Heli Raappana

Ethics in Data-Driven Marketing

School of Technology and Innovations
Master's thesis in Information Systems Science
Master's Programme in Digital Business Development

Vaasa 2020

UNIVERSITY OF VAASA
School of Technology and Innovations
Author: Heli Raappana
Title of the Thesis: Ethics in Data-Driven Marketing
Degree: Master of Science in Economics and Business Administration
Programme: Master's Programme in Digital Business Development
Supervisor: Ahm Shamsuzzoha
Year: 2020 **Number of Pages:** 121

ABSTRACT:

In today's digitalized world, people leave trails of data about themselves when operating online. While marketers are eagerly utilizing this data for marketing insights, many consumers have a growing fear of their personal data ending up in the wrong hands and losing their privacy. This thesis aims to understand the role of ethics in current marketing based on consumer data and to examine how data could be used ethically in marketing. This thesis met these research aims through an extensive literature review related to the issue and through empirical research. The empirical research was conducted by analyzing secondary data from a 2018 survey about the use of digital services in Europe. This analysis was combined with a content analysis of Twitter tweets about data ethics.

This research produced a number of key findings: First, privacy, confidentiality, and transparency are critical issues for consumers. Still, a large number of people are unaware of their data rights or think that is not important to change e.g. privacy settings. Also, some of the people feel that changing privacy settings has no effect. Second, people want to have the power to decline the selling of their personal data to third parties and to delete or adjust their personal data. Third, the lack of trust prevents individuals from using digital services, and the main factors for increasing trust are the security and reliability of the service. Other important factors are transparency and ease of use. Personalization of services based on the previous usage was the least important factor for a digital service. The content analysis showed that organizations are heavily criticized for jeopardizing privacy and confidentiality, whereas praise was given to organizations that focus on transparency.

The main conclusions that can be drawn from these findings are that the need for ethics is even more apparent in marketing than ever before, because of the broad access to personal data. The power of Big Data does not come from the raw data itself but from the ability to combine and merge data, thus creating detailed insights about consumers. There also lies the foundation for ethical dilemmas. The most crucial challenge for organizations in the use of personal data is how to build and maintain trust. Consumers' use of personalized services is dependent on whether the organization is perceived to be trustworthy or not, and both the empirical research and literature have shown that the lack of trust reduces the use of digital services. Transparency and clear privacy policies can then be seen as a key ingredient in trust-building. Giving the consumers the ability to opt-out from data use and data selling without it affecting the use of services can create an enormous competitive advantage for organizations in the future. Still, it can be challenging to combine the need for profitability with ethics. However, success in the long-term requires organizations to integrate ethics with the operations throughout the organization.

KEYWORDS: Marketing, Business Ethics, Marketing Ethics, Data Ethics, Big Data, Consumer Behaviour, Privacy

Contents

Figures

Tables

1 Introduction

We live in a world that is significantly influenced by marketers. Marketing stimuli surrounds us in many forms, both online and offline, and much of what we learn from this world is filtered by marketers. Consumers rely on marketers to sell them products and services that are safe and operate as promised, and also to tell the truth about what they are selling. (Solomon, 2016, pp. 14.) Data about consumers is used to personalize marketing messages as well as to gain a better understanding of the consumers' needs, ultimately improving business effectiveness and increasing profits.

Every day massive amounts of data are created as we leave a trail of data from millions of messages, emails, photos, and videos shared online or by mobile devices. In addition, the data from sensors like the GPS is gathered from our mobile devices. (Marr, 2016.) This abundance of data means that marketers can utilize an enormous amount of consumer data rather easily. Data and the insights received from it are hence at the core of data-driven marketing.

In this setting, ethical issues are more critical than ever before. Consumers demand more individualized and personalized content, which requires using data about individual consumers. The balance between personalization and privacy is a critical issue facing marketers. On the one hand, consumers want personalized communication, but on the other hand, they are concerned with the implications it has on their privacy. Privacy as a concept is also closely related to ethics, especially data ethics. Flaherty (1999, pp. 27) argues, that privacy has always been threatened in the Western world, resulting from surveillance ideology, technology, or both. Technological innovations, such as the telegraph, the telephone, and the camera were all considered a high threat to privacy upon their introduction. Therefore, concern for privacy can be seen as a fundamental aspect of humanity. (Flaherty, 1999, pp. 27-28.) However, the amount of personal data gathered in the modern digitalized world is massive, and therefore the threat to personal privacy is more apparent than before. For example, Amnesty has expressed hard criticism towards the leading data giants, Google and Facebook, accusing that their business models are a

threat to privacy and even to human rights. (Amnesty International, 2019). Many marketers utilize these companies' marketing tools, so this is a matter that needs to be taken seriously by marketers as well.

The use and storage of personal data can be regulated through laws and regulations, like the GDPR in Europe. However, merely abiding by the law is not enough for organizations. The law cannot keep up with the rapid pace of technological innovations, so the ethical practices need to be taken into account at an organizational level. Therefore, it is essential for marketers as well as consumers to understand how data can be used in an ethical way.

This thesis will study the ethical considerations regarding data-driven marketing. It will describe how data has affected the business and consumer environments, and it will explore how data is used in marketing. Finally, this thesis will examine the role of ethics in current marketing and the ethical challenges facing data-driven marketing. The thesis consists of a literature review regarding consumer behavior, data-driven marketing, and business ethics. In addition to the literature review, a mixed method study is conducted to research consumers' views on the ethical use of data.

1.1 Research Gap

The concept of privacy has been widely researched before. For example, Chellappa & Sin (2005), Garcia-Rivadulla (2016), and Jackson (2018) have all studied the relationship of personalization and privacy. Martin & Murphy (2017) have also studied the role of data privacy in marketing. However, though privacy is an integral part of marketing ethics, the ethical challenges in marketing are not restricted merely to privacy. Nor is marketing limited to personalization, albeit being an essential part of modern marketing. Ethics in marketing is gaining attention, and there is some earlier research on data ethics. However, most of the research regarding data ethics in marketing focuses on marketing research. There are also some studies about targeted marketing, but most of them concentrate either on targeting to underaged children or senior citizens. This research aims

to gain a holistic view of the subject, combining both the consumer view on data use and its implications on marketers. The consumer view is taken into account, as marketers need to understand the customers and their behavior in order to satisfy their needs, and consumer behavior is the key driver in the collection of marketing and consumer data.

1.2 Overall Research Aim and Individual Objectives

The overall aim of this research is to gain understanding of the ethical challenges in data-driven marketing. It aims to examine the changes in the business and consumer environments and explore the role of data and privacy in current society. Moreover, this research will look at the ethical implications that marketers need to take in consideration regarding data and privacy.

The individual research objectives aim to answer the following questions:

1. How Big Data affects the business and consumer environments?
2. What kind of challenges are involved in the use of personal data?
3. What is the role of ethics in current marketing?
4. What kind of implications does the concept of privacy pose to marketers?
5. How data can be used in an ethical way in marketing?

1.3 Boundaries of the Research

To sustain the usability of this research, it concentrates on the ethics of data-driven marketing on a general level without focusing on specific platforms or technologies. As platforms and technologies evolve and change rather quickly, it is not useful to concentrate on any particular platform and doing so could make the research quickly outdated. Ethics, on the other hand, is an important issue regardless of what platform is used.

In this research, the emphasis is on using Big Data for customer analytics, focusing on Business-to-Consumer (B2C) marketing and online marketing, where data is used for tar-

geting and personalization. According to Kumar et al. (2013), marketing data can be categorized into traditional, digital, and neurophysiological data. This thesis focuses on the digital data, as it is the largest in volume in current marketing data sources. Also, the use of digital data is the source of most of the ethical dilemmas in contemporary marketing.

The concept of privacy is critical when examining ethical considerations. This thesis will explore privacy more from an individual and socio-cultural point of view rather than from a political level. As the concept of privacy and ethics are culturally related, this thesis will focus on exploring these concepts from a European perspective. According to Ess (2014, pp. 66), compared to other cultures, the concept of privacy in Europe is strongly deontological, i.e. privacy is seen as a fundamental right which must be protected even at an economical cost. Ethical theories are only briefly examined in this research, and the focus is more on the practical implications of ethics in business.

1.4 Definition of Key Concepts

This chapter explains the key concepts related to this thesis.

Big Data refers to massive data sets with large and complex structures. The data is obtained from various sources, e.g. emails, videos, audios, images, clickstreams, search queries, sensors, and mobile phones and their applications. (Sagiroglu & Sinanc, 2013.)

Business Ethics means the application of ethics, i.e. what is considered right or wrong, into business practices (Paliwal, 2006, pp. 7). The concept of business ethics can be expanded further into marketing, information, and data ethics.

Data-Driven Marketing utilizes data from various resources, analyzing that data in order to gain insights about customer behavior, and using those insights to create marketing strategies (Arthur, 2013, pp. 48).

Data mining means that statistical methodologies are used to analyze data in order to solve business problems and refine strategies. Data mining allows organizations to discover hidden opportunities that cannot be detected with other methods.
(Chiu & Tavella, 2008, pp. 6.)

Personalization means using technological solutions to create personalized communication and services to individual consumers (Young, 2011, pp. 407). Personalization is mostly used in advertising, but increasingly also in other media content, such as music services and video games (Turow, 2017, pp. 135).

1.5 Structure of the Thesis

The thesis is divided into seven chapters, which are described in detail below. The structure of the thesis is shown in Figure 1.

This Introductory chapter introduces the background for the research and provides justification for this study. After the Introduction, the theoretical framework for this thesis is explored in chapter 2-4. Chapter 2 investigates how consumer behavior has changed in recent years and what is the role of consumption in modern society. It also examines the consumer's need for personalization and discusses how privacy is connected to consumption. Chapter 3 examines the concept of data-driven marketing. This chapter defines data-driven marketing and discusses the sources and uses of data in marketing. It also discusses the challenges regarding the use of data in marketing and why marketers need to be concerned about privacy and data security issues. Finally, chapter 4 examines the ethical considerations and challenges organizations are currently facing. It explores the concept of ethics especially from the data point of view.

Chapters 5-6 are concerned with the empirical part of this thesis. The empirical research is a mixed method study. It consists of a survey "The Use of Digital Services" made in 2018 by the Finnish Innovation Fund SITRA, and a qualitative content analysis of the conversations about data ethics in Twitter. Chapter 5 explains the research methodology and

justifies the research strategy and data gathering methods. Chapter 6 examines the research results and findings of the empirical research.

The final chapter of this thesis is Chapter 7, the Conclusions, which will summarize the findings regarding each individual research objective and will provide conclusions based on those findings. In addition, recommendations for future research are discussed along with the limitations of this research.

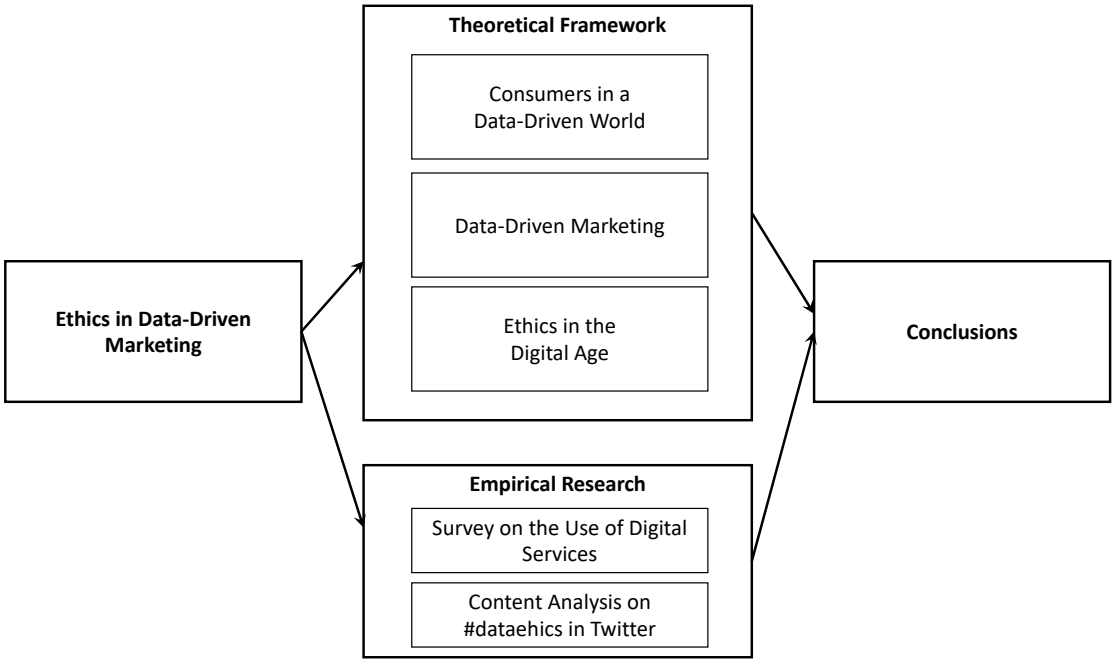


Figure 1. Structure of the Thesis.

2 Consumers in a Data-Driven World

The need to study consumer behavior is vital for marketers, as understanding the consumers and the consumption culture is good business for organizations and essential for success. Organizations exist to satisfy needs, and the needs of the customers cannot be satisfied unless marketers understand the customers and their behavior, the societal dynamics, and how these affect organizations. (Solomon, 2016, pp. 8; Firat et al., 2013.) Modern consumers are empowered by technology, and marketing has become more interactive as consumers share feedback about organizations online. Brand loyalty has decreased as consumers can choose from a wide variety of products and services. Therefore, the only way to ensure the success of marketing activities is to be customer-oriented and deliver value to the customers. Consumers are likely to choose a product or service that provides the best perceived value for them. Hence, a customer-driven marketing strategy requires sufficient understanding of the consumers and the marketplace. (Kotler et al., 2016, pp. 13-14, 19-20.)

This chapter examines what kind of implications the changes in society and consumer behavior pose to marketing and what consumers expect from organizations. Especially crucial for marketing are the increase in consumption activities, technological advances, and the need for individualized communication. This chapter will also explore what kind of qualities consumers value the most in organizations.

2.1 The Role of Consumption

The society we live in today can be described as a consumer society. Obviously, we live in a world that is full of things, but more importantly, the role of consumption has increased. Earlier people were defining themselves mostly through their role in the production process, i.e. work, whereas today consuming is becoming more prominent in constructing our personal and social identities. For example, people's personalities, values and aspirations can be expressed through the products they buy and the services they use. Consumption also keeps the economy going, because as long as consumers

continue buying, the producers keep on producing. The idea of a consumer society is based on this notion that our social lives are organized increasingly according to our roles in the consumption system. (Solomon, 2016, pp. 36, 39.) At the heart of consumer society is the assumption that consumers exercise their freedom and make their lives more meaningful by purchasing. Consumption can also be seen as the ultimate goal instead of an element of fulfilling needs. (Firat et al., 2013.) In this societal context, marketing plays an important role. Marketing creates jobs by increasing demand for products and services. It also helps organizations introduce product innovations and improvements to help or enrich consumers' lives. (Kotler & Keller, 2011.)

2.2 The Internet and Social Media

The Internet can also be seen as an essential aspect of modern society, and it has influenced consumer behavior tremendously. The amount of information available for consumers means that marketers cannot always trust that consumers will find them, or that the conversations about a brand are always favorable. The Internet has become an abundant source of marketing content, as well as criticism and complaints. The competition has also expanded with the globalization of services, and organizations need to take into account the industry, competitors, and business opportunities at a global level. (Kotler et al., 2016, pp. 26, 29.)

Social media has made it possible for people to build relationships without geographic and demographic barriers, and it gives people a sense of belonging in a community. (Kotler et al., 2017, pp. 22.) Creating and sharing content is easy for all age groups living in different parts of the world because of the widespread access to a variety of technological devices. The information flow has shifted from the downward movement from organizations to individuals to across individuals as well. Social media also enables a culture of participation because people are contributing instead of just joining social media platforms. (Solomon, 2016, pp. 16-17.)

