source, in any normal situation, companies will always get a better prediction to run their business, whatever prediction that may be. Two of the interviewees also further specified that the

advantages of using third-party data are much higher for smaller companies and for companies operating in smaller markets, and these companies also depend on third-party data sources more than others. An example provided by one of the interviewees (I2) summarizes such a situation where a company should add third-party data to their internal data to reach better predictions.

"... especially in smaller markets like in the Nordics this is actually relevant, because if you are a Finnish coffee chain and you have maybe 10 locations compared to Starbucks that has thousands and thousands of locations, so you are not able to get a good picture just based on the amount of data." (I2, Chief Growth Officer)

Two of the interviewees also stress that using third-party data is an option when there is not enough data and sometimes it is even the only option available. This supports the finding that some small companies are more dependent on third-party data. Companies and freelancers can even run their business based on third-party data alone, if necessary.

"... there are many data sources where even one freelance mobile developer can make you great apps utilizing the external data sources, even if that freelancer does not have any own data." (I5, CEO)

To conclude, third-party data that can complement the existing data or substitute nonexistent internal data should be used to run effective AI campaigns for marketing. Sources like social media, national statistics and weather data, can provide meaningful insights when combined with internal data to gain better predictions. The importance of third-party data and what third-party is used in AI campaigns varies based on the purpose what is wanted to reach and on the size of the company and market. Trustworthiness of third-party data is important and understanding its quality is relevant if the data sets are meant to be used as such or combined with internal data for marketing purposes through AI.

4.2 Data collection and management

To understand the role of data collection and management in running successful AI-driven marketing, existing literature recognizes the importance of proper data ecosystems and the cooperation between marketing and IT departments as key factors. Henke et al. (2016) disclose data generation and collection, data aggregation and data analysis as key categories of data ecosystems. As this research aims to find answers to what type of data is required to run effective AI campaigns, the focus is placed more on the first two categories defined by Henke et al.

Majority of the interviewees stress the current data collection and management methods are not good enough and most companies should improve these methods. Chui (2017) and Henke et al. (2016) emphasize the importance of having proper data ecosystems in place in order to deliver successful AI

campaigns. The findings from the interviews agree and show that young companies are considered to be more successful with their data ecosystems in general, while mature companies are considered not as mature and data information driven. However, one of the interviewees (I2) also discussed the benefits mature companies may have through having longer customer relationships versus the challenges of not having these existing relationships that can ease the pressure of running successful businesses through data. Thus, for younger companies, having proper data ecosystems in place is more vital compared to those that have longer customer relationships, and not having these systems in place could lead to bankruptcy.

"I think that younger companies are better at this for two reasons: they do not have the legacy systems that make it in the real world very difficult to incorporate modern machine learning AI technology. The second thing is, many of the older corporations are able to survive because they have such long customer relationships even if their effectiveness is pisspoor. But what took them here, they are still utilizing that whereas if these younger companies would be so poorly managed, they would go bankrupt immediately. So there is the advantage of not having the legacy and then there is the pressure of doing better business because you do not have the long standing relationships from customers." (I2, Chief Growth Officer)

Major challenges the interviewees raised concerning companies' current data collection and management methods were the fact that they do not have data sets in place to begin with or data sets that are not a part of the daily business are often missing, and on the contrary, a phenomenon of having a lot of data but no one either knowing about it or caring about it. Not having data sets in place was evidently seen as a showstopper for running AI campaigns in marketing. Not having data apart from that generated through the daily business can be concluded to diminish the success potential of these campaigns as the interviewees stressed a combination of both internal and external data provides better results. Kietzmann et al. (2018) further stress the challenge of connecting data from various sources and managing it properly. The interviewees agree and further stress data quality and storing processes are a challenge within companies' data managing practices. For the data collection methods to be successful, it was highlighted that any data collection should begin from understanding the business goals and knowing what it is that is wanted as a result from running the AI campaign instead of collecting data with no further reason than to be on the safe side.

"... most importantly, before you start with anything with AI please think about the use cases, what is it that is wanted to be improved." (I5, CEO)

"When you know the business goals, then you know the tools or AI tools or machine learning tools that are applicable and when you know the tools you will know the data sets that are required. And we always like to come from the top down method, and then we see that the companies actually have the correct data sets in place."

Because if they don't, that is always or in many cases a show stopper. But I don't want to encourage anybody towards that kind of approach, let's build a huge data warehouse or huge data sets just in case. Because you will never know where they are needed and if you are still collecting certain data, will that be applicable for the machine learning solution in question." (I6, CEO and Co-Founder)

However, one of the interviewees saw data collection of any kind to have potential benefits for business operations and innovations in the future, as explained below.

"One good example is our library system. Library collected information for hundreds or thousands of years and they didn't know why they collected it. But a big part of our scientific advances are based on the fact that we have this information, even though at the time we didn't know what we collected it for and how we use it." (I9, Founder and Chairman)

This type of approach may not offer effective solutions for the time being but may be of great assistance in the future. To further enhance the data collection and management methods, the interviewees emphasized consistency in the chosen methods and keeping a simple infrastructure in place. Henke et al. (2016) saw the need of substantial infrastructure for data and analytics to be imperative for running successful operations around data. The importance of such infrastructure was described essential by the interviewees. Building a huge data warehouse was not considered to be the correct approach for companies, as explained by one of the interviewees above (I6). Instead, the interviewees considered starting with the business goals and the use case, then determining what data is required, what tools are required and finally deciding how to label the data as the correct approach. Getting to know what data is already there and understanding the purpose and context to what any data has been collected for were also considered to be important. Additionally, cleaning the data before running AI was raised as an important factor in terms of data quality. The process and reasons for cleaning data are discussed in more detail in chapter 4.4.

Data quality was one of the challenges raised by interviewees regarding data management. It was discussed that more often than not, the issue is that the data that is actually needed for a specific purpose does not exist and to overcome that challenge, the existing data sets are then used, which further reflects on the data quality. The data might not be updated, or it is wrong data for that purpose. To solve this issue, one of the interviewees (I1) recommended that after the business goal is identified and data sources are mapped, the current data quality situation should be evaluated and then further enhanced to ensure the best quality of the data.

"I would say, start identifying what you want to do with AI, what's the use case, and then map what kind of data sources or data assets you need to do that, and then check the current state of quality and consistency of that and then put in place actions to increase the quality." (I1, Head of Data)

Sleep and Hulland (2016) stressed the importance of cooperation between IT and marketing departments in ensuring successful operations and avoiding conflicts between the operations. It was not specified by the interviewees that a strong cooperation between the IT department and marketing department would be important. However, it can be concluded that many of the challenges and concerns related to data collection and management could be decreased by such cooperation based on the arguments made by the interviewees. Henke et al. (2016) argue organizations often overlook the link between departments and thus practical knowledge of business is not communicated to data scientists or other specialists working closely with data and vice versa. The interviewees stressed the importance of understanding the business goals and the desired outcome of any AI campaign before defining the required data that needs to be collected in the first place. Additionally, the issue of having a lot of data but not having anyone knowing about its existence or caring about its existence could potentially be eliminated through more active cooperation between the departments. The interviewees also raised the challenge related to data quality that a lack of understanding of preferred outcomes of these AI campaigns can deliver. Thus, it can be concluded that understanding the business goals is the starting point of all data collection and management activities regardless of which department is handling these matters more closely. Lack of such understanding can lead to unsuccessful delivery of AI campaigns in marketing and thus communication of the business goals and use cases between those responsible for such activities is vital.

To conclude, the current methods for data collection and management are not good enough to run effective AI campaigns, and companies need to improve these methods. Younger companies have more pressure in terms of data collection and analysis as new customers expect more from these younger companies when marketing their products or services to them. A data ecosystem should always be in place and all activities related to data collection and management should begin with knowing and understanding the business goals as well as the use cases. Simple infrastructure, simple processes and consistency in all activities enhance a successful delivery of data collection and management while also increasing data quality. It is recommended that companies should first start collecting data from their daily business and further adding other data to complement the existing internal data. Cooperation between marketing and IT departments was not raised as a key factor for determining how data should be collected and managed but understanding the business goals was seen as the key starting point for these activities. Business goals determine the actions that follow, such as what data is required and thus collected, what tools are used and how the data is labeled and analyzed. Thus, building a successful data ecosystem begins with understanding the business goals. Data that is required to run effective AI campaigns for marketing is clean, quality data that is collected and managed consistently according to the business goals.

4.3 Data privacy

As discussed in chapter 2.4.4, it is important for companies to understand and comply with the current data privacy requirements and other regulations that guide data collection and management practices. Davenport et al. (2020) stress the importance of ensuring these processes do not stand in the way of gathering relevant data insights of customers and creating new innovations. This conclusion can also be drawn from the interviews, as the interviewees agreed the privacy requirements and regulations guide the operations of companies. Sterne (2017) argues GDPR to be the most significant change in data privacy regulations in recent years. The interviewees also stressed its role in marketing practices today. However, the interviewees saw regulations such as GDPR to have a positive effect on the incorporation of AI into marketing practices.

"I am 100 % convinced that it will affect it in a good way. Data privacy is a huge issue nowadays already and it will be more so in the future. And now when we have GDPR in place, it will force the companies to concentrate on these privacy issues already in the beginning. It is not possible to, if something is already up and running, it is very difficult to take into account certain privacy things. But if we have those rules in place already in the beginning, it will be a great thing." (16, CEO and Co-Founder)

"I think GDPR was the best thing that could happen in the data business... I would say that these regulations help create fair and sustainable AI business. Fair is kind of the same as ethical but an easier concept." (19, Founder and Chairman)

As seen above, GDPR is considered to guide companies to the right direction regarding consumer data privacy from the very beginning. One of the interviewees (I2) emphasized that GDPR is a major factor when it comes to incorporating AI into marketing practices and that it has a negative impact only when there is a lack of understanding involved. He explained that it will divide those who are capable of using data and analytics in a smart, benevolent way to help their customers and to carry out various solutions and those who are afraid of GDPR due to lack of understanding of what it allows, and thus do not perform such actions. Hence, placing GDPR as a positive driver for marketing practices for those who educate themselves with the matter. These findings support the arguments of Kietzmann et al. (2018) as they disclose privacy regulations to create challenges for marketers while also creating new opportunities for more effective marketing practices.

Another interviewee (I1) explained that thanks to GDPR, consumers are now more aware of what data is collected of them and also understand their own responsibility related to the data they provide for any company at any given circumstance, which can be seen as a good thing. However, customers also expect to gain value from giving out their personal data through receiving better services. This also refers back to the findings regarding the expectations consumers have on younger companies especially, when considering marketing

activities that are carried out based on collected data such as personalization, as discussed earlier in chapter 4.2. She also highlighted that by complying with GDPR or other regulations, companies can create better services and increase their overall brand image as well as confidence of certain partners.