Marketing communications have an impact on brand and product awareness, but they are not as trusted as the opinions of fellow consumers (Yarrow, 2014, pp. 28). Today, the key drivers for customer satisfaction and loyalty are peer reviews and user ratings, whether they are accurate or not (Solomon, 2016, pp. 336). Customer communities are becoming more powerful and vocal in sharing stories about brands. The credibility of random stories about brands is viewed more trustworthy than targeted advertising, and social communities have more influence on consumer behavior than marketing communications or even personal preference. Consumers are more eager to believe in friends, families, and social media communities than in marketing communications, and people turn to social media when they need advice and trust them more than advertising or expert opinions. Therefore, organizations should view consumers more than targets, and instead of peers and friends of the brand. (Kotler et al., 2017, pp. 20-26.) As consumers are turning to the Internet in their information search, building long-term relationships with customers and engaging in interactive marketing is vital for organizations (Cetinã et al., 2012).

Kotler et al. (2017, pp. 29) also claim that in the history of marketing, connectivity has been the most important change. Mobile devices have changed the way consumers behave as they enable consumers to compare prices and make buying decisions anywhere at any time. Connectivity also affects the perceptions of organizations' competitors and customers, as collaboration with competitors and co-operation with customers is essential. Organizations need to realize that in order to succeed, they need to collaborate with external sources. Connectivity should not be viewed merely as a technological issue, but a more holistic view is necessary. The Internet should be seen as a means to deliver superior customer experience and to build customer communities. (Kotler et al., 2017, pp. 28-30.) The Internet can indeed be seen as the cornerstone of customer-centric marketing, and its importance is crucial in the decision making phase of information seeking. To succeed in this phase, marketers need to attract customers with improving digital assets like product-related websites, word-of-mouth applications, and customized marketing. (Court et al., 2009.)

2.3 The Need for Personalized Communication

One problem facing consumers is the vast amount of choices they have due to the growing number of information. A real challenge for consumers is how to narrow search results when enormous amounts of information about products and organizations are available in seconds. (Solomon, 2016, pp. 325, 335.) Consumers do not want irrelevant information, but instead they want individualized communication matching their own needs (Grandhi et al., 2017). Every organization needs to add value for its customers, and the value received from personalization is the fit of the product or service to the customer's needs and the convenience of it being delivered proactively (Chellappa & Sin, 2005; Kotler et al., 2016, pp. 12).

As the amount of information has increased, the consumers have actually narrowed down the brands they are willing to consider when making a purchase decision. Consumer-driven marketing bases on the notion that consumers are increasing their control over information and "pull" relevant information to them actively. (Court et al., 2009.) As marketers operate as links between the organization and its customers, they need to shift away from the traditional one-way marketing perspective, where organizations publish content to their audiences. To succeed in the new digitalized business environment, marketers need to engage with customers in their preferred place of consuming and provide customers with personalized and individualized content. (Petersen et al., 2014.)

2.3.1 Personalization and Trust

The implementation of personalization depends on the organization's ability to obtain and handle data about consumers, and on the consumers' willingness to share their personal data and use personalized services. Though many organizations do have the ability to track consumer behavior, the concerns for privacy have inhibited consumers in sharing their data. Therefore, trust plays an essential role in consumers' willingness to share their information and in their buying intent. (Chellappa & Sin, 2005.) König (2012, pp. 29) also

mentions that consumers' buying decisions are increasingly basing on ethical issues. From the consumer's point of view, everything that an organization does affects the brand. To stay competitive, organizations need to do more than follow the minimum requirements of legislation and accept their roles as important players affecting the society at large. (Kotler et al., 2016, pp. 30.)

One of the biggest challenges for organizations in the modern marketplace is how to overcome the pitfalls of short-term thinking. Organizations cannot be lulled into believing that current sales or profits are the best measures or objectives for business, that sales and marketing always produce new customers for the organization, and product and services differentiation always creates value. Instead, organizations should understand that ensuring customer satisfaction and loyalty creates value to the organization in the long term. (Peppers & Rogers, 2011, pp. 22.)

The next chapter takes a more detailed look at how marketers use data to gain competitive edge. The aspect of privacy and other ethical implications are further explored in chapter 4.

3 Data-Driven Marketing

A marketing strategy's success is measured by consumer response, and organizations can define the market and identify business threats and opportunities by using consumer data (Solomon, 2016, pp. 8). Information itself has little value, but the value comes from customer insights drawn from that information. Creating value for customers and building long-lasting relationships requires delving deep into the wants and needs of the customers. These insights can then be turned into better marketing decisions. (Kotler et al., 2016, pp. 101.)

However, it can be very challenging to obtain customer and market insights, as customer needs and buying motives can be difficult to analyze. Creating customer relationships takes a lot of work from identifying the customers and their needs, designing offerings and setting prices, to promoting and product development. Hence, in order to gain effective customer insights, organizations need to manage marketing information from various sources. (Kotler et al., 2016, pp. 13, 101.)

By definition, data-driven marketing is concerned with collecting and combining data from online and offline sources, then analyzing the acquired data and gaining insights about the customers and their behavior. Thus it enables highly personalized communication with the target audiences. The trend in data analysis is shifting towards predicting the future, i.e. using the data for strategizing and anticipating customer needs, where technology plays an important role. Building predictive models helps companies in establishing customer-centric processes, and data can be used to identify customer needs and factors that influence the consumer decision-making process. Integrating the analysis of internal and external data can help companies in the development of products and services. The gains can be richer content for customers, acquiring new customers, and preserving current customers. Ultimately, this can lead to cost avoidance or reduction and increased productivity and efficiency. (Grandhi et al., 2017.)

Companies also need a high degree of innovativeness to be able to differentiate from the competition. This requires developing new products continuously to satisfy customer needs. Profits are also influenced by the degree of innovation. If companies are not able to launch socially accepted, up-to-date products, they are left behind in the market, and consumers will begin to look for a substitute. (König, 2012, pp. 23.)

So far, we have explored the key drivers for using consumer data. The challenge with data is not so much the volume, but the quality of information and how the obtained information is used (Kotler et al., 2016, pp. 101). In the following subchapters I will look at the sources and uses of data in more detail.

3.1 Sources of Data

Kumar et al. (2013) categorize the sources of marketing data into three groups: traditional, neurophysiological, and digital data, shown in Figure 2. I have adapted the original figure by adding Big Data to the data sources. In the figure, the sources of data are grouped by the volume of data, Big Data being the largest in volume, and neurophysiological data the smallest. In the following subsections I will look at each data source starting from the smallest data, i.e. neurophysiological data. Though the focus of this thesis is on digital data, traditional and neurophysiological are explained briefly in order to get the full scope of data sources in marketing. Also, Big Data will be examined in more detail.

3.1.1 Neurophysiological Data

Neurophysiological data is becoming increasingly popular in marketing though it is an expensive method for obtaining data and has ethical challenges. Organizations can use neurophysiological data for reading customer's emotional response, thereby creating emotional insights that are more accurate and scalable. For example, eye tracking can be used in marketing to examine visual attention in outdoor or print advertising. Eye tracking can produce more accurate results on how brands are gaining attention in a

supermarket shelf compared with consumers' own reports. In addition, various physiological measuring techniques can be used in combination to measure the positive or negative response to advertising or media engagement. (Kumar et al., 2013.)

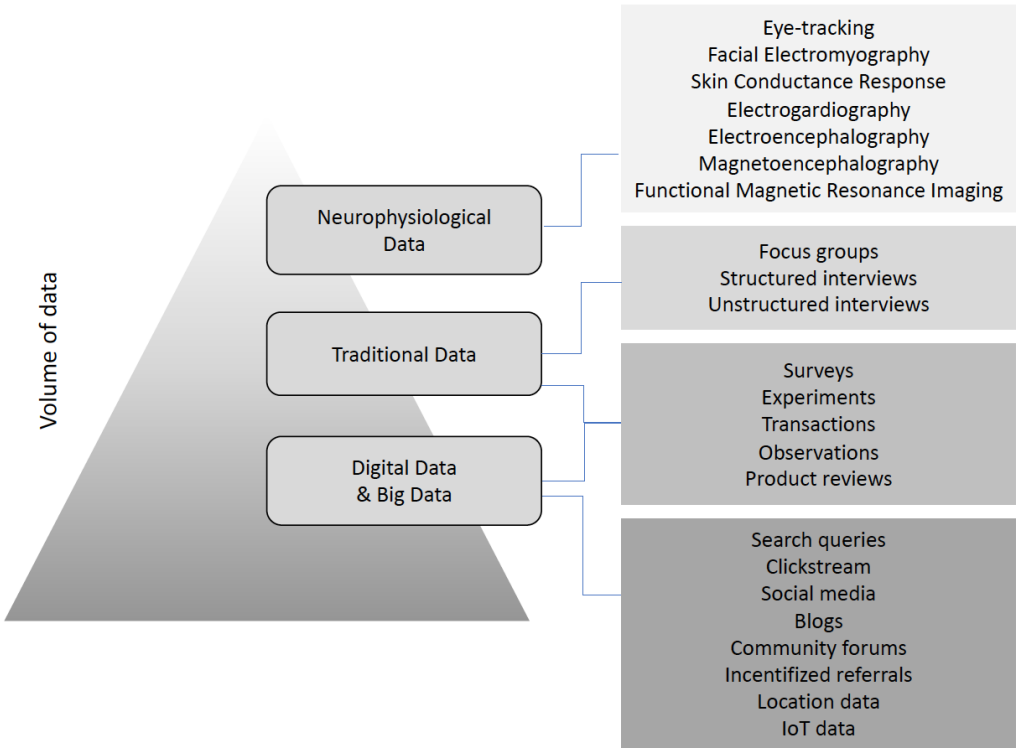


Figure 2. Sources of Marketing Data (Adapted from Kumar et al., 2013).

3.1.2 Traditional Data

Marketing research produces customer insights regarding specific marketing decisions, e.g. purchase behavior and satisfaction, market share, or market potential (Kotler et al., 2016, pp. 105). Traditional data sources include surveys, observations, focus groups, and interviews. Many of the traditional data sources can be overlapping with digital data as they can be conducted online as well. For example, observation can be executed as a physical observation e.g. in a store or online. (Kumar et al., 2013.) Chiu & Tavella (2008, pp. 6) also mention that traditional marketing research can be used to create high-level strategies on a macro level, and combining this with a more detailed information from data mining can result in effective insights, thus increasing competitive advantage.

3.1.3 Digital Data

Organizations have access to a vast amount of digital data, both internal and external. Many organizations have extensive internal databases, which collect consumer and market data. Harnessing the information well from various sources can produce effective insights about customers and create competitive advantage. For example, the internal database can collect information e.g. on customer transactions and buying behavior, customer satisfaction or service problems, sales records, inventories, and competitor activities. Challenges with these kinds of internal data is that the data is gathered for different purposes and requires skill to integrate all the data into marketing insights. In addition, the amount of data is enormous and can become outdated rather quickly. (Kotler et al., 2016, pp. 103-104.)

Organizations can also improve strategic decision-making by utilizing marketing intelligence. Collecting and analyzing public information about consumers, competitors, and the market enables organizations to gain more understanding of the consumer environment and assess competitors and business opportunities and potential threats. Marketing intelligence can be implemented e.g. by observing customers and monitoring online conversations or competitors' activities. (Kotler et al., 2016, pp. 104-105.) Kumar et al. (2013) mention a variety of ways organizations can use digital data, e.g. search and clickstream data, and data from social media and blogs. Social media channels have become very popular in marketing, but marketers need to understand that effective social media marketing campaigns need to be tied with the organizational strategy and performance measures. In addition, organizations need to understand, which social media platforms are used by their current or potential customers. (Kumar et al., 2013.) One form of digital data is Big Data, which will be discussed further in the following subchapter, due to its significance in business.

3.1.4 Big Data

In addition to the digital data sources mentioned above, some of the most emerging sources of Big Data are location data from mobile devices and machine data from IoT applications. (Brown et al., 2011; Jiménez-Zarco, 2019). What is characteristic of Big Data is the massive volume of data and the rapid pace and variety of data being gathered. Hence, Big Data is commonly characterized by volume, velocity, and variety. Volume can be seen as the primary characteristics of Big Data. Enormous amounts of data is created every day through online transactions, emails, videos, images, etc., and the amount of data is said to double in size every two years. The Internet of Things (IoT) is one of the main reasons for the substantial increase in the volume of data, as different devices ranging from cars to toys and appliances are computerized. Velocity refers to the rapid production of data. Through Big Data, organizations can get insights on customer behavior rapidly and can react to any changes instantaneously. In order to maximize the value from Big Data, velocity is also required for all the processes. The sources of Big Data are varied and that makes it so big in volume. Big Data varies from structured data (e.g. files and databases) to semi-structured or unstructured data (e.g. social media data). Big Data consists more of unstructured behavioral data than traditional data. (Erevelles et al., 2016; Sagiroglu & Sinanc, 2013.)

In addition to volume, velocity, and variety, Big Data can be characterized with veracity and value. Veracity underlines the importance of the quality of data. As the data created ever increases, the quality of data is becoming more and more critical, as not all data is accurate. Data needs to be objective, truthful, and credible, and information security needs to be ensured. Mutually important is the issue of value. With the volume of data, an important question is how to ensure the usefulness of data, and for organizations data must bring financial benefits. Therefore, eliminating unimportant and irrelevant data is needed. (Erevelles et al., 2016; Stieglitz et al., 2018.)

3.2 Uses of Data

The importance of data depends on its ability to influence marketing decisions. Through traditional and digital data along with Big Data, organizations can get a vast amount of information e.g. on market trends and transitions as well as customer segments. This information helps organizations make strategic decisions. (Kumar et al., 2013.) Data needs to be turned into information that can be used for actionable business tactics. Therefore, data mining using statistical methods is needed for processing and analyzing the large quantities of data. To gain the most effective insights, organizations need to combine both traditional market research and data mining. Market research provides insights on a macro level, but data mining helps to identify hidden information that cannot be obtained through traditional methods. (Chiu & Tavella, 2008, pp. 6, 139.)

One important aspect of data use is that in order to succeed, organizations need to create full portraits of their customers instead of just a series of snapshots. This means that organizations need to gather and combine all the data from their customers together, from basic customer data to transactions and browsing history. (Van Bommel et al., 2014.) Also, a big challenge is how to combine the acquired data with marketing strategy. Organizations can improve branding by using customer data about their buying behavior, and then increasing engagement with more personalized marketing. In addition, organizations can use Big Data analytics for evaluating which marketing tactics work and which not and adjust the strategy accordingly. (Jiménez-Zarco, 2019.)

Digital data and Big Data have a strong influence on segmentation and targeting, and they have enabled the personalization of marketing communications. These areas of marketing are therefore explored further in the following subsections.

3.2.1 Segmentation and Profiling

The traditional way to start with marketing is segmentation, i.e. categorizing consumers into groups based on different kinds of factors like demographics, psychographics, and

behavioral profiles (Kotler et al., 2017, pp. 47). Profiling, on the other hand, means creating unique descriptions of the discovered segments. Understanding the audiences is important for effective advertising and communication. Segmentation can be considered as one of the most important analytics, because it enables the organization to identify groups of people who share similar interests and to match the right products, messages, and incentives to the right audiences. Segmentation needs to be objective-driven to be effective, and the objectives can be e.g. to understand audience behavior, needs, or values. Moreover, data can be used to profile individual customers in a way that finds the most suitable products and services targeted to the customer. (Chiu & Tavella, 2008, pp. 140, 195; Fan et al., 2015.) McDonald (2011, pp. 30) notes that in addition to specifying target customers, correct market information enables organizations to effectively measure market share and growth, recognize relevant competition, and formulate a market strategy. However, Fan et al (2015) state that segmentation is becoming increasingly challenging because of Big Data and the volume and variety of data.