While the positive role of GDPR was emphasized by the interviewees, there were few challenges disclosed concerning the limitations towards AI-driven marketing practises that such regulations may cause. One of the challenges described by Business Lead AI (I4) reflect back to the data collection and management processes.

"...You have to understand what data you are using and how you are using it and whether the usage of that data is purposeful as for what the data was created." (I4, Business Lead AI).

As discussed in chapter 4.2., the existence of a proper data ecosystem and the understanding of business goals as well as the use case for any data that is collected is crucial. Thus, by understanding what any data is used for and managing it consistently as recommended by interviewees, this challenge can be expected to be tackled. Martin et al. (2017) also highlighted the significance of properly taking care of data management to prevent negative effects from occurring. Another interviewee (I5) stressed the debilitating impact of data anonymization when following certain regulations such as GDPR. The potential value of consumer data on an individual level could provide a company could be much higher than insights from a higher level.

"I guess that the GDPR was created to protect the privacy of people and I think the intention was good and positive. But in the AI world, it might complicate things. You need to take it to account and you need to anonymise the data so that you are not violating the regulation and that is something, if it is just a high level, do you get the added value of the data that is the question." (I5, CEO).

Another interviewee (I11) agrees with GDPR complicating things for marketing practices. However, he emphasized the role of trust between the company and its customers correspondingly to the viewpoint of rising privacy concerns among consumers raised by Davenport et al. (2020) and Martin et al. (2017), as discussed in chapter 2.4.4.

"So GDPR has certainly stopped B2C organizations from getting a full advantage of an AI model...I believe that companies that have stronger relationships with consumers and customers would be able to collect all the needed data that will enable them to run these algorithms in the most effective way. So if your consumer customers trust you as a company, they are willing to provide any sort of data you need because they understand it goes to their own advantages. But when it comes to collecting data in a subtle way in an unclear and advantageous way, using this data for purposes that do not go to the advantage of the consumer, this is definitely a way to destroy that trust. And if there is no trust, companies will struggle in

finding a way to collect this data because of the regulations such as GDPR in Europe.” (I11, AI expert).

Respectively to the opinion presented above, Martin and Murphy (2017) also stress the role of providing customers some control over their data which they believe to build trust between a customer and a company and increase the likelihood of customers providing their data to the company. Thus, building trust between a company and its customers can be identified as an important factor for receiving customer data while complying with certain regulations such as GDPR. Another interviewee (I13) raised a situation that companies may phase in the future where they will need to have full explainability of algorithms to the public and believes that this kind of transparency is an emerging theme among businesses leveraging AI. Yet another interviewee (I1) also emphasized the importance of transparency and explainability in AI-driven marketing, as anything a company does that touches upon an individual level easily becomes a bit scary from consumers' perspective. The Director of Analytics (I7) interviewed also discussed transparency as a factor when using AI for marketing practices. She highlighted that everything that is done with data collection and storing begins with following rules and regulations that apply, and when processing the quality of that data, the transparency factor comes into picture in terms of being able to backtrack the data and thus understand what is really being done. Correspondingly to the interviewees' opinions, Aguirre et al. (2015) recommend transparency in terms of when and how consumer data is being collected. Tucker (2018) also discloses the challenge of using data that a company has stored for years and using it later on for purposes that may vary from what it was originally collected for. In such situations, transparency, as brought up by the interviewees, takes up a role and can be expected to affect the trust relationship between a company and its customers when using their personal data for marketing purposes. Thus, it can be concluded that in order to build trust between a company and its customers, transparency with data related matters is required.

4.4 Data cleanness and reliability

The interviewees stressed the importance of data cleanness and reliability at various occasions throughout the interviews. These two qualities of data were identified as pivotal factors in running effective AI-driven marketing and were brought up in relation to data collection and management, data privacy and preventing inaccurate or biased outputs. The interviewees agreed with the arguments of Davenport et al. (2020) in terms of considering data bias to be a common outcome of an AI-run system if the data sets used contain information with expected bias. Tucker (2018) also discusses that there have been cases where AI-driven marketing practices, such as targeted advertising, have shown biased outcomes. Thus, to prevent biased outcomes, the interviewees emphasized the

data used for AI-driven marketing should be clean and reliable. Biased outputs as such were not raised as a key factor in determining what type of data is required to run effective marketing campaigns with AI, but the cleanness and reliability of data were seen as the relevant components and key factors that may cause bias if not properly managed.

The interviewees defined reliable data as data that can be traced back to its origin. Factors such as where it was created and in which circumstances and by whom should be known when using data sets to run AI-driven marketing. Clean data was defined as data that is correct both in terms of context and by being up to date. One of the interviewees (I1) explained companies often use customer record data with their marketing practices and sometimes there are parts of this data that are correct and parts that are no longer accurate. This type of data is unclean as such, and if used in AI-driven marketing practices, it may provide ineffective or inaccurate results. She further explained data should be managed with consistency so that there is nothing missing, and all data points and time stamps created are consistent with one another and thus make the data quality accurate for any specific use. In accordance with the findings discussed previously in chapter 4.2, she pointed out that understanding the business perspective is the starting point in these processes as well.

“That’s really the starting point; to understand from the business process perspective that when we do our daily work like this, this is the kind of data that we are able to catch and collect. And then the other thing is to really understand from the business perspective what is the quality requirements, so what data quality level is needed and for what data sets and then make sure that that’s followed and there are KPIs also in place to monitor and check that.” (I1, Head of Data)

However, another interviewee (I9) stressed that the definition of clean and reliable data depends entirely on the use case. He explained that in cases like suggesting the next best song on Spotify, any data can be good enough to make accurate outputs, but if a psychological profile of the customer is wanted based on their previous song history, the list of played songs is not valid enough alone to create such a profile.

One of the interviewees (I2) stressed that data which is generated through machinery is easy to clean while data created by humans has many variants in the data. If, for example, humans collect data to a customer relationship management software, it will always be filled with limited amounts of data and the data is, more often than not, all over the place. Another interviewee (I7) further discussed the complexity of data by reminding the data used for any AI campaign can also be in the form of a picture, a voice or any other unstructured data. It was suggested by the interviewees that companies should depend on data scientists or data engineers to oversee the data is clean and reliable enough, as they will always be able to identify when the data is accurate enough for any specific use case. One of the interviewees (I1) emphasized that companies should have data governance in place to ensure these matters are properly disclosed.

The interviewees also raised the concern of bad data providing bad results, especially in terms of precision and predictions when using AI for marketing practices. Understanding the data that is being used in terms of its quality, reliability and cleanness, was highly emphasized to ensure best possible results. One of the interviewees (I7) explained the importance of the matter as follows:

“That is a super important thing, it must be the responsibility of every organization that uses AI to understand their data that they are using. Or if we think about machine learning, what data is used for machine learning to learn. And that is a part of our job when we do this AI work, it is our data scientist’s job to insert the data set and have the understanding if that is the proper quality data that can be used.” (I7, Director of Analytics)

As mentioned above and discussed in chapter 2.1.2, machine learning algorithms learn from the data that is put into the system. Correspondingly, Chui (2017) stresses the potential biased outcomes of AI systems that a lack of training with comprehensive data sets can cause. Based on the interviewees’ opinions, understanding what data is used is extremely important and as discussed earlier in this chapter, data scientists or data engineers are suggested to make these decisions. The Director of Analytics interviewed (I7) also reminded the importance of understanding what the business goal is and thus, what is wanted as the outcome of any AI campaign when deciding whether the data that is planned to be used is accurate or not. This way it can be ensured that the data represents the desired outcome. She continues her point of view by discussing the theme of potential bias:

“...for example, if we think about the B2C world, the data should represent all the ethical groups in my world, in my target group, so it does not give the wrong results. The data should be structured related to the question...” (I7, Director of Analytics)

As mentioned above, data that is used should include all relevant data for the question that is wanted to be answered or the outcome that is wanted of any AI campaign. It was pointed out by the interviewees that when AI campaigns are run without all relevant data, it becomes a challenge for companies as the outcomes of these campaigns can be inaccurate. As discussed above, companies need to pay attention to what data is put into AI systems as, for example with machine learning, the algorithm learns from the data that is put into it. If the original data sets are lacking information, all analyses that come later on are not entirely accurate. One of the interviewees (I12) stressed the role of missing information as a big ethical issue connected to AI, especially as companies automate their processes based on the output of AI systems, which naturally are a result of the initial input. Hence, as previously discussed, ensuring all data is clean is vital for it to be reliable enough to provide best possible results. It was recommended by one the interviewees (I11) that companies should start thinking about kind of information would they need if they wanted to anticipate consumer

trends, predict individual behavior, personalize messages and so on at the very beginning of their data collection and generation to ease the entire process and ensure effective delivery of AI campaigns.

Thus, while the existing literature discussed data bias as an important factor when it comes to practicing marketing through AI and was seen as such by the interviewees as well, it was not recognized as an important factor in determining what type of data is required for these purposes. However, if data is not clean and reliable, and is thus “bad data” as described by the interviewees, it can give inaccurate and even biased outcomes. Acknowledging these outcomes can be identified as important and proper actions should be taken into consideration in order to ensure data is clean and reliable when collecting, managing and analyzing data.

5 DISCUSSION

This chapter discusses the contributions of this study to both theory and practice. First, the theoretical contributions are discussed by modifying the established framework according to the empirical findings. Next, managerial implications are proposed to provide guidance for marketers for AI-driven marketing in terms of data related matters. Finally, limitations of the study and future research suggestions are presented.

5.1 Theoretical contributions

AI is seen as one of the biggest game changers in marketing and is believed to be reshaping marketing as it is today (Vishnoi et al., 2018). AI has the potential of providing competitive advantage, but companies need to take a proactive approach into their AI efforts (Chui, 2017). As AI requires data to run, data has become an essential part of marketing today (Campbell et al., 2020). Furthermore, as marketing is becoming more data-oriented, marketers are finding themselves working in an era where understanding these modern technologies and becoming more data information driven has become a part of running successful operations within marketing (Sleep, Holland & Gooner, 2019).

This study aims to answer the research question “What type of data is required to run effective AI campaigns in marketing?” and to provide a framework for marketers to guide them in considering the right factors when planning and running these AI campaigns. In chapter 2.5, a framework based on the existing literature is presented. In figure 3 below, this framework is modified based on the findings of this study. The modifications to the initial framework are bolded in the figure.