3.2.2 Targeting and Personalization

Data has been used for a long time for segmentation and targeting purposes. However, Big Data has increased the effectiveness of data use by enabling real-time personalization. Big Data is especially useful in retailing, as the industry has access to a large variety of data from online purchases, social media conversations, and location data from mobile devices. Individual consumer's behavior can be tracked and modeled in real-time to recognize when a customer is about the make a purchase decision. The customer can then be nudged into completing a purchase with bundling products and offering reward program savings. (Brown et al., 2011.)

The difference between targeting and personalization is that targeting aims to reach specific groups of consumers, whereas personalization aims to reach individual consumers. Targeting is based on consumer segments and profiling, and personalization utilizes personal data like name, address, and email address. Moreover, personalization is mostly used in an organization's own media like newsletters, whereas targeting is usually used

in paid media like website banners and social media advertisements. (Strycharz & Smit, 2019.) The benefits of personalization for marketers include improved customer satisfaction and customer loyalty, making it more difficult for competitors to lure the customers away. (Chellappa & Sin, 2005).

The techniques that use personalization include online behavioral targeting, email marketing, social media advertising, applications and notifications, on-site personalization, customization, and price differentiation. Online behavioral targeting means that an individual's online behavior is stored in cookies, i.e. text files stored on PCs and mobile devices. In addition, users' data from social media sites can be used for advertising. Behavioral data can be used to find out which topics are interesting to an individual and advertising can then be adjusted accordingly. Online behavioral targeting is employed using automatic algorithms, and with these algorithms, personal and contextual data rises to be central in personalization. In email marketing, personalization is a standard process. It is employed by including information about the recipient, such as adding the person's name in the greeting or sending special offers on the customer's birthday. In addition, emails can be personalized in terms of content, based on demographic or behavioral data. The aim of this kind of personalization is to make communication more meaningful. Personalization of emails is an effective method as it increases the open rates of emails. (Strycharz & Smit, 2019.)

Social media advertising is mostly based on general targeting, as personalization activities in social media generates more negative reactance from consumers. Personalization of content in mobile applications is not widely used because of users' privacy concerns. Still, in-app personalization is more accepted by consumers than personalization of notifications. On-site personalization means that a website's look, feel, and content can be adjusted according to the individual's personal preferences. This type of personalization is often too costly, and therefore personalization is mainly used only in landing pages from emails. Customization is also sometimes used, meaning that individuals can for example adapt a website to their own personal needs by using filters. Personalization is

then self-driven by the customers instead of being automated by the organization. Finally, the price of an online product or service can be personalized based on information about an individual consumer. (Strycharz & Smit, 2019.)

Cloud computing has enabled effective personalized filtering, which means that information can be filtered by different criteria for individual users. The system can then predict what kind of information is relevant to the user and will be filtered out. Google, for example, uses previous search words, location, and social media data to customize search results. Facebook uses personalized filtering based on how users interact with each other. Other services like Amazon and Netflix are also based on heavy personalized filtering, and in the future, personalized filtering it is bound to increase even more. (Bozdag & Timmermans, 2001.) In addition, location data can be seen as an important source of personalized marketing information. Mobile technology enables organizations to utilize location-based services to personalize communication to specific locations at specific times. In advertising, location can be used to deliver advertisements or product recommendations based on the user's current or predicted location. (Fan et al., 2015.)

3.3 Challenges with the Use of Data

Though marketers gladly embrace consumer data, and intelligent data use can be seen as a way to increase competitive edge, the use of data also poses many challenges for organizations. Still, the risks and potential harms of using consumer data often remains unrecognized. The essentialness of technology in modern life and the convenience of tailored marketing messages may leave consumers unaware of the consequences of their information sharing. (Walker & Moran, 2019.) The knowledge about different data collection methods and personalization techniques are scarce, making the consumers more vulnerable and unable to control the use of their data. (Strycharz & Smit, 2019.)

For organizations, using Big Data may result in information chaos. With massive amounts of data from various sources, organizations may struggle with identifying important data

from unimportant one and finding the right metrics. Also, organizations need to be very careful with data storage and security so that consumer data is not compromised. (Grandhi et al., 2017.) Often, data protection choices are made based on the tradeoff between profitability and privacy, whether consciously or unconsciously. This is especially the case when organizations share their data with outside parties. This is often done to create additional revenue, but the threats to privacy decrease brand value and trust, and can also result in legal penalties. (Schneider et al., 2017.)

If we think about segmentation, its pitfall is that it can sometimes lead to stereotyping by ignoring the subtleties of human character. Therefore advertising can paint a too simplistic image of a certain group of people. In addition, technology has enabled increasingly sophisticated marketing methods, and consumers may not be able to recognize advertising, which in turn decreases transparency. (Bivins, 2009, pp. 197-198.) There are even more serious threats posed by consumer surveillance and data-driven advertising, such as increasing market discrimination and marketers' ability for social control (Nadler & MacGuigan, 2018). The use of Big Data also raises the question of the effectiveness of algorithms. Predictions made by algorithms may be incorrect. Hence consumers may be exposed to unwanted, unappreciated, or offensive content. In addition, marketers may put too much emphasis on technologies instead of the customer relationships, therefore increasing the risks of communication. (Walker & Moran, 2019.)

3.3.1 Challenges with Personalization

Some researchers speak of the personalization-privacy paradox, which means that personalization can either enhance or diminish consumer engagement. If personalized communication raises concern for privacy, then engagement with the company decreases. (Aguirre et al., 2016.) Personalization does impose unique challenges in regard to other marketing techniques. While consumers are willingly accepting the benefits of targeting, personalization of marketing is not accepted so enthusiastically. The most critical issue with personalization for consumers is the loss of privacy as well as the possible misuse of data. Privacy concerns often lead to consumers feeling skeptical about advertising,

hence avoiding it. Though some marketers believe that younger audiences do not value privacy as high as older consumers, research has shown that privacy is a critical issue for all age groups. Individuals lack control over their personal information, because they may not be aware of data use, or they do not understand fully how their information is being used. Many organizations do not have clear privacy policies. Another risk is that organizations have much personal information about individual consumers, and therefore people's weaknesses can be exploited, or the organization can employ social sorting, i.e. refusing to advertise to riskier targets. (Strycharz & Smit, 2019.) One critical pitfall of personalization is that the algorithms for filtering content are not transparent, and the filtering of information occurs silently before reaching the consumer. In order to increase transparency, users should be notified when and on what basis filtering occurs. (Bozdag & Timmermans, 2001.)

Advances in technology and the changes in marketing practices can have serious impacts for consumers, marketers, and the society. These challenges go beyond privacy issues, and more problems can arise with the ongoing evolvement of technology, such as the Internet of Things, artificial intelligence, and virtual or augmented reality. (Walker & Moran, 2019.) Some of the challenges related to data use can be controlled through regulation. The following subchapter explains how the European regulation aims to ensure data privacy. However, as described in the next chapter, merely following regulations is not enough, and the law cannot keep up with the rapid development of technology. Therefore, organizations also need to establish ethical procedures for data use.

3.4 Data Regulation

Regulatory policies regarding data privacy are rules defined by governments. The EU has taken a strict view on data privacy protection, and it has defined that personal and sensitive information includes e.g. health status, religion, and sexual identity (Ess, 2014, pp. 65). The current General Data Protection Regulation, GDPR, which came into force in May 2018, aims to protect the citizens from privacy and data breaches. It applies to all companies processing and holding the personal data of the EU residents, no matter

where the company is located. Organizations that do not comply with the GDPR may be heavily fined. If an organization processes data for another organization, the regulations of GDPR apply for it too. (GDPR.eu, 2018.) According to the GDPR, individuals have the right to know when their information is being collected and to review that information. Furthermore, individuals need to give consent to the collection and use of their personal information. (Ess, 2014, pp. 65.) In addition, opposed to the earlier privacy laws in the EU, the GDPR requires organizations to address data privacy at a high level early on in the design process, instead of being an add-on (Garcia-Rivadulla, 2016).

The GDPR has received some criticism as well. Its implementation varies among different countries, as some countries have taken a stricter line than others, and the strategies on penalties vary from country to country. One of the biggest challenges is how to handle technology coming outside of the EU. Big data and AI are seen as the most significant threats, especially when China is increasing its role in both of the two technologies. (Computer Weekly, 2019, pp. 7-10.) In addition, according to a recent survey by Sitra, as the operating environments of digital services are complex and multilayered, it is often impossible for individuals to find out all the third parties that have access to their data, therefore making it hard to control the uses of personal data (Nupponen, 2019).

4 Ethics in the Digital Age

So far, we have examined how the changes in society and customer behavior have changed marketing and how data is used in marketing. This chapter examines what kind of ethical implications these changes bring for organizations.

On a general level, ethics is concerned with what is right or wrong, and it provides a standard for human behavior and morals. One of the most important aspects of ethics is trust, as ethical behavior builds trust among people. If we think about the balance between ethics and law, both of them are concerned with creating criteria for acceptable behavior, but ethics is a much broader concept than the law. The law is concerned with the minimum regulation for ensuring public order, whereas ethical behavior may not be covered by the law at all. Ethical behavior bases on voluntary action of shared values. For example, taking care of the elderly or not telling a lie are examples of socially acceptable conduct, which is not covered by the law. (Paliwal, 2006, pp. 6-7, 26; Painter-Morland, 2004, pp. 3.) Following regulations does not mean that we are behaving ethically, and that is why in addition to regulation, we need to set up ethical rules for business as well.

Ethics can be viewed from two different perspectives: consequentialism and deontology. They differ in the way morality with its objectives and principles is defined. Consequentialism emphasizes the consequences of actions, whereas deontology is concerned with duty and the decision-making process. In consequentialist theories, the probable outcome or consequences determine the morality of an action or a decision. One of those theories is utilitarianism, which is the most popular of ethical theories. Utilitarianism is based on the notion that moral choices should always create the most happiness for the largest group of people. At the heart of utilitarianism is the belief that most people desire ultimate good, e.g. happiness or pleasure. More recently, satisfaction has grown popular as the ultimate goal for utilitarians. Deontological theories, on the other hand, are based on actions and decisions governed by rules or principles. The most persuasive of these

theories is Kantianism, created by the philosopher Immanuel Kant. Instead of consequences, the morality of a decision is based on the rightness of rules. Universal rights, such as the freedom of speech, freedom of consent, and the right to privacy, are in accordance with Kantian rules. (Singh & Mishra, 2018.)

Both of these approaches have strengths and weaknesses. Utilitarianism is an objective and a very practical theory, being easy to apply and able to suit human behavior and decision-making. All people are treated equally and instead of individuals, the emphasis is on the wellbeing of a large number of people. Utilitarianism is thus concerned with maximizing social welfare. On the other hand, the challenge of this approach is that the concept of happiness is vague and broad, and there is no objective way to measure whether an action is right or wrong. Then again, Kantianism as a theory is logical and well-structured, emphasizing reason and freedom. However, on a practical level, it is too strict and does not take circumstances into account. Solving complex dilemmas may become complicated if the individual does not know which rules to follow. In addition, some rules may be contradictory, making ethical decision-making challenging. (Singh & Mishra, 2018.)

4.1 Ethics in Business

Business ethics means that general ethical rules are applied to business behavior. These rules of business can be used in order to judge the appropriateness of business activities. Ethics in business is needed because organizations are part of society and they need to operate as responsible corporate citizens in the society. (Paliwal, 2006, pp. 7-8.) As regulations cannot cover all situations and are hard to enforce, organizations need social codes and rules for professional ethics to govern their activities. Acting for the long-term interests of customers and the environment is essential for socially responsible organizations. Sustainability does not mean merely caring about the environment, but it also involves social aspects, i.e. building a long-term sustainable society. (Kotler et al., 2016,

pp. 30, 90.) Many organizations have, therefore, employed corporate social responsibility strategies as the driving force of business, combining commercial responsibility with social responsibility (Solomon, 2016, pp. 58).

Merely developing ethical codes of conduct is not sufficient; they also need to be followed (Kotler et al., 2016, pp. 90). Painter-Morland (2004) states that although currently organizations are wakening up for the importance of ethics, it has proven difficult to incorporate ethics into business practices. Sometimes ethics are viewed with a "check-box" mentality and treated only as insurances against corporate liability. With this mentality, ethics is seen as a compulsory set of principles that need to be applied to business activities, becoming the last thing to be considered after decisions are made. When ethics are integrated into business operations, it guides the actions from early on and is an important part of the organizational culture. Instead of merely identifying and applying relevant rules, ethics in business should embrace practical wisdom, which is intuitive and continuous and enables organizations to make decisions in situations which are not covered by regulation. (Painter-Morland, 2004, pp. 1-3.)

Ethical dilemmas in business are usually complex with multiple choices and alternatives. The outcomes of ethical decisions may be uncertain, or they can be favorable one way but unfavorable the other way. For example, an organization may need to choose between short-term or long-term outcomes. Often, organizations have to balance between economic and social performance. Ethical dilemmas can also have significant consequences on different stakeholders, therefore making decision-making risky. (Paliwal, 2006, pp. 85-86.)

There are multiple ways to handle ethical dilemmas. Earlier, we examined the concepts of consequentialism and deontology. According to Singh & Mishra (2018), these traditional theories do not suit well in the business domain. Brinkmann (2002) proposes that business and marketing ethics should be viewed as a form of professional ethics. The

four approaches to business ethics and decision-making are, therefore, the moral conflict approach, the professional code approach, the professional role morality approach, and the moral climate approach. These will be discussed further below.

The moral conflict approach is especially suitable for business, as it is full of conflict. As ethics can be quite an abstract issue, the moral conflict approach makes ethics easier to understand. In this approach, ethics is not considered until faced with an urgent conflict or dilemma. This approach, therefore, focuses on acute ethical conflicts. *The professional code approach* handles ethical dilemmas through rules or codes. The challenge with this approach is that codes can be imprecise and difficult to enforce. *The professional role morality approach* bases on rights and obligations attached to the role itself. Professionals are more responsible for their actions, but as morality is dependent on the role, it may reduce moral responsibility. *A moral climate approach* is the most holistic of these approaches, and bases on the idea that ethics is part of the overall climate or culture, and its participants both shape and are shaped by it. Because the cultural and social context is included in the decision-making process, issues and conflicts can be examined on a broader level. All of the aforementioned approaches have strengths and weaknesses, and these approaches are not mutually exclusive but rather complimentary and can be combined. (Brinkmann, 2002; Bivins, 2009, pp. 194-196.)

4.2 Marketing Ethics

Marketing has raised some of the broadest and disputed ethical issues in business. The marketing field has been accused e.g. of dishonesty, manipulation, and invasion of privacy. (Jamnik, 2011.) Earlier, ethical problems in marketing focused on product safety, price fixing, deceptive advertising, and unethical information collection. Today, due to environmental change, the marketing field is challenged by issues regarding fairness, honesty, and product and human resources management, which are important factors influencing the perception of product quality. (Lee & Jin, 2019.) Nantel & Weeks (1996) point out, that if the ultimate goal of marketing is really to satisfy customer needs, then marketing needs to focus on issues beyond short-term gains, therefore underlining the

need for ethical considerations. Kotler et al. (2016, 90) also state that ethics and social responsibility are involved in almost every aspect of marketing.

If we think that the final objective of marketing is mainly to satisfy customers' needs, then we are looking at marketing from a utilitarian aspect. Therefore, marketing ethics can be seen as utilitarian ethics. If then the core of marketing is utilitarian, marketing itself could be seen as a highly ethical field. But the concept of happiness or satisfaction as the ultimate goal is complexing. For example, if we think about sugary products, they can respond to the immediate needs of consumers, but on the other hand, they can cause medical or dental issues. Therefore the ethical challenge in marketing is what needs are to be satisfied and of what people. Customer satisfaction often aims at fulfilling short-term needs, and hence answering those needs can actually cause problems for the consumers and the society in general in the long-term. (Nantel & Weeks, 1996).

Nantel & Weeks (1996) propose that marketing could benefit from a more deontological approach, which often bases on a code of ethics. Successful organizations understand that consumers have other needs that outweigh the fulfillment of immediate needs. The manufacturing of a product or the marketing process impact the consumers' willingness to buy. For example, environmental factors or manufacturing conditions can affect buying intent even if the end product or service is of good quality and is able to satisfy customer needs. Consumers are willing to engage in boycotts and protests in order to influence business conditions. As the needs of the customers have evolved in this manner, it brings justification to move towards a deontological approach. Therefore, organizations should have a set of explicit rules to define the ethicality of actions. Of course, the existence of rules does not necessarily mean that the actions of an organization are ethical. (Nantel & Weeks, 1996.)

Murphy et al. (2007) continue Nantel & Weeks (1996) ideology, stating that sometimes good intentions and sense of honor are not sufficient in building long-lasting relationships. Instead, they propose that marketing should embrace virtue ethics. In order to

sustain long-lasting relationships with their customers, organizations need to employ virtues, i.e. good habits, in their operation. Instead of focusing on ethical dilemmas, virtue ethics focuses on the individual and the organization. There are a number of virtues associated with relationship marketing, e.g. honesty, fairness, reliability, integrity, commitment, and trust. Of these virtues, especially trust is seen as an essential element of successful relationship building. Building trust requires honest and reliable product or service claims, as well as integrity and consistency. In addition, the wellbeing of consumers needs to be a priority. If the consumers are viewed merely as targets or objects of marketing, the organization cannot be customer-oriented, nor is it able to build long-lasting relationships. In addition to trust, ethical relationship marketing requires commitment and diligence. Trust leads to commitment, which means that promises are made and kept. Keeping promises is therefore an act of commitment. Diligence is also crucial in long-lasting relationship building, i.e. organizations must be persistent in maintaining ethical decision-making. In addition to the above-mentioned virtues, organizations also need integrity, fairness, respect, and empathy in their relationship management. (Murphy et al., 2007.)

It has to be noted that these virtues can be difficult to operationalize in practice. Virtues can be seen as the ideals which marketers should aspire, but real-life situations are often complex and multifaceted. (Laczniak & Murphy, 2019.) It is the complexity of ethical issues that brings fundamental challenges to organizations' ethical decision-making. Hence it can be challenging to apply pre-determined rules into real-life situations. (Reinecke & Ansari, 2015.) Ethical dilemmas often arise when a marketer's view differs from the views of the customers or the general public (Cui & Choudhury, 2003). The following subsections examine some of the concrete ethical dilemmas from the point-of-view of the traditional marketing mix, i.e. product, price, promotion, and place. Technological advances have brought new ethical dilemmas to these aspects, especially to pricing and promotion.

4.2.1 Product

Some of the product-related ethical issues are concerned with product quality, safety, eco-friendliness, packaging, and branding (Lee & Jin, 2003). For example, organizations may mislead consumers into thinking the quantity of the product is larger by packaging products into larger packages. In addition, the packaging may include wrongful or misleading information. Packaging has a strong effect on purchase intent and consumers are well aware of the organizations' capability to mislead consumers with packaging. (Parıltı et al., 2015.)

Another important issue are harmful products. If an organization has knowledge that its product is some way harmful but fails to take any corrective actions, it can be seen that it intentionally harms consumers. However, the evaluation of a product's harmfulness is not always easy, and the interaction between consumer characteristics and marketing practices affect the ethical evaluation of some products. For example, some products may become harmful if they are abused or misused. (Cui & Choudhury, 2003.)

4.2.2 Price

The price of a product or service is one of the most important attributes to consumers. Traditionally, ethical challenges have arisen when using false reference prices when organizing a sale, thus misleading the customer. (Nantel & Weeks, 1996.) Other ethical challenges with price are e.g. misleading or deceitful price advertisements, price fixing, and price discrimination (Sihem, 2013).

Technological advances have brought new ethical implications concerning price in the form of price variation, which means that organizations may use information about customers to differentiate prices. Though regulations forbid unlawful discrimination, it might be unclear for organizations when price differentiation becomes unethical or even unlawful. It is especially attractive for online stores to attract customers and increase

profits, and flexible and dynamic prices are often used in the airline and hospitality industries. However, some consumers see this kind of tactics as unethical. One important dilemma for marketers is what kind of data can be used for price differentiation without it becoming unlawful. (Strycharz & Smit, 2019; Murphy, 2017.) It also has to be noted that fair pricing enhances the trust of customers and strengthens customer relationships (Lee & Jin, 2019).

4.2.3 Promotion

Ethics related to promotion have a direct impact on brand loyalty, and they have the most crucial role in building strong customer relationships (Lee & Jin, 2019). Much of the ethical issues in marketing are focused on advertising, being the most visual element of marketing. Many ethical issues in advertising are linked to content and techniques, but advertising brings forward broader societal issues, like consumption being associated with happiness. Advertising can also promote values that are not appropriate for a healthy society, e.g. by focusing on material gain and promoting over-consumption. In addition, advertising is often seen as attempting to bypass rational thinking. (Bivins, 2009, pp. 186, 196-197.)

The new digitalized business environment has increased the ethical challenges of advertising with the introduction of online behavioral targeting (OBT), which means that consumers' online activities are tracked and used to create profiles of individuals (Murphy, 2017). Marketers are able to use these extensive consumer profiles to create targeted messages by analyzing data from website visits, used search terms, and purchase behavior, and combining this information with demographic and geographic data. Though OBT is a very effective advertising method compared with traditional marketing, and it enables consumers to receive messages relevant to their needs, the method poses threats to privacy and raises the issue of control. (Nill & Aalberts, 2014.) Another challenge of online advertising is the use of native advertising, i.e. portraying advertising as editorial

content, in the forms of advertorials and branded content. The ethical dilemma concerning native advertising is that consumers may not be able to differentiate between advertising and editorial content. (Murphy, 2017.)

Targeting brings forward also the ethical dilemma of marketing to vulnerable consumers. For example, children and the elderly are groups that can be seen as vulnerable consumers. Targeted marketing can be successful and welcomed in some instances, e.g. when marketing special promotions to children and their parents. On other instances targeting can be viewed as unethical, like when giving sweepstakes to the elderly or marketing high-interest loans to consumers with poor credit history. (Cui & Choudhury, 2003.)

4.2.4 Place/Channel

Place related ethics mean the honesty in distribution channels and customized distribution. Place-related ethics have a direct impact on brand loyalty, i.e. if place-related ethics are high, brand loyalty increases. In addition, place-related ethics can improve both the relationship of the consumers to the brand as well as perceived product quality. (Lee & Jin, 2019.)

One issue related to place that has received increased attention is the supply chain form manufacturers and retailers. Ethical issues related to the supply chain include employee safety, wages, and working conditions. The most powerful players in the chain should bear the strongest responsibility, and many global brands have received criticism for not ensuring proper supply chain policies of their contractors. (Murphy, 2017.)

4.3 Data Ethics

As data is at the core of data-driven marketing, we need to look more into the concept of ethics from the viewpoint of data. The ethical problems related to data arise from collecting and analyzing large datasets. Some of the key issues in data ethics are that

different datasets can be linked, merged, or re-used to identify individuals. The identification of individuals may lead to discrimination or even violence towards certain groups of people. Other critical issues are trust and transparency. (Floridi & Taddeo, 2016.) Richards & King (2014) propose that that Big Data should not be considered merely as a technical issue, but a societal one, as the decisions based upon data can have huge societal impacts. Therefore, it is important that when organizations engage in data projects that have an impact on human life, their focus is on social acceptability or preferability (Floridi & Taddeo, 2016).

Richards & King (2014) propose four principles that should be followed in order to advance the ethics of big data: recognizing privacy as information rules, recognizing that shared private information can remain confidential, recognizing that big data requires transparency, and lastly, recognizing that big data can compromise identity. In addition, in their earlier work they proposed that Big Data brings forward the issue of power (Richards & King (2013). All these five issues will be discussed in more detail in this chapter

4.3.1 Privacy

The perception of privacy is not fixed but rather fluctuates between individuals, context, and culture. The individual's perception of privacy depends on many aspects. First, the political philosophy of the society, e.g. democracy versus authoritarian society, affects how privacy is defined. In addition, the person's power and social status affect the perception of privacy. Finally, on an individual level, the changing needs and desires of an individual affect the need for privacy on a daily level. The perception of privacy is then constantly shifting depending on the situation. (Westin, 2003.)

Privacy can be seen as one of the most vital and important issues in the modern society based on technology (Chellappa & Sin, 2005, Richards & King, 2014). While collection of data is vital for the success of organizations, they also need to recognize their responsibility in ensuring individual privacy (Kumar et al., 2013). Though some claim that privacy is a lost cause, Richards & King (2014) say that privacy is very much alive,

but the concept of privacy is changing along with the changes in society. Undoubtedly, the amount of information gathered about each individual is massive, and in that sense, privacy is diminishing. In addition, the social acceptance of shared information is increasing. Instead of thinking privacy as keeping secrets, the focus on privacy today should be in the ways we can manage the information flows. The definition of privacy is an important matter, and we need both legal and social rules for the use of information. The use of Big Data actually increases the importance of privacy. (Richards & King, 2014.)

Another issue with privacy is the ability of individuals to control the trade and uses of their information. Individuals should able to manage their personal data so that they can weigh the benefits and costs of information use. In practice, as also required by the GDPR, data processors need to keep consumers informed of what they are doing with personal data, and consumers also have to have the ability to opt-out of uses of their data. The problem with this kind of privacy management is that people often do not have an actual choice of opting out of information use if they wish to use the services offered. In addition, not many individuals have the time or the skill to go through the complex terms and conditions nor to revisit them after consent has been given. (Richards & King, 2014.)

If we think about marketing, privacy issues arise with personalization as it always indicates some loss of privacy. Though consumers value privacy high, they are willing to share their personal information if there are some benefits from it. The benefits can be monetary, but also intangible ones, like convenience. If the value for personalized services is high, then consumers are more willing to accept the loss of privacy. (Chellappa & Sin, 2005.) It also has to be noted that trust has an enormous impact on consumer's willingness to share personal information, as well as increasing purchase intent and advertising acceptance. In addition, consumers are more willing to accept personalized and targeted advertisements when they have greater control over their privacy settings. (Martin & Murphy, 2017.) However, Jackson (2018) points out that as the data

collection methods are becoming more sophisticated, consumers are not always aware of the collection and use of their data, hence decreasing their control over data.

Martin & Murphy (2017) list practices that enhance the trust of consumers. First, organizations need to prioritize data privacy in an authentic way. Second, organizations need to have a dialogue with their customers about information privacy issues, and the communication needs to remain open and transparent. Third, data privacy practices need to be aligned across all functions of the organization. Fourth, organizations need to focus on what they are doing right and highlight their own strengths instead of focusing on competitors' insufficient practices. Fifth, organizations need to commit to data privacy as a long-term objective. Seeing privacy as a strategy builds consumer trust, but organizations need to bear in mind that building trust takes time. On the other hand, trust can easily be lost if the organization does not take care of the above privacy practices. (Martin & Murphy, 2017.)

4.3.2 Confidentiality

Privacy should not be seen as an on-or-off matter, meaning that information is not merely public or private. Instead, almost all of information lies in between completely private or completely public. Much of the private data individuals share, is shared with the trust that it remains confidential. Therefore, confidentiality is based on trust. (Richards & King, 2014; Cavoukian, 1999, pp.121.) In practice, confidentiality means that personal information is accessed only by those who have the permission to access it and that there are sufficient safeguards to protect the information from unauthorized access. This requires adequate data security measures. (Cavoukian, 1999, pp. 121.)

One important notion is the distinction between primary and secondary uses of data. Personal information collected for one purpose should not be used for other purposes without the individual's consent. (Cavoukian, 1999, pp. 122.) However, Big Data has increased the secondary uses of data and personal information may be shared extensively and in unexpected ways. It is largely the secondary uses of data that make Big Data so

powerful, as they can be used to make new predictions and conclusions. (Richards & King, 2014.) It is often challenging to isolate the primary reason for data collection so that data use could be restricted. Also, if an organization collects different types of data for different purposes, the data should be segregated so that access to different parts of data could be restricted from those who do not need to access it. Moreover, identification of individuals has become the new norm, and identification is used and recorded in databases in situations like ordering goods online, subscribing a magazine, or joining a club. This protocol is not even questioned, but as anonymity can be considered as a key element of privacy, the options for conducting transactions anonymously should be increased. (Cavoukian, 1999, pp. 122-123.)

4.3.3 Transparency

Big Data is composed of small data information, which is created by information of people and their locations from sensors, cell phones, and web behavior. These small datasets are integrated to create a bigger picture of the individual consumer. The Transparency Paradox related to Big Data is that data is used to make the world more transparent, but on the other hand, data is collected silently and invisibly. Some level of secrecy is reasonable, as profitability can depend on trade secrets. Still, individuals should have the right to know the basis of decision making if organizations use personal information to make decisions. Therefore, transparency includes balancing between openness and secrecy. (Richards & King, 2013, 2014.)

Big Data increases the need for transparency. Just like confidentiality, the benefit of transparency is that it enhances trust. (Richards & King, 2014.) Transparency has a strong impact on consumer's willingness to compromise privacy, and transparency is needed not only for data as a product but for data collection processes as well. However, limitations to data access are often reasoned with privacy issues, but there are also economic benefits of restricting data access. As data is used for gaining insights about customers and their behavior, organizations gain competitive advantage from it, and data becomes a valuable commercial asset. Therefore consumers often just need to trust that their

data is being used responsibly. (Richterich, 2018, pp. 37.) Also, Floridi & Taddeo (2016) note that it can often remain unclear how transparency can be employed in practice: what data to include and to whom to give access to this information.

4.3.4 Identity

Big Data transforms the way information is processed in society, and in the future it may affect how people see the world (Cukier & Mayer-Schoenberger, 2013). Identity bases on the notion that people have the right to choose who they are. Big Data can be seen as a threat to identity as it enables organizations to gather people's phone records and social media posts as well as search or buying history. Through this extensive knowledge individuals may be pushed with tailored messages into directions the organizations want to take them. The ethical dilemmas with Big Data are therefore concerned with its ability to persuade, influence, and even restrict identity. If filters and personalization overcome intellectual choices, people's identities may fade and taken even further, this may even influence the aspect of democracy. (Richards & King, 2013, 2014.) The danger with personalization is that individuals are grouped into pre-determined categories, thus gathering like-minded individuals into echo chambers (Tene & Polonetsky, 2012).

One challenge with data and identity is that all the data gathered of individuals creates a virtual identity, which may differ from the real-world identity of an individual. Still, virtual identities may influence real-world situations, like job offerings, credit ratings, or risk profiles. Though predictive analysis is good for law enforcement and national security, it can be seen as especially problematic when dealing with sensitive information like health, race, or sexuality. Virtual identities may also affect a person's self-perception, especially with younger age groups. Moreover, if a person would want to change his virtual identity, it would be very difficult, because there are no tools for managing all the data. Even if a person would be able to delete his data from a certain service provider, there is no guarantee that the information already sold to third parties gets deleted as well. (European Economic and Social Committee, 2017.)

4.3.5 Power

Big Data also creates issues with power. As the Big Data sensors and pools are in the hands of institutions and not individuals, there is a risk that organizations become more and more powerful. The power balance between those who create data and those who utilize it is unequal, and in order to manage with this inequality, stronger protections are needed to ensure privacy, transparency, and identity. (Richards & King, 2013.) The question of transparency also lies in the heart of the power issue. Commercial Big Data remains mostly inaccessible. Data sharing in the private sector is obviously limited due to privacy concerns as well as reasons related to competitiveness. Still, this creates a power asymmetry, where individuals do not have access to data to see exactly what kind of information is gathered about them and how it is used. In practice, the power related to data is in the hand of data monopolies, which constitute of a few internet and technology companies. (Richterich, 2018, pp. 40-41.) Also, individuals often use their digital identities e.g. from Facebook or Google to access other, third-party services. While this allows for quick and easy identification, at the same time it diminishes the individual's awareness of how his personal data is used and by whom. The service providers get a very detailed and personal information of individuals, but individuals' power and freedom reduces. (European Economic and Social Committee, 2017.)