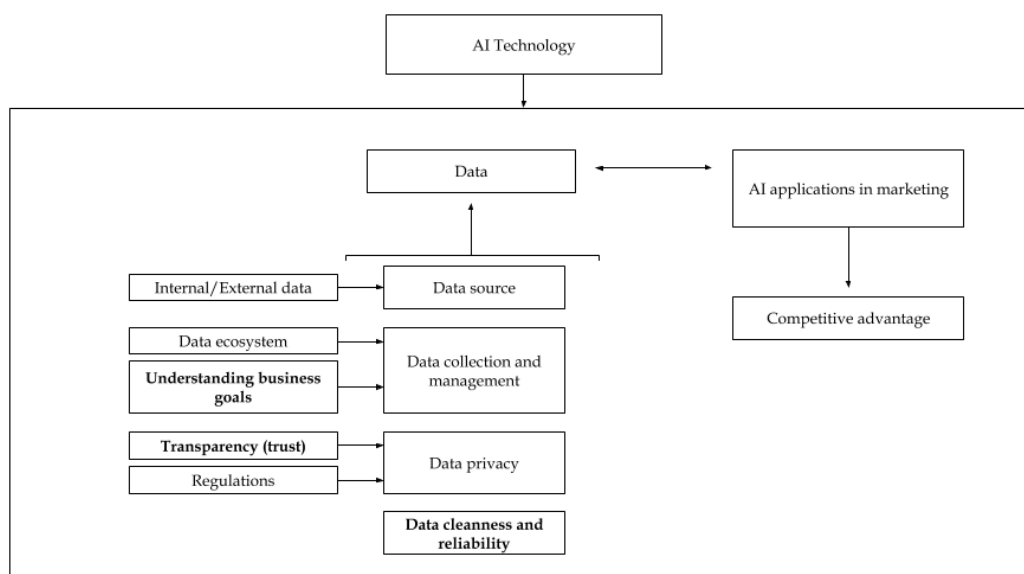


FIGURE 3. Final framework

The framework demonstrates the factors the interviewees agreed with in the empirical analysis. The findings support the initial theoretical framework for most parts, which shows the empirical data supports the literature review almost to the full extent. However, some changes were made to the initial framework as seen in figure 3.

In general, the importance of understanding relevant factors involved with data collection, management and analyses were emphasized. When examining the data source used for AI driven marketing, the existing literature discusses companies using both internal and external data to run AI (Chui, 2017). As seen in table 1, Campbell et al. (2020) disclose various internal and external data sources that companies use for their AI-powered marketing practices such as sales data, brand perceptions, demographics, third-party data and social media data. The findings reinforced the importance of both internal and external sources of data when running AI campaigns in marketing. Internal data was seen as the priority if available, and external data was believed to compliment the internal data to provide deeper insights and better predictions. It became evident from the interviews that third-party data such as social media was considered valuable in terms of providing a bigger picture of customers and allowing effective personalizations. There is no one source for data that can be identified as key data to run effective AI campaigns, it rather depends on the individual situation of any company running these campaigns.

When discussing the second factor in the framework, the data collection and management methods, understanding the business goals was identified as the key factor to run these processes and thus to also determine the data ecosystem which should be put into place. According to the existing literature, cooperation between IT and marketing departments are seen as an important matter in order to practise AI-driven marketing effectively. Sleep and Hulland (2019) see this cooperation as essential part of ensuring successful operations and to avoid

conflicts from arising. The empirical data did not specify cooperation between these two departments, but did however, emphasize the importance of understanding the business goals and what is wanted as an outcome of any AI campaign. Henke et al. (2016) fears organizations often rely on data scientists to cover these activities as they have knowledge and understanding of technology and insights, all the while ignoring the practical knowledge of business that marketing managers possess. Thus, theory also recognizes the challenge the lack of understanding of business goals can provide. The crucial component to effective data collection and management based on the empirical data is the understanding of business goals and therefore, the cooperation between departments was modified to understanding the business goals within the framework. Data ecosystems' role as key success factor for data collection and management were emphasized by Chui (2017) and Henke et al. (2016) as well as the interviewees.

Both existing literature and the empirical data show companies are required to follow current regulations within their data practices. However, according to Wedel and Kannan (2016), over half of the consumers presume companies do not comply with the current privacy regulations. This assumption together with the growing amounts of data collected of customers is shown to affect their image of the company. Davenport et al. (2020) and Martin et al. (2017) stress privacy concerns of customers to be an emerging theme within data collection. The empirical data showed that transparency in data collection and management diminishes these concerns consumers have and builds trust between a company and its customers, which in the future will potentially have a big role in data collection. While Martin and Murphy (2017) stress consumers may falsify information about themselves to feel more in control of their data, the interviewees considered transparency to encourage consumers to provide their data to companies. Thus, consumer concern was modified to transparency within the framework, as it builds trust and diminishes concerns while ensuring companies keep receiving data from customers to then use with their AI campaigns.

Theory recognizes data bias as a common outcome of AI-run systems when the data put into the systems contains information that can create such an outcome (Davenport et al., 2020). While the interviewees agreed with theory, data bias was not considered to be a factor in determining what type of data is required to run effective AI campaigns in marketing. Rather, it is one outcome of unclear and unreliable data. Data should be clean and reliable to ensure accurate, effective and unbiased outcomes of any AI-driven campaign. These two features of data were also stressed in relation to data collection and management in general as well as with data privacy and can thus be identified as highly important factors in determining what type of data is required to run effective marketing campaigns with AI. Therefore, data bias was modified to data cleanness and reliability within the framework.