Big Data can be used for good societal purposes like fighting against terrorism and cyber threats, but this produces the dilemma of institutional surveillance. In many ways, this kind of control is necessary. But as government surveillance aims to protect us, it also means that a very detailed and extensive data of individuals is gathered. This kind of automated surveillance can be a threat to identity because it can moderate or even determine people's identities. It is therefore crucial, that individuals gain more privacy and organizations less, in order to balance the power asymmetry. It is essential for the future of Big Data societies that data collection and utilization are made more transparent, and organizations collecting and utilizing data are made more accountable. (Richards & King, 2013, 2014.)

An essential issue related to power is data ownership, which is also an area of much debate. Al-Khouri (2012) proclaims that data in itself has no value, but the information that can be drawn from it. If data has no value, then the ownership of data should not be an issue. Also, when data is generated, it is also stored somewhere. Data ownership therefore refers to the storage process, and the data is hence owned by the storage owner. (Al-Khouri, 2012.) In other words, according to this view, individuals lose the ownership of their data once they allow organizations to access it. Janeček (2018) also admits that the line between personal and non-personal data is vague, and at some point and in different contexts, non-personal data can become personal. Also, the terms are vague even in the GDPR, which at some point speaks about personal data and at another point, personal information, which are different concepts. The important question that remains open is whether information can be protected by the law. (Janeček, 2018.) Next, I will examine the consumer's view about data use and data ethics through empirical research.

5 Research Methodology

The objective of this empirical research is to gain insights on how consumers view the use of personal data and whether it affects their buying behavior. In addition, this research aims at understanding what kind of thoughts consumers have on data ethics. This chapter provides details of the chosen research strategy and data collection methods as well as the chosen methods of analysis.

5.1 Research Strategy

I have chosen a mixed method research, also referred as triangulation, as a research strategy. Mixed method research is often used to gain a fuller picture of the researched phenomena (Silverman, 2013, pp. 65). It combines both quantitative and qualitative methods to provide a better understanding of the research problem. It is appropriate to use mixed methods when using quantitative and qualitative data together will provide a better picture of the issue than either type by itself. The mixed method approach can also be used when you want to get an alternative perspective on the research problem. (Creswell, 2005, pp. 535.) The main point for choosing this methodology is to gain a good understanding from different perspectives of the overall research objective. In addition, I feel that the individual studies I have selected for this research could not answer the research questions comprehensively on their own.

The research uses a parallel research design, where both qualitative and quantitative studies are equally dominant and address different aspects of the research problem (Hesse-Biber, 2010, pp. 70). The process of the research can be seen in Figure 3. The two individual studies in this research are analyzed independently. Both of the studies are weighed equally, and the integration of the data is completed at the end of the research to identify the similarities and differences of the results.

1. The first part of the study uses secondary data from a 2018 survey “The Use of Digital Services”. This research was commissioned by the Finnish Innovation Fund

Sitra and conducted by Kantar TNS Oy. More than 8000 respondents took part in this survey in Finland, the Netherlands, France, and Germany. The survey's aim was to investigate how people are using data services and whether people trust services that collect and use data.

2. The second part of the research aims to explore perceptions about data ethics by analyzing English language content from Twitter, which is a popular social media platform. The study is a qualitative content analysis of tweets using the hashtag #dataethics. The aim of this content analysis is to gain more insight on people's views on the ethical use of data.

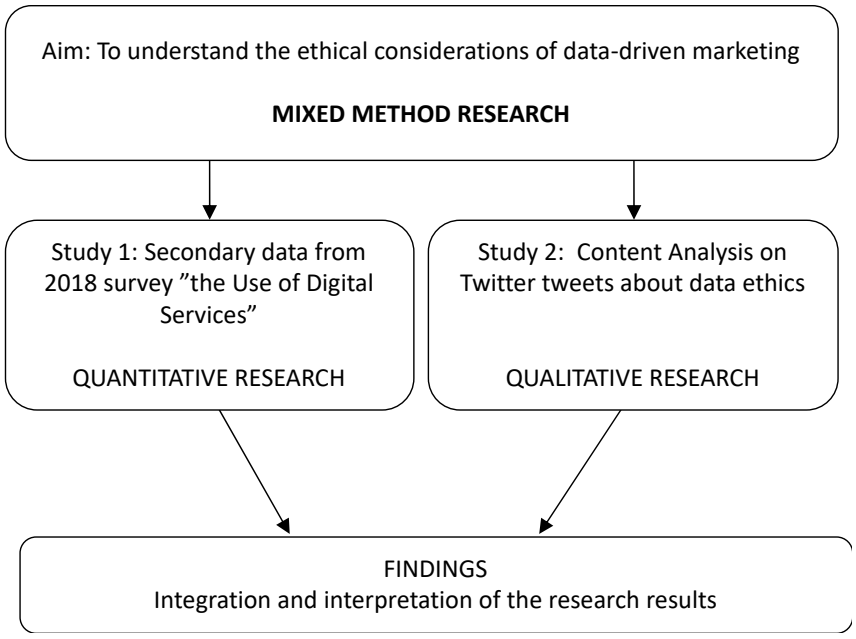


Figure 3. The Process of the Mixed Method Research (Adapted from Hesse-Biber, 2010).

5.2 The Use of Digital Services Survey

The first part of my research is the "Use of Digital Services" survey commissioned by the Finnish Innovation Fund Sitra and conducted by Kantar TNS Oy. The aim of the survey was to examine how digital services are used and if data leakages have affected people's trust in service providers that utilize data. (Hyry, 2019.) I chose to use this secondary data

in my research because it suits well with the aims of my research, and the number of respondents is large enough so that generalizations can be made. The survey was made in 2018, so it is quite recent, and the respondents were from four different European countries, which also suits well for my research.

The main advantage of using secondary data is the low cost of research. Also, it enables the researcher to access large datasets rather easily, saving time and effort. Secondary data is often created by governmental institutions or other professionals, so the quality of data is usually high. (Smith et al., 2011.)

5.2.1 Data Collection

The data collection was conducted by Kantar TNS Oy. The data was collected through online surveys in November-December 2018. The target population of the survey was individuals between the ages of 18 to 65 years in Finland, Germany, the Netherlands, and France. The number of responses gathered from the survey was altogether 8 004, of which 2000 were from Finland, 2004 from Germany, 2000 from the Netherlands, and 2000 from France. The survey consisted of demographic information, closed-ended, and open-ended questions. (Hyry, 2019.) All of the survey questions can be found in Appendix 1. The data that was used for this research is publicly available.

5.2.2 Framework for Data Analysis

The original survey analysis was implemented individually for different countries. In this research, the focus is on analyzing the responses by country and by age. Analysis by country is essential, to see if there are common opinions about data use in different European countries. Analysis by age was incorporated because much talk about data privacy claims that younger people do not care about data privacy as much as the older individuals. This research aims to discover if that is true, and therefore the results are analyzed by age as well. The closed-ended questions were analyzed using Excel, and the open-ended questions were analyzed with the help of Voyant Tools. The data analysis

for this survey focuses on descriptive analysis, i.e. summarizing the data and describing what is occurring in the dataset (Thompson, 2009). The data from the survey is divided into five sections: the background information of respondents, perceptions about terms of use and rights to personal data, trust towards service providers, information management, and fair use of data. The analysis of this survey was made from the publicly available raw data.

5.3 Content Analysis on #dataehics in Twitter

The second part of my research is a content analysis of #dataethics on Twitter. Online conversations can be analyzed in multiple ways. Content analysis was chosen for this research, as it aims to create a holistic view of the content and themes of online conversations. The emphasis of content analysis can be either quantitative or qualitative, and in this research, the focus is on qualitative content analysis, which aims to categorize data into words or themes (Laaksonen et al., 2013, pp. 208; Bengtsson, 2016). Still, Schreier (2012, pp. 36) points out that qualitative research often includes quantitative elements, e.g. by presenting the frequencies of the themes. Content analysis enables the analysis of social interaction, as it looks directly at communication and allows to study real, unaltered opinions of people. It is also an unobtrusive method of research, as the researcher does not influence the communication or opinions of the target audience. (Hakala & Vesa, 2013, pp. 223; Mazur, 2010, pp. 78.) Online conversations apply well in studying societal phenomena as they reflect the opinions of citizens quite accurately (Laaksonen & Matikainen, 2013, pp. 208). In addition, the populations in online conversations are large and diverse (Mazur, 2010, pp. 78).

Currently, Twitter is one of the most popular social networking sites in the world. It has over 330 million monthly active users worldwide, and the number of tweets sent every day is 500 million (Omnicore, 2020). Twitter is a popular social media platform for discussing current issues. People express their opinions freely, and most of the tweets are publicly available. I chose content analysis as part of my research to examine what kind of themes are discussed on Twitter about data ethics. The aim of this content analysis is

to understand the opinions of individuals regarding data ethics and to examine if there are any specific recurring themes that arise about the topic. The hashtag #dataethics was chosen because much of the current ethical challenges in marketing are related to data, and the emphasis of this research was on data ethics. In addition, the researched hashtag interlaces well with the secondary analysis data.

The data collection and analysis in this research followed the social media analytics framework by Stieglitz et al. (2018). The framework shown in Figure 4 was adapted for the purposes of this specific research. The aim of the content analysis was to find out recurring themes regarding data ethics, so the approach that was chosen was topic-related.

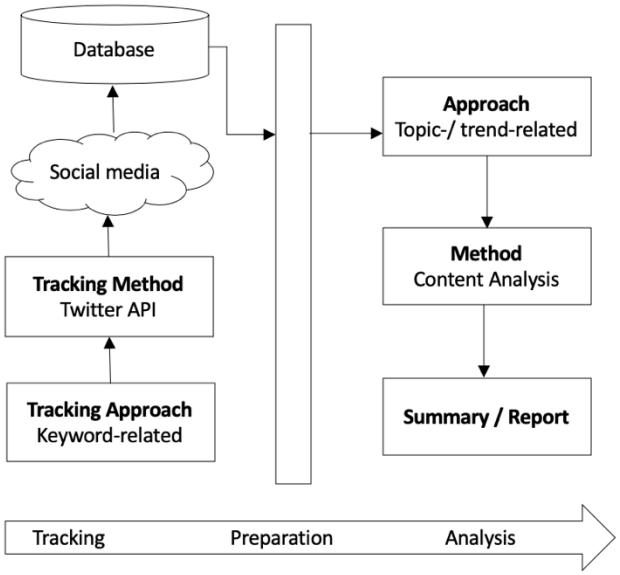


Figure 4. Data Collection and Analysis (Adapted from Stieglitz et al., 2018).

5.3.1 Data Collection and Preprocessing

For the content analysis, public tweets using the hashtag #dataethics were collected. The data collection period was four months from the 25th of September 2019 to the 24th of January 2020. The total number of tweets collected during that period was 5,692, of which 5,007 were unique tweets. The data was collected with TAGS, which is a free

Google Sheet template that allows automated collection of search results from Twitter. TAGS collects tweets from the Twitter API (Application Programming Interface). In qualitative research, the goal is to understand a phenomenon instead of creating generalizable results. Therefore, this study uses purposeful sampling, which means that the researcher intentionally selects certain individuals or sites to gain a comprehensive understanding of the phenomenon. (Creswell, 2005, pp. 206.)

Social media data needs to be cleaned before it can be analyzed. This is a very important step in data analysis in order to make the data more understandable and to make the analysis more valid. Uncleaned data may give unreliable results. (Batrinca & Treleaven, 2015.) Cleaning the data from non-English tweets is important in order to get results that can be analyzed properly, so this was the first step in the cleaning process. The next step was to delete retweets and duplicate tweets from the data. Some tweets may be tweeted several times or they can be created by bots, so it is important to delete these. After this phase, those tweets were deleted, which contained only a link and no other text. Finally, Twitter allows users to specify their location in their profiles. As this research concentrates on ethics from a European perspective, only those tweeters were selected for the study, which stated their location to be in Europe. Those tweets were deleted, which were coming from outside of Europe, or the location was not specified.

As a conclusion, the inclusion criteria for the tweets were: 1) publicly available tweets, 2) use of the hashtag #dataethics, 3) posted between 25th of September 2019 to 24th of January 2020, 4) text in English, 5) Location in Europe. The exclusion criteria for the tweets were: 1) non-English tweets, 2) re-tweets, 3) duplicate tweets, 4) tweets containing only a link, 5) location other than Europe or unspecified.

The total number of tweets after the cleaning phase was 579, and these tweets were included in the content analysis stage. Therefore, the final sample size for the analysis is 11 % of the original 5,007 unique tweets. Figure 5 shows the stages of the cleaning phase.

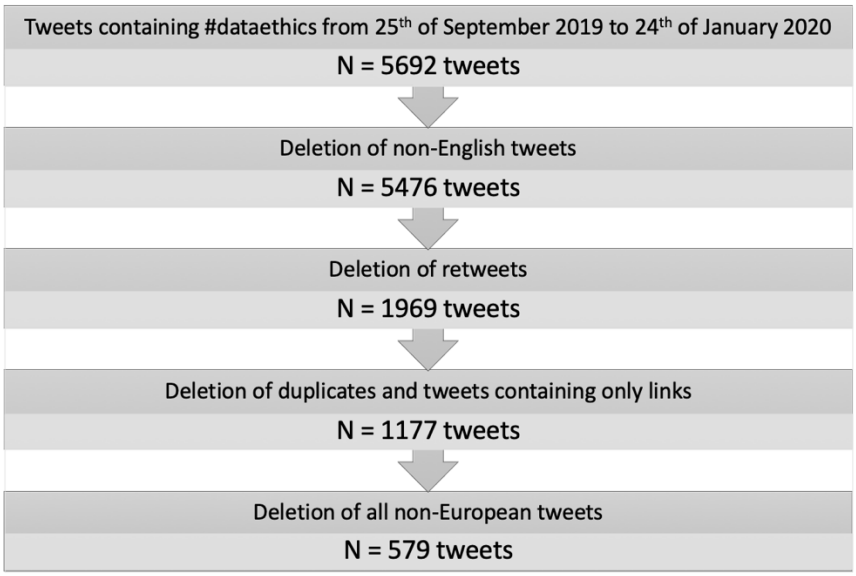


Figure 5. Selection of Tweets.

5.3.2 Framework for Data Analysis

The base for data analysis in this research lies in inductive reasoning, which means that the data is analyzed with an open mind without any predetermined subjects. Inductive reasoning therefore aims at identifying subjects based on the text. (Bengtsson, 2016). Figure 6 shows the data analysis process, which has been adapted from the data analysis framework from Bengtsson (2016). This research uses latent analysis, i.e. the focus is on interpreting the text to find the underlying meaning of it. In latent analysis, the researcher needs to look at the context for interpretation. (Bengtsson, 2016; Schreier, 2012, pp. 15.) Tweets often include links to external sources like news articles or blog posts. Often the links explain the tweet's theme more, and therefore links need to be taken into consideration when assigning tweets into categories. In addition, some of the analyzed tweets were replies to other tweets, and therefore the context needs to be taken into consideration in the analysis.

The analysis process consists of four different phases: decontextualization, recontextualization, categorization, and compilation. The first phase in the process is decontextualization, which means that the researcher needs to become familiar with the data to be

analyzed in order to understand the whole dataset. After that, the data can be broken into smaller meaning units, which are further labeled with a code. The codes are used to identify concepts to find patterns in the data. In this research, codes are created inductively, i.e. the coding is an iterative process, and the codes may change during the research process. The second phase of recontextualization means that the researcher has to check that the original text matches the code list made in the previous phase and whether there is unimportant data that can be excluded from the research. The categorization phase includes identifying themes and categories. Categories can also be divided into subcategories. Often there are several categories in the beginning of the process, but the number of categories usually reduces during the analysis process. The final phase of compilation includes the analysis of the found categories or themes. In addition, the researcher needs to examine how the findings correspond to the literature. (Bengtsson, 2016.) The data analysis stage was implemented manually using Microsoft Excel. In addition, Voyant tools was used to examine the occurrence of the most popular words in the dataset.

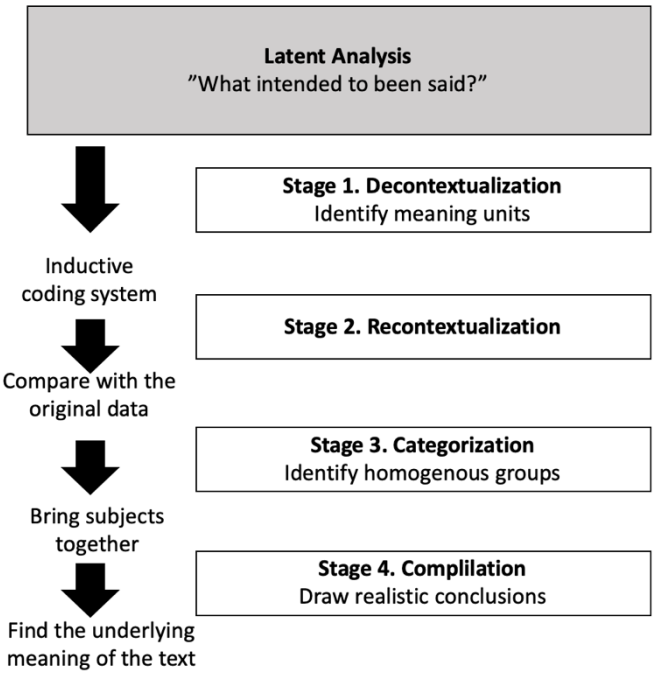


Figure 6. Data Analysis Framework (Adapted from Bengtsson, 2016).

6 Research Results and Findings

In this chapter, I will analyze the results of the empirical research. The objective was to gain insights on the consumers' view on the use of personal data and trust towards service providers. The research consisted of two studies; a secondary data survey and a content analysis. I will look at the results of both researches individually, and in the Findings subsection I will look at the similarities and differences of the results.

6.1 The Use of Digital Services Survey

We will first look at the demographic information of the respondents. The demographic information included Age, Gender, Region Type, Level of Education, and Occupational group or situation. Figure 7 shows the percentage of Age groups in different countries. The largest group to answer the survey was aged between 55 to 65. The youngest group aged between 18 to 25 was the smallest in size. There are some variations in the age groups between countries. For example, in German respondents, the number of respondents under the ages of 25 was significantly lower than in other countries.

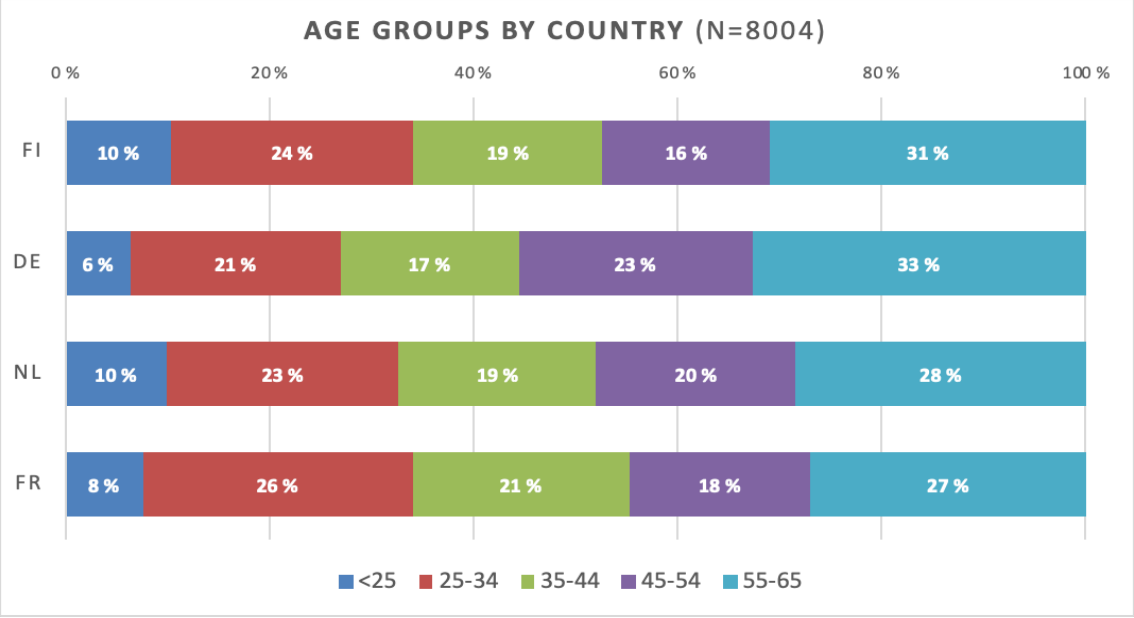
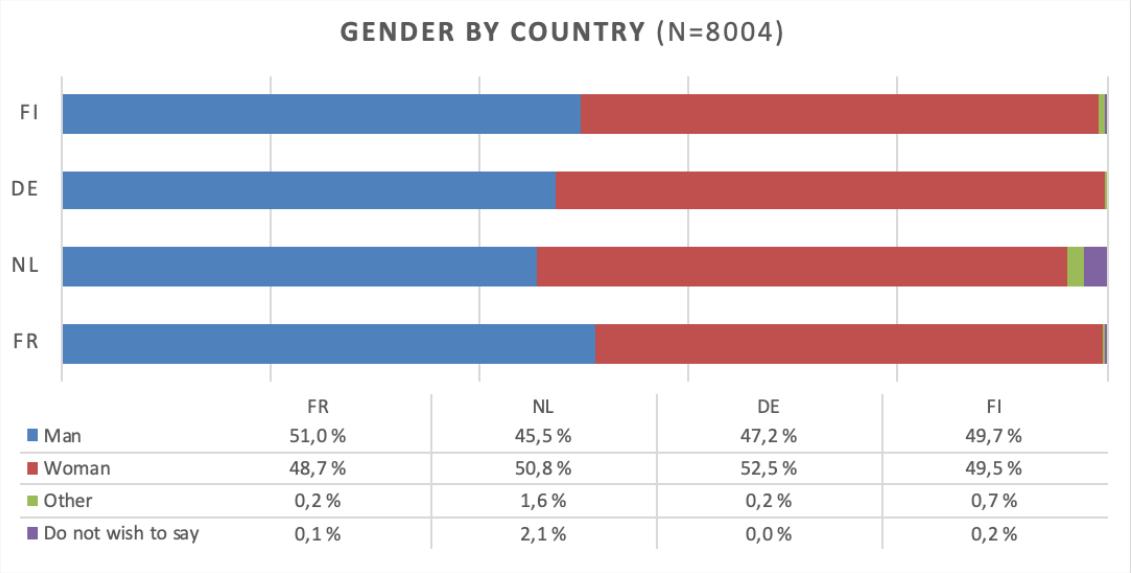


Figure 7. Age Groups by Country.

Of the total population, 48% were men and 50 % women. 2 % of the respondents stated "other" as their gender or did not wish to say. Figure 8 shows the percentage of Gender variable by each country. In all countries, the gender composition was almost equal between men and women.

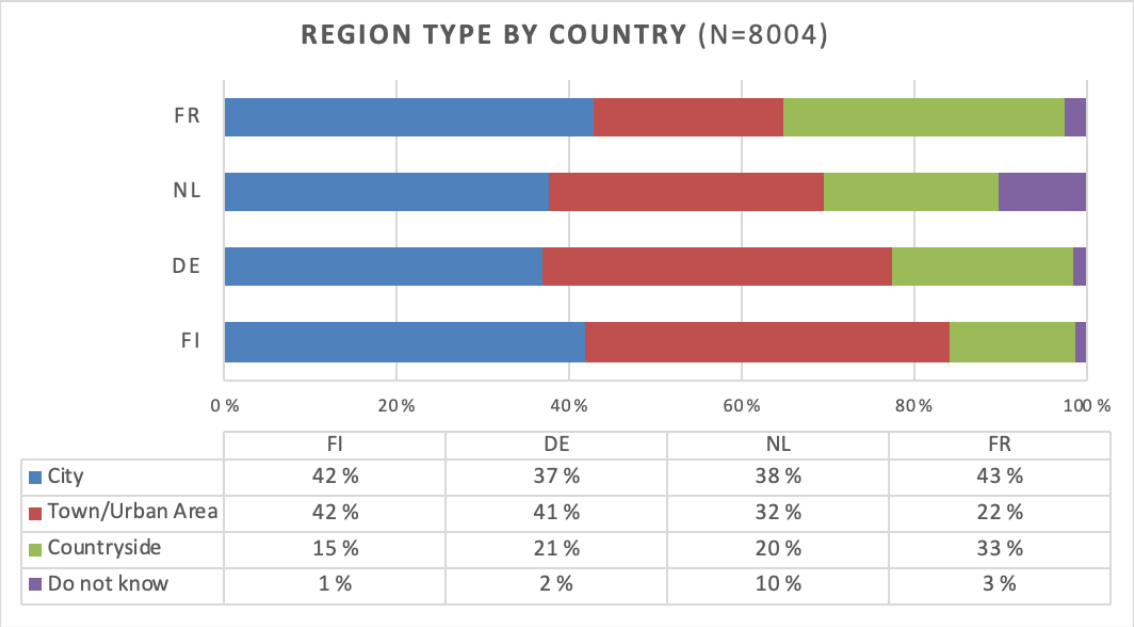


	FR	NL	DE	FI
Man	51,0 %	45,5 %	47,2 %	49,7 %
Woman	48,7 %	50,8 %	52,5 %	49,5 %
Other	0,2 %	1,6 %	0,2 %	0,7 %
Do not wish to say	0,1 %	2,1 %	0,0 %	0,2 %

Figure 8. Gender by Country.

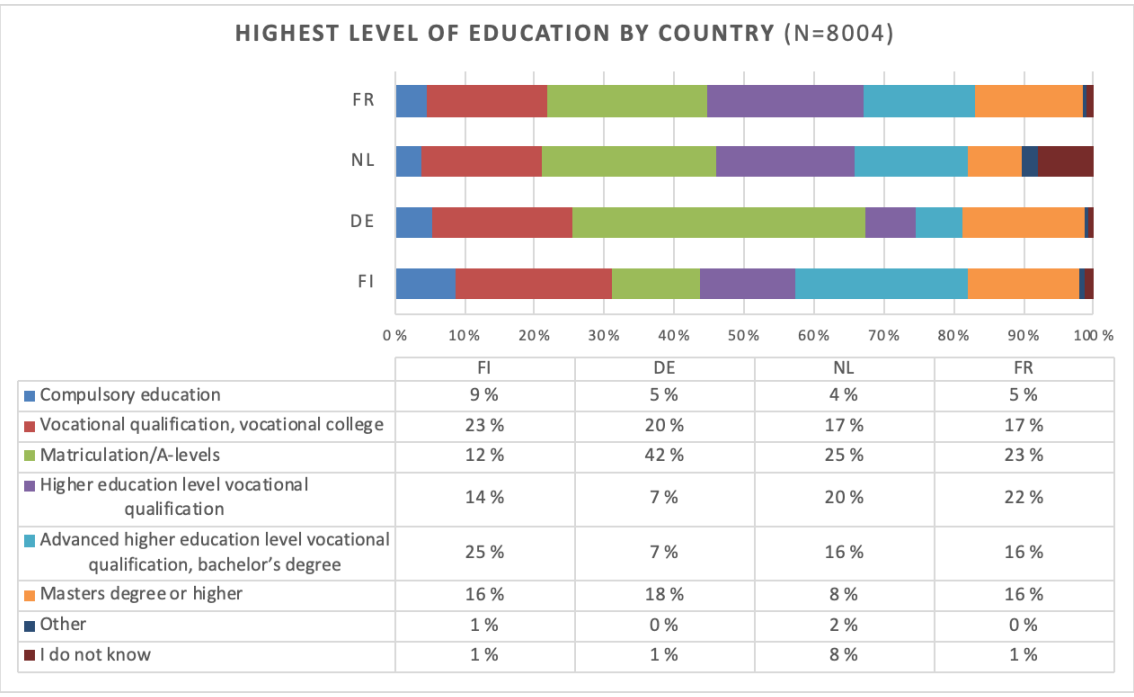
The respondents were also asked what type of region they live in. Figure 9 shows the percentages of region types by country. We can see from the figure that most of the respondents were from cities or urban areas. However, in France, the number of respondents from the countryside was higher than in other countries whereas Finland had the smallest number of respondents living in the countryside.

Figure 10 shows the highest level of completed education. There is some variation in this variable between different countries. Finland has the highest number of respondents with higher education qualification, with 41 % of respondents having a bachelor's or master's degree. In Germany, 42% of respondents have stated Matriculation or A-levels as their highest level of completed education.



	FI	DE	NL	FR
City	42 %	37 %	38 %	43 %
Town/Urban Area	42 %	41 %	32 %	22 %
Countryside	15 %	21 %	20 %	33 %
Do not know	1 %	2 %	10 %	3 %

Figure 9. Region Type by Country.

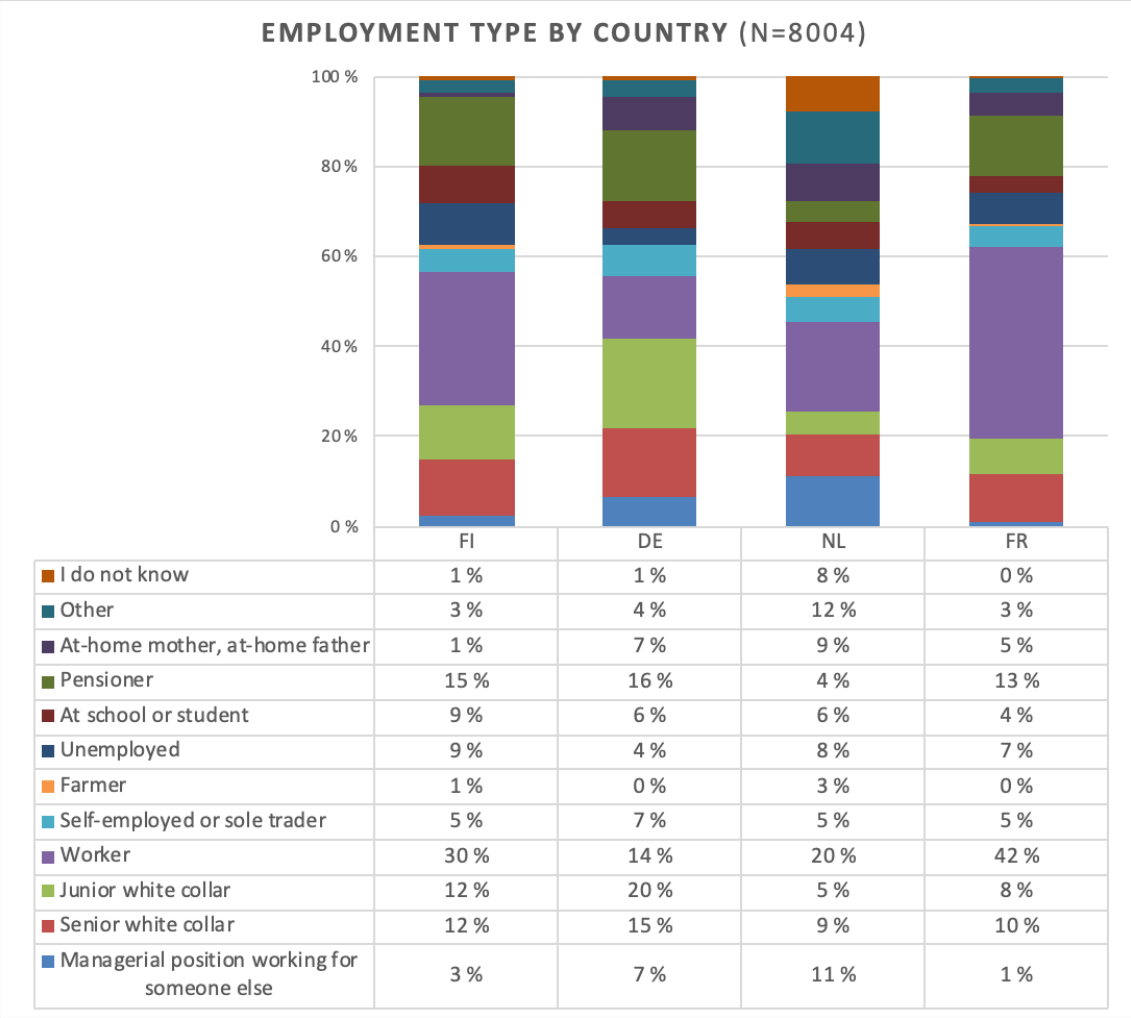


	FI	DE	NL	FR
Compulsory education	9 %	5 %	4 %	5 %
Vocational qualification, vocational college	23 %	20 %	17 %	17 %
Matriculation/A-levels	12 %	42 %	25 %	23 %
Higher education level vocational qualification	14 %	7 %	20 %	22 %
Advanced higher education level vocational qualification, bachelor's degree	25 %	7 %	16 %	16 %
Masters degree or higher	16 %	18 %	8 %	16 %
Other	1 %	0 %	2 %	0 %
I do not know	1 %	1 %	8 %	1 %

Figure 10. Highest Level of Education by Country.