The objective of this study is to find out what type of data is required to run effective AI campaigns in marketing. Two sub-research questions were included

in the study to further compliment this objective. The first, “How should this data be collected and managed?” relates to the decisions marketers make at early stages of various activities and to guide them to consider relevant matters throughout their processes. The second sub-question “What is the role of third-party data?” relates to finding an answer as to what benefits if any does third-party data bring to AI campaigns in marketing. As the aim is to offer marketers practical suggestions on how to enhance their data collection and management practices and/or understanding of the factors involved, the managerial implications that these two sub-research questions provided in the empirical analysis, are presented in the following chapter 5.2 together with those related to data privacy.

5.2 Managerial implications

This chapter further evaluates the findings discussed in the previous chapter while presenting the managerial implications on how marketing managers can ensure the data used for AI campaigns in marketing is good enough to provide effective results. The findings provide four main implications for managers. First, the results of the empirical study show that combining external data with internal data is the best way to run successful marketing campaigns with AI. Third-party data such as social media data complements the existing internal data sets in terms of providing deeper insights, better predictions and enhances the possibilities for personalization efforts. There is no simple answer as to what data is the most effective data in any campaign, instead the desired results need to be identified beforehand to know which data brings the most effective results.

Secondly, understanding the business goals is a key starting point for determining any data collection and management processes, and thus communicating the desired outcomes to whoever is responsible for managing these processes is priority. While the findings did not emphasize the importance of cooperation between departments, if the business goals and desired outcomes are communicated to the IT department, the starting points for data collection and management could be expected to enhance. Additionally, the findings revealed there is often data that no one knows about or cares about. Thus, if those handling the IT departments educate marketing departments of existing data, new use cases might emerge.

Thirdly, building trust between customers in terms of data collection and management is important as it may have long-term effects on the success of marketing efforts. The findings show that without having the trust of the customers, they might not provide their information to companies to begin with due to privacy concerns. By being transparent about these matters, companies can build this trust between their customers and thus keep receiving valuable data. However, customers also expect to get value back in exchange for the valuable data they provide. This is also relevant for managers, as this expected value can be delivered through successful AI campaigns.

Finally, data scientists and/or data engineers are needed to ensure data is clean, reliable and the quality of data is what is required to succeed with the AI efforts in marketing. What managers can do to enhance these qualities of data is to be consistent with any activities regarding data. The findings pointed out data especially within customer relationship management systems is more often than not, all over the place. Filling in all relevant data ensures the data this accurately reflects the existing customers' profiles.

5.3 Limitations of the research

There are certain limitations to this research. This study aims to provide understanding of the factors that make data appropriate to run effective AI campaigns in marketing. However, the data collection was conducted via semi-structured expert interviews on a rather broad scale instead of focusing on a specific factor. Although the semi-structured interviews offered the possibility for new aspects to emerge, this thesis was not able to offer a detailed description but rather a general overview of the factors that make data appropriate for practicing effective AI campaigns in marketing.

Additionally, the sample size of the interviewees was enough to identify commonly emphasized factors but could have been bigger in terms of including more marketing experts with some understanding of data related matters. By doing so, the responses of those working more closely with marketing and those working more closely with data could be compared to distinguish differences and similarities between the two, and thus identify potential challenges more comprehensively. Especially regarding factors such as cooperation between departments and transparency in data privacy matters. In this research, most of the interviewees have stronger expertise in data related matters out of these two areas, which can affect their opinions about certain aspects of data. However, this research aims to identify data related matters and thus, expertise in data matters is crucial.

Additionally, if more precise results are wanted in terms of each theme covered in this research, some amount of quantitative data could also be added to provide comparison to the qualitative data. Findings such as a combination of internal and external data creating more effective campaigns than focusing on one data source alone, for example, could be examined through actual case studies where analytics of such campaigns could be examined. Furthermore, this study did not focus on a specific industry such as retail but offered a general view of the important factors of data for any company regardless of its industry. Findings showed some answers are not possible to generalize for any company and thus further research on a specific industry or situation of a company could provide more thorough results.

5.4 Future research suggestions

To create a more thorough and precise guideline for what type of data should be used for AI campaigns in marketing, further research on the topic is required. As this research provided a broad overview of the type of data required and the underlying factors that make data appropriate for these marketing practices, there is room for future research on each theme covered in this research. As briefly discussed in the previous chapter 5.3, future research could compare findings from expert interviews to quantitative data from case studies in terms of identifying which data source provides best outcomes with AI campaigns. Future research should also focus on a specific industry or situation of a company to offer precise guidelines for marketers working in the field in question.

Few of the underlying factors that relate to data being appropriate to run effective AI campaigns also require further research. Although the findings did not emphasize a strong cooperation between marketing and IT departments or between marketing managers and data scientists as discussed in the existing literature, the results clearly showed that there were several factors that emphasize on one hand the importance of understanding the business goals and on the other hand the understanding of data governance. This research did not focus on defining the required cooperation between these areas of expertise but focused on identifying whether such cooperation is a factor in determining the appropriate data for running effective AI campaigns in marketing. However, the findings show signs of importance in the matter and thus, the cooperation and its value to successful AI campaigns should be further researched.