Figure 11 shows the employment type of respondents by country. There is much variation in this variable. It has to be noted that in France, 42 % of respondents have stated

“worker” as their employment type. The biggest number of respondents are workers also in Finland and the Netherlands, but in Germany the largest group is junior white collars.



	FI	DE	NL	FR
I do not know	1 %	1 %	8 %	0 %
Other	3 %	4 %	12 %	3 %
At-home mother, at-home father	1 %	7 %	9 %	5 %
Pensioner	15 %	16 %	4 %	13 %
At school or student	9 %	6 %	6 %	4 %
Unemployed	9 %	4 %	8 %	7 %
Farmer	1 %	0 %	3 %	0 %
Self-employed or sole trader	5 %	7 %	5 %	5 %
Worker	30 %	14 %	20 %	42 %
Junior white collar	12 %	20 %	5 %	8 %
Senior white collar	12 %	15 %	9 %	10 %
Managerial position working for someone else	3 %	7 %	11 %	1 %

Figure 11. Employment Type by Country.

6.1.1 Rights to Personal Data and Perceptions About Terms of Use

The first part of the survey was to find out people’s perceptions about their rights to personal data, as well as how they perceive the terms and conditions of applications or services. Figure 12 shows what kind of rights people believe to have concerning personal data that has been gathered by service providers. The responses of the whole population as well as individual countries are shown. This question was a multiple-choice question,

so the respondents could pick more than one option from the list. Over half of all the respondents believed they have the right to access or delete any personal data collected of them. However, in the Netherlands, these beliefs were significantly lower than the average. Also, nearly 60 % of all the respondents believed they have the right to know how and for what their data is being used. 35 % of all the respondents believed that they have the right to get notified when their data is being sold or handed over to a third party. This amount is quite high, regarding that the statement is false. There was less belief on the right to move data from one system to another, with only 21 % of respondents believing so. 6 % of respondents believed they do not have any rights regarding personal data. This belief was higher in Finland (9%), whereas in the Netherlands, only 3 % believe they have no rights. 15 % of all the respondents did not know about their rights.

What rights do you believe you have concerning the personal data the service provider has gathered of you?

STATEMENT	Finland	Germany	The Netherlands	France	TOTAL
Right to have access to any personal data the service provider has collected of me	60 %	61 %	57 %	54 %	**58 %**
Right to have any personal data collected of me corrected or updated	51 %	54 %	**39 %**	51 %	**49 %**
Right to have any personal data collected of me deleted	51 %	62 %	**49 %**	56 %	**55 %**
Right to move any personal data from one system to an other	19 %	22 %	24 %	20 %	**21 %**
Right to know how and for what purpose my personal data is used	**65 %**	64 %	56 %	**49 %**	**59 %**
Right to get a notification from the service provider when my personal data is being sold/handed over to a third party	30 %	**43 %**	37 %	31 %	**35 %**
Right to restrict handling of my personal data, especially where automated decision making is conserned	33 %	43 %	33 %	37 %	**36 %**
Right to be informed when there is a service provider data breach involving my data	38 %	47 %	43 %	41 %	**42 %**
I do not believe I have any rights	**9 %**	6 %	**3 %**	5 %	**6 %**
I do not know	12 %	13 %	18 %	18 %	**15 %**

Figure 12. Perceptions About Rights to Personal Data.

Figure 13 shows the attitudes towards terms and conditions of applications or services. Most of the respondents in different countries read the terms before accepting them. However, in France, over 30 % of respondents accept them without reading them first. Less than 15 % of all the respondents always read the terms.

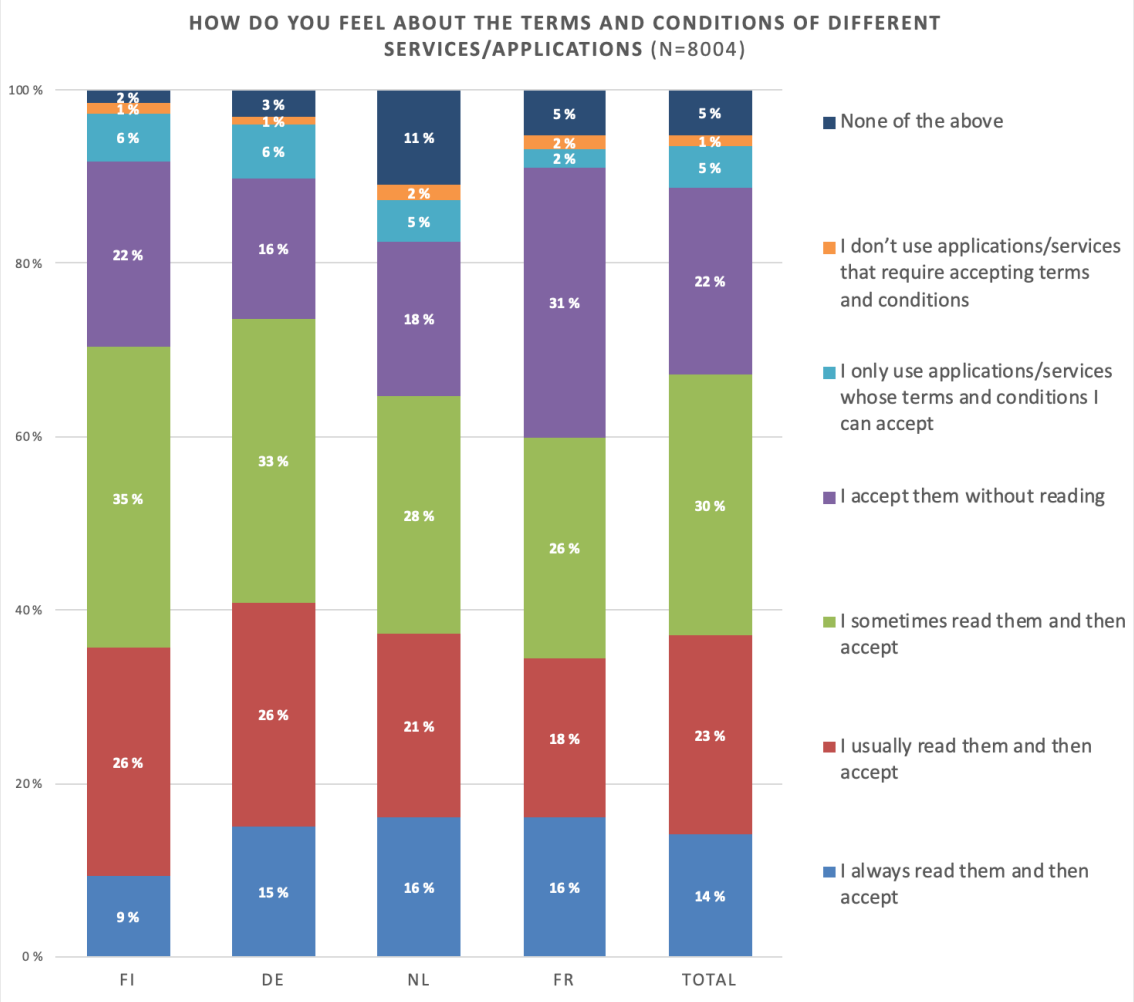


Figure 13. Attitude Towards the Terms and Conditions of Services or Applications.

The respondents were also asked how well they understand the terms and conditions. Figure 14 shows that there is variation in the understanding between different countries. Of all the respondents, 33 % understand the terms and conditions very well or fairly well. The most common response is still "somewhat well", and nearly half of German respondents have chosen this response. 25 % of all the respondents state that they do not understand the terms and conditions very well or not well at all.

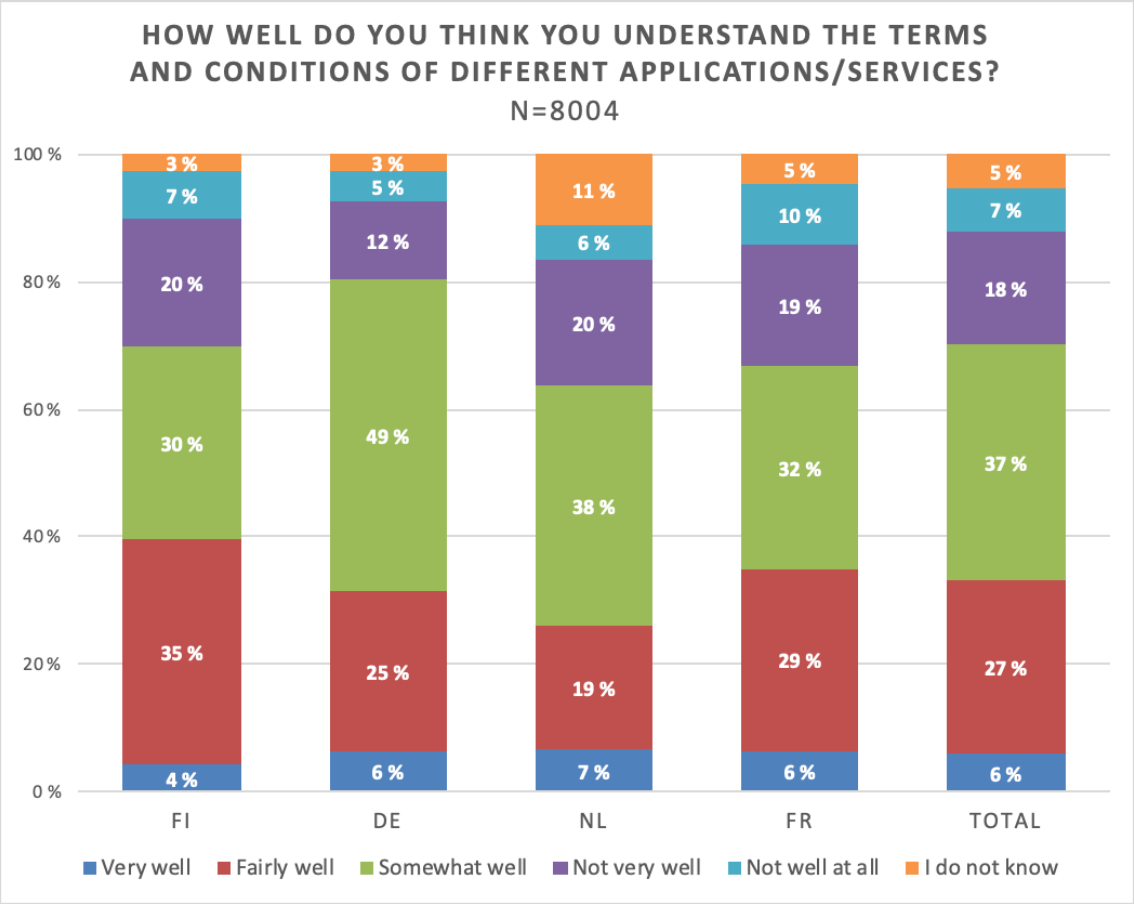


Figure 14. Level of Understanding About Terms and Conditions by Country.

If we look at the same question from age point-of-view, Figure 15 shows that there isn't much variation between different ages in how well individuals know the terms and conditions. 35% of the respondents under the age of 35 know the terms or conditions very well or fairly well. The percentage of these responses is only slightly smaller in older age groups. Still, only 4% of the respondents aged 55 to 65 state that they know the terms and conditions very well. In addition, the youngest group of respondents under the age of 25 have the highest number of respondents saying that they do not know the terms and conditions very well or not well at all. 27% of the respondents in the youngest group feel this way.

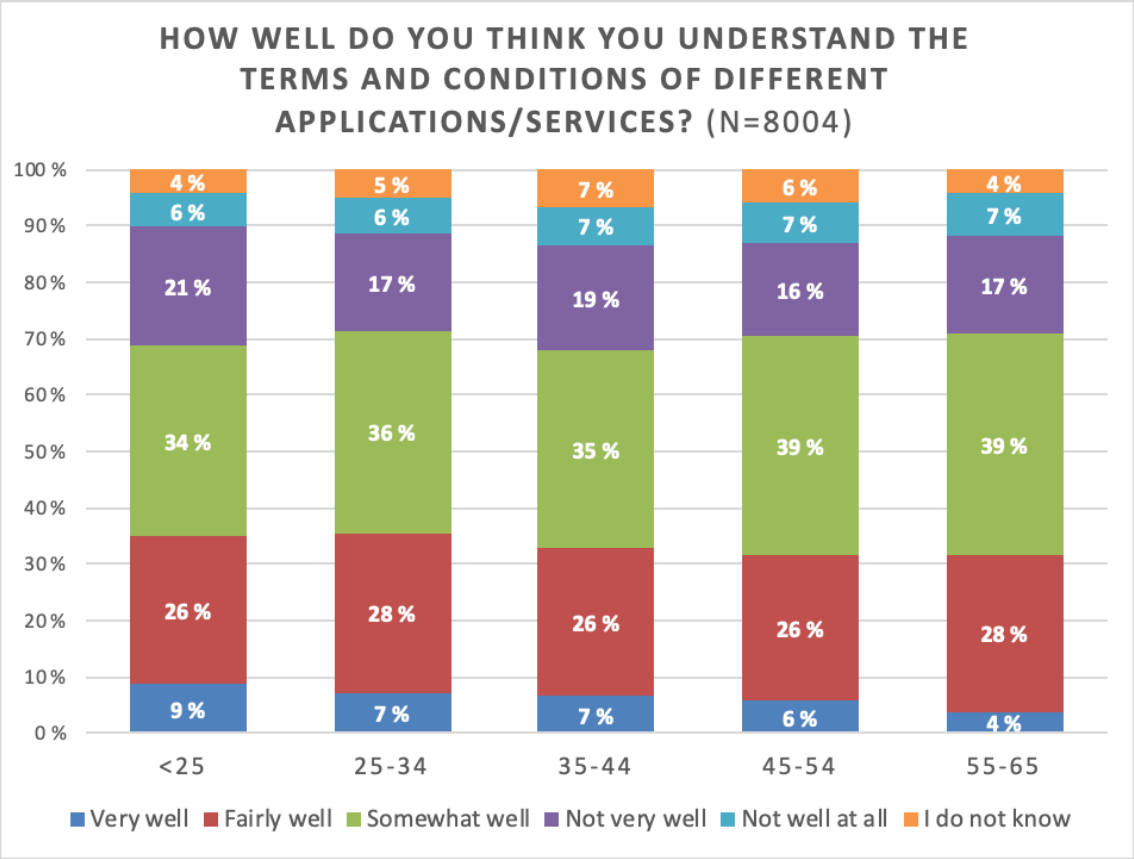


Figure 15. Level of Understanding About Terms and Conditions by Age.