Another factor that requires further research is the transparency factor related to data privacy. The findings reveal the importance of building trust between companies and customers as privacy issues are rising among consumers. While the findings show regulations like GDPR are seen as a good thing for companies and transparency is identified as a factor in making data appropriate as customers provide accurate and valuable data of themselves when they trust a company, further effects of the matter are not covered. Future research should examine the effects of transparency in data matters on the effectiveness of AI campaigns in marketing.

REFERENCES

- Adams, J., Khan, H. T. A. & Raeside, R. 2014. *Research methods for business and social science students* (Second edition.). New Delhi: SAGE Response.
- Aguirre, E. Mahr, D., Grewal, D., de Ruyter, K. & Wetzels, M. 2015. Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing*, 91(1), 34-49.
- Alreck, P.L. & Settle, R. B. 2007. Consumer reaction to online behavioural tracking and targeting. *Database Marketing & Customer Strategy Management*, 15(1), 11-23.
- AlSheibani, S., Cheung, Y. & Messon, C. 2018. Artificial intelligence adoption: AI-readiness at firm level. *Twenty-Second Pacific Asia Conference on Information Systems*, Japan 2018.
- Alshura, M. S., Zabadi, A. & Abughazaleh, M. 2018. Big data in marketing arena. *Big opportunity, big challenge, and research trends: An integrated view. Management and Economics Review*, 3(1), pp. 75-84. doi:10.24818/mer/2018.06-06
- Braun, V. & Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2), 77-101.
- Bughin, J., Chui, M. & McCarthy, B. 12.7.2017. What every CEO needs to know to succeed with AI. [Referred 14.4.2020]. Available at: <URL: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/digital-blog/what-every-ceo-needs-to-know-to-succeed-with-ai>>
- Cam, A., Chui, M. & Hall, B. 2019. Global AI survey: AI proves its worth, but few scale impact. Available at: <URL: <https://www.mckinsey.com/~media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Global%20AI%20Survey%20AI%20proves%20its%20worth%20but%20few%20scale%20impact/Global-AI-Survey-AI-proves-its-worth-but-few-scale-impact.ashx>>
- Campbell, C., Sands, S., Ferraro, C., Tsao, H-Y. & Mavrommatis, A. 2020. From data to action: How marketers can leverage AI. Elsevier, *Business Horizons*, 63. 227-243.
- Chui, M. 2017. Artificial Intelligence. The Next Digital Frontier? McKinsey and Company Global Institute.
- Council of Europe. 1950. European convention for the protection of human rights and fundamental freedoms, as amended by protocols nos. 11 and 14, November 1950, ETS.5. Available at: <URL: https://www.echr.coe.int/Documents/Convention_ENG.pdf>
- Davenport, T., Guha, A., Grewal, D. & Bessgott, T. 2020. How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48, 24-42.
- Dawar, N. & Bendle, N. 2018. Marketing in the age of Alexa. *Harvard Business Review* May-June 2018.
- Eriksson, P. & Kovalainen, A. 2008. *Qualitative methods in business research*. SAGE Publications Ltd., London.

- European Union. 2012. Charter of fundamental rights of the european union, 26 October 2012, 2012/C 326/02. Available at: <URL: http://data.europa.eu/eli/treaty/char_2012/oj>
- Fan, S., Lau, R.Y.K., Zhao, J.L. 2015. Demystifying big data analytics for business intelligence through the lens of marketing mix. Elsevier, *Big Data Research* 2, 28-32.
- Flick, U. 2009. An introduction to qualitative research. London: SAGE publications Ltd.
- General Data Protection Regulation. 2016. Official Journal of the European Union, 2016/679, 1–88. Available at: <URL:<https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679>>
- Gentsch, P. K. 2019. AI in marketing, sales and service: how marketers without a data science degree can use AI, big data and bots. Cham: Palgrave MacMillan.
- Hair, J. F., Wolfinbarger, M. E., Money, A. H., Samouel, P. & Page, M. J. 2015. The essentials of business research methods (Third edition.). London: Routledge.
- Harvard Business Review. 2020. Artificial intelligence and the future of sales and marketing. *Harvard Business Review Analytic Services*, 1-12.
- Huang, M.-H., & Rust, R. T. 2018. Artificial intelligence in service. *Journal of Service Research*, 21(2), 155-172.
- Jarek, K. & Mazurek, G. 2019. Marketing and artificial intelligence. *Central European Business Review*, 8(2), 46-55.
- Jarrahi, M.H. 2018. Artificial Intelligence and the future of work: human-AI symbiosis in organisational decision making. Elsevier, *Business Horizons* 61, 577-586.
- Kaplan, J. 2016. Artificial Intelligence: what everyone needs to know. New York: Oxford University Press, Incorporated.
- Kielmann, J., Paschen, J. & Treen, E. 2018. Artificial intelligence in advertising. How marketers can leverage artificial intelligence along the customer journey. *Journal of Advertising Research*, 58(3), 263-267.
- Kietzmann, J. & Pitt, L.F. 2020. Artificial intelligence and machine learning: what managers need to know. *Business Horizons*, 63, 131-133.
- Kietzmann, J., Paschen, J. & Treen, E. 2018. Artificial intelligence in advertising. How marketers can leverage artificial intelligence along the consumer journey. *Journal of Advertising Research*, 58(3), 263–267.
- Kothari, C. R. 2004. *Research methodology: Methods & techniques* (2nd rev. ed.). New Delhi: New Age International (P) Ltd., Publishers.
- Levesque, H.J. 2017. Common sense, the Turing test, and the quest for real AI: reflections on natural and artificial intelligence. MIT Press.
- Madathil, M. 2017. Music Recommendation System Spotify - Collaborative Filtering. *Reports in Computer Music*. Aachen University, Germany.
- Marr, B. 3.7.2018. How Harley Davidson Is Using AI and Robots To Prepare For the 4th Industrial Revolution. Available at: <URL: <https://www.forbes.com/sites/bernardmarr/2018/07/03/how-harley-davidson-is-using-ai-and-robots-to-prepare-for-the-4th-industrial-revolution/#4c3e01025271>>

- Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: effects on customer and firm performance. *Journal of Marketing*, 81(1), 36-58.
- Martin, K. D., & Murphy, P. E. 2017. The role of data privacy in marketing. *Journal of the Academy of Marketing Science*, 45(2), 135-155.
- Maytom, T. 27.6.2018. L'Occitane boosts mobile conversions by 159 percent with AI-powered personalisation. Available at: <URL: <https://mobilemarketingmagazine.com/locitane-boosts-mobile-conversions-by-159-per-cent-with-ai-powered-personalisation>>
- Ozcelik, A.B. & Varnali, K. 2019. Effectiveness of online behavioral targeting: a psychological perspective. *Electronic Commerce Research and Applications*, 33, 1-11.
- Patton, M. Q. 2014. *Qualitative research & evaluation methods: integrating theory and practice*. Sage publications.
- Redhat. 2015. Driving digital transformation: new skills for leaders, new role for the CIO. *Harvard Business Review Analytic Services*, 1-16.
- Shanahan, T., Tran, T.P. & Taylor, E.C. 2019. Getting to know you: social media personalization as a means of enhancing brand loyalty and perceived quality. *Journal of Retailing and Consumer Services*, 47, 57-65.
- Sterne, J. 2017. *Artificial Intelligence for marketing. Practical implications*. John Wiley & Sons Inc.
- Tucker, C. 2018. Privacy, algorithms, and artificial intelligence. In *the economics of artificial intelligence: An agenda*. University of Chicago Press.
- Shahid, M. Z., & Li, G. 2019. Impact of Artificial intelligence in marketing: a perspective of marketing professionals of Pakistan. *Global Journal of Management and Business Resear*, 19(2), 26-33.
- Shankar, V. 2018. How artificial intelligence (AI) is reshaping retailing. *Journal of Retailing* 94 (4, 2018), 6-11.
- Siau, K.L. & Yang, Y. 2017. Impact of artificial intelligence, robotics and machine learning on sales and marketing. *MWAIS 2017 Proceedings*. 48. Available at: <URL: <http://aisel.aisnet.org/mwais2017/48>>
- Sleep, S., & Hulland, J. 2019. Is big data driving cooperation in the c-suite? The evolving relationship between the chief marketing officer and chief information officer. *Journal of Strategic Marketing*, 27(8), pp. 666-678.
- Sleep, S., Hulland, J., & Gooner, R.A. 2019. THE DATA HIERARCHY: factors influencing the adoption and implementation of data-driven decision making. *AMS Review*. pp. 1-19.
- Tran, K. 12.7.2018. Google is capitalizing on AI in marketing. Available at: <URL: <https://www.businessinsider.com/google-uses-ai-to-enhance-ad-campaigns-2018-7?r=US&IR=T>>
- Vishnoi, S.K., Bagga, T., Sharma, A. & Wani, S.N. 2018. Artificial intelligence enabled marketing solutions. A review. *Indian Journal of Economics & Business*, 17(4), 167-177.
- Wedel, M., & Kannan, P.K. 2016. Marketing analytics for data-rich environments. *Journal of Marketing*, 80, 97-121.

- Wirth, N. 2018. Hello marketing, what can artificial intelligence help you with? *International Journal of Market Research*, Vol. 60(5), 435-438.
- Wright, L.T., Robin, R., Stone, M. & Aravopoulou, E. 2019. Adoption of big data technology for innovative B2B marketing. *Journal of Business-to-Business Marketing*, 26 (3-4), 281-293.
- Youyou, W., Kosinski, M., & Stillwell, D. 2015. Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences of the United States of America*, 112(4), 1036-1040.
- Zhang, H., Zhao, L. & Gupta, S. 2018. The role of online product recommendations on customer decision making and loyalty in social shopping communities. *International Journal of Information Management*, 38(1), 150-166.

APPENDICES

APPENDIX 1 SEMI-STRUCTURED INTERVIEW QUESTIONS

Question 1

What kind of important forms of data can a company use to run AI? *For an example customer purchase history etc.*

Question 2

Do you think firms' internal data management and collection methods are good enough to run effective AI campaigns? If not, what can we do to solve this issue?

Question 3

What is clean or reliable data? How can organizations make sure that data they collect should be reliable to run effective AI solutions in marketing? In other words, what reforms a company has to take in existing data collection methods in order to get reliable data?

Question 4

What are the biggest sources for third-party data? Can organizations trust on third-party data or should they depend upon company's own internal data collection methods?

Question 5

What are the advantages of using third-party data?

Question 6

Do you think that unclean or unreliable data can give wrong results if it is put in an AI system? What do you suggest in order to collect clean data for correct AI outputs (predictions, recommendations etc.?)

Question 7

How can Government regulations (like GDPR) or consumer data privacy requirements affect the adoption process of incorporating AI into marketing practices? How would you suggest solving this issue?