The respondents were asked if they had changed the settings for a service or application to meet their own needs, e.g. for privacy. Figure 16 shows that 61 % of all the users had changed all, most or some of the settings. Still, there are 25 % of those who have not changed any settings. In Finland, the changing of settings was more common than in other countries, while in France, almost one-third of respondents had not changed any settings.

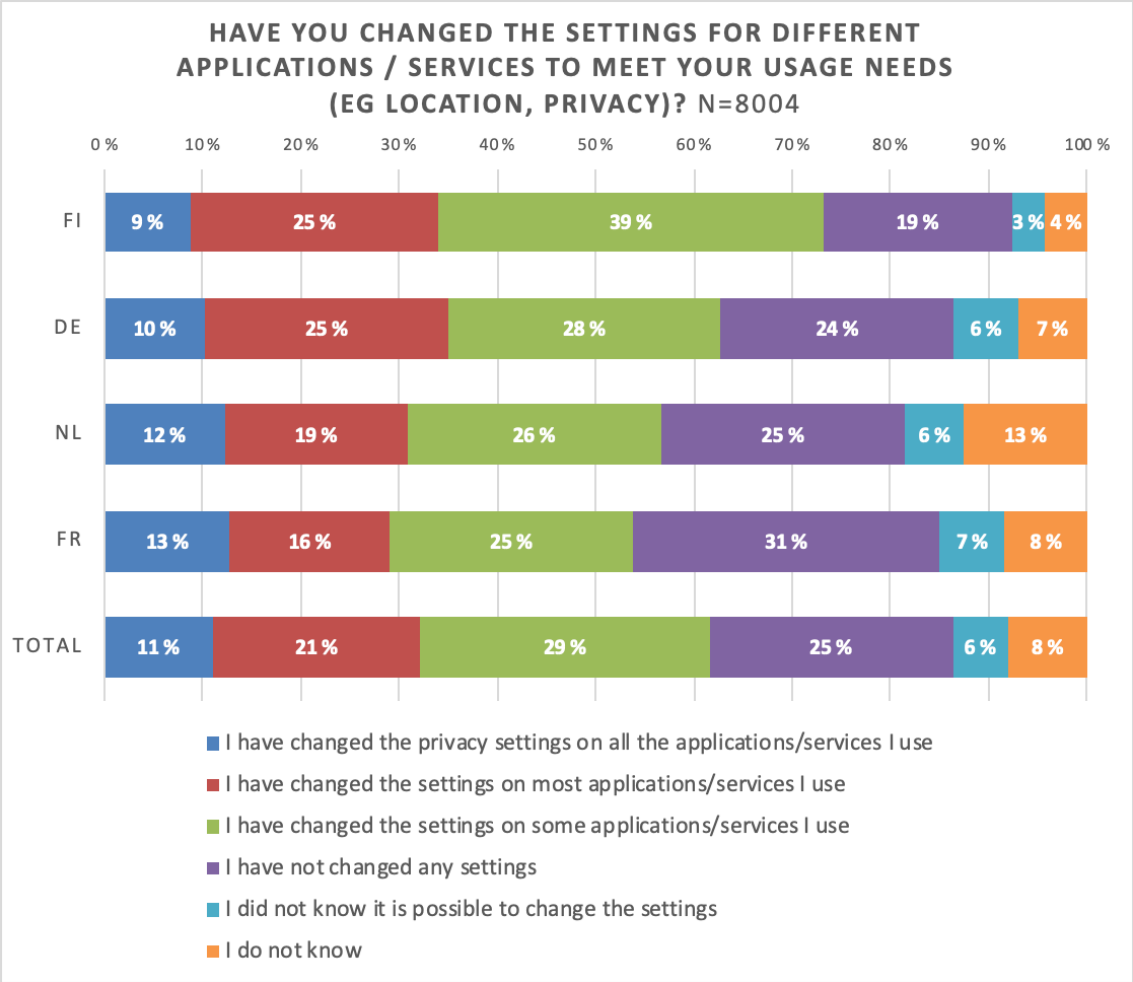


Figure 16. Changing the Settings for Applications or Services by Country.

If we look at the same question form age point-of-view, we can see from Figure 17 that the older people get, the less they have changed settings on applications or services. Almost half of respondents under the age of 25 has changed settings on all or most the services or applications they use, but only 22% of respondents aged between 55 to 65 have done the same. 32 % of the oldest age group had not changed any settings. In addition, the number of respondents stating that they do not know how to change settings increased slightly with age.

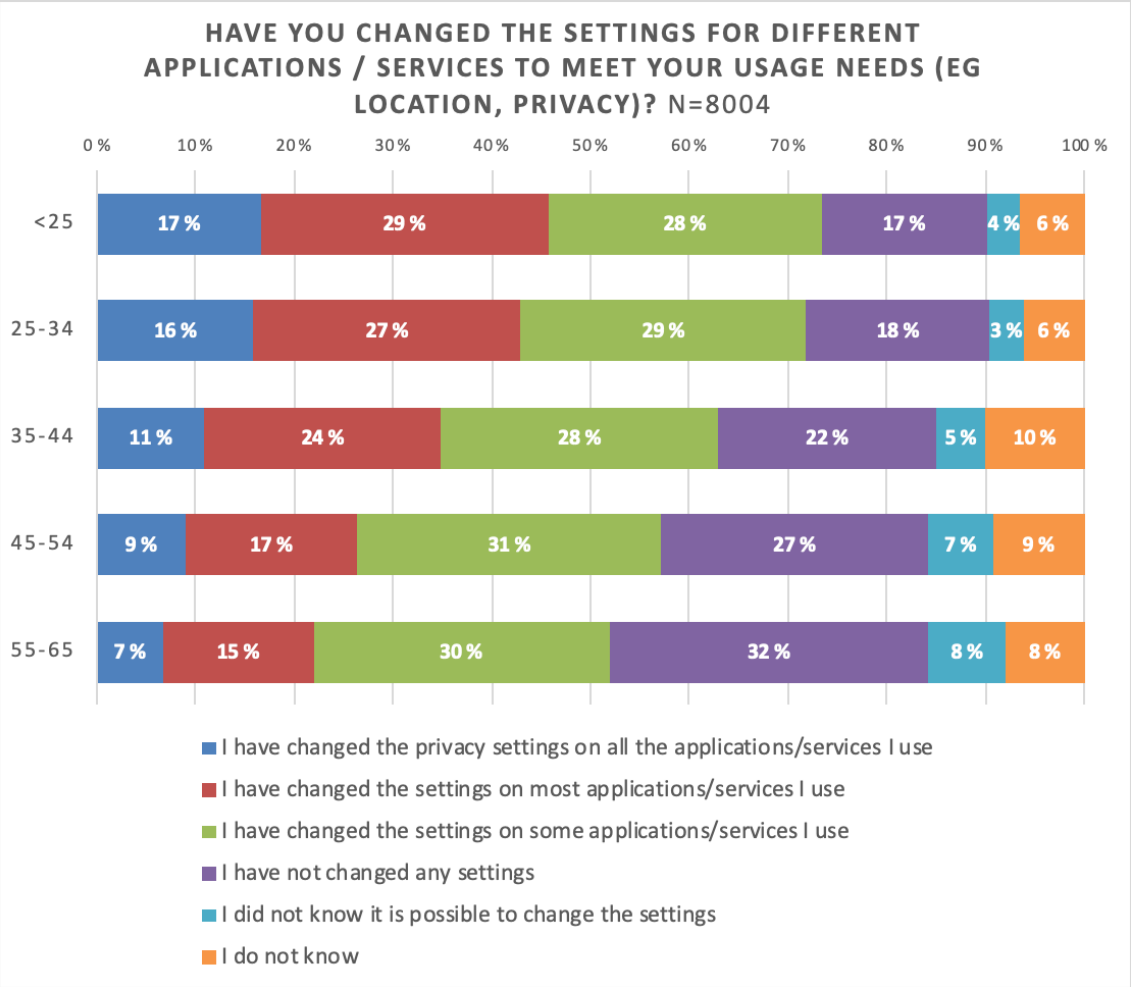


Figure 17. Changing the Settings for Applications or Services by Age.

Those who had not changed settings or were not aware that they could be changed, were asked what the reasons were for not changing settings. Figure 18 shows that the most important reason for not changing settings was that the respondents felt that it does not have any effect. Of all the respondents 25 % felt this way, and in Germany and the Netherlands this option was selected almost by 30 % of respondents. The exception here is Finland, where 28 % of respondents feel that it is not important to change any settings. The second most important reason for not changing settings was that respondents are unaware of how the settings can be changed.

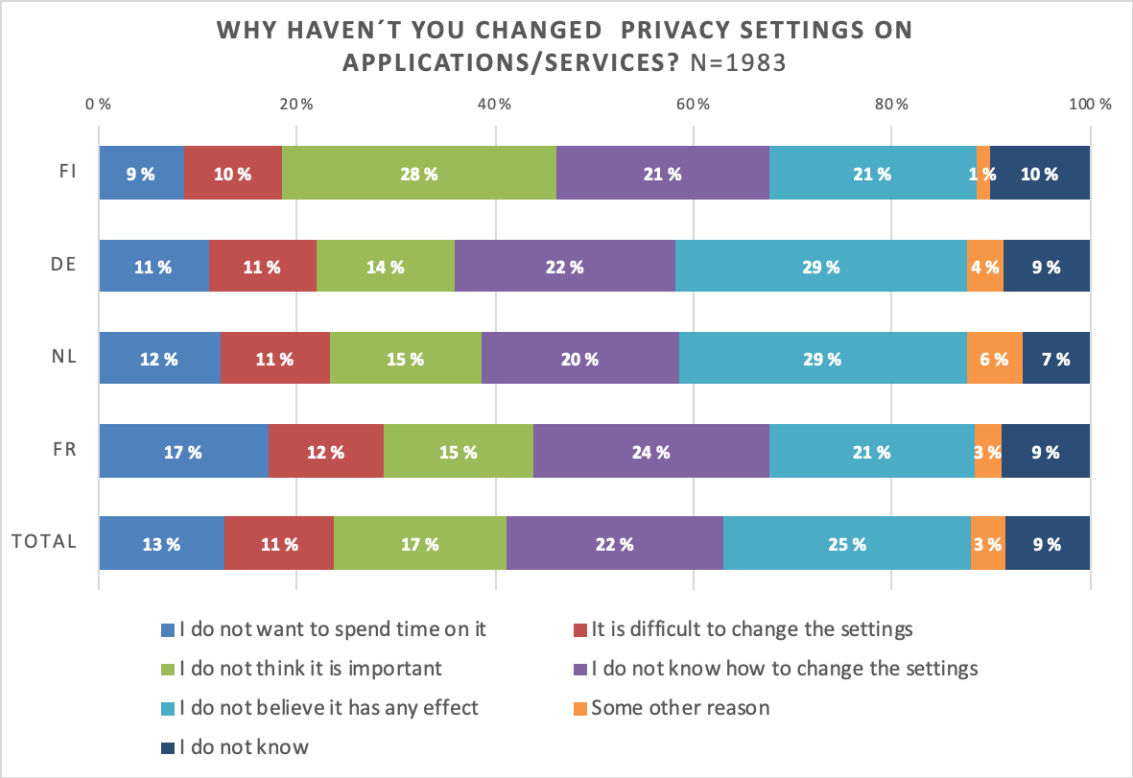


Figure 18. Reasons for Not Changing Settings by Country.

Figure 19 shows the answers to this question by age. The responses vary in different age groups. For the youngest respondents under the age of 25, the most popular option (25%) was that it is not important to change the settings. For the other age groups, the most popular option was that changing the settings have no effect. In addition, 18 to 25 % of the respondents in different age groups didn't know how the settings could be changed. Also, especially the younger age groups did not want to spend time on changing the settings.

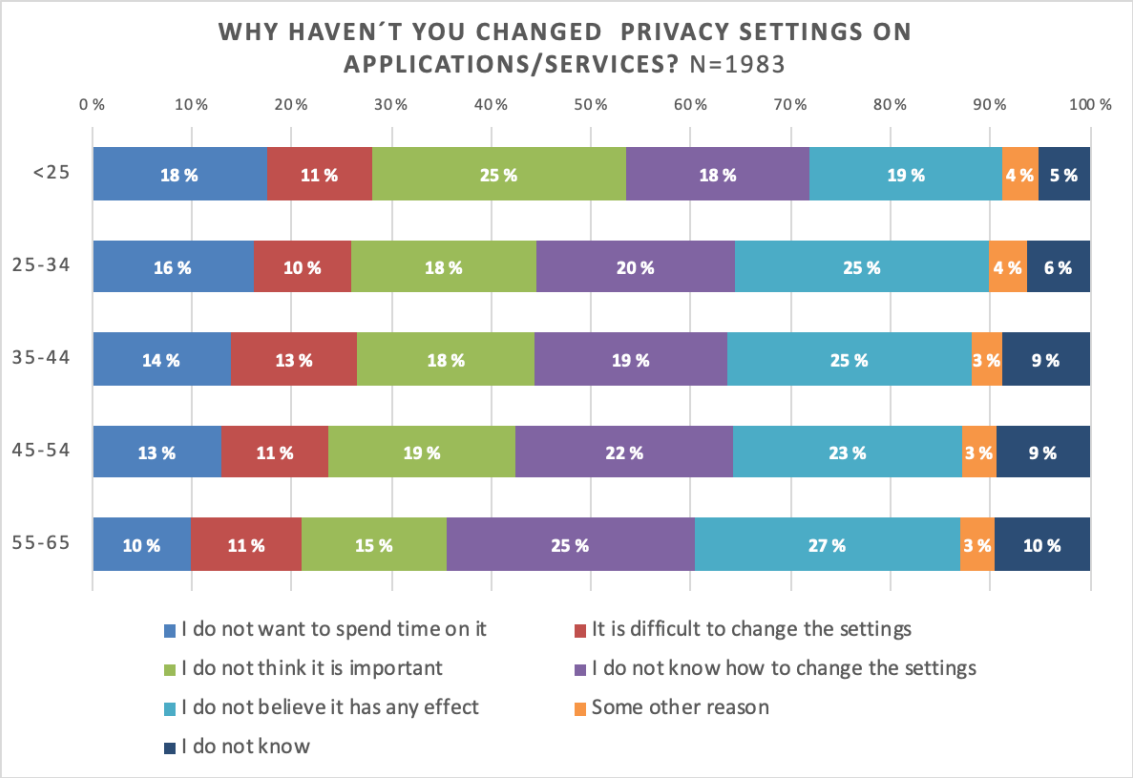


Figure 19. Reasons for Not Changing Settings by Age.

6.1.2 Trust Towards Service Providers

The second part of the survey examined trust towards service providers. The first question in this category was how the recent news of data leakages had affected the respondents' use of different services. The respondents could choose more than one option. Figure 20 shows that the responses vary in different countries. The most popular option in Finland (36%), France (34 %), and Germany (29%) was that data leakages have not had any effect on the use of digital services. In the Netherlands, the most popular option was adjusting settings in some services, with 28 % of the respondents stating so. Still, of all the respondents, 38 % had either stopped or reduced using some services because of the news of data leakages. Figure 21 shows the responses by age. We can see that majority of respondents under the age of 35 have either stopped or reduced using some services or adjusted their privacy settings. The majority of the older age groups felt that the news of data leakages haven't had any effect.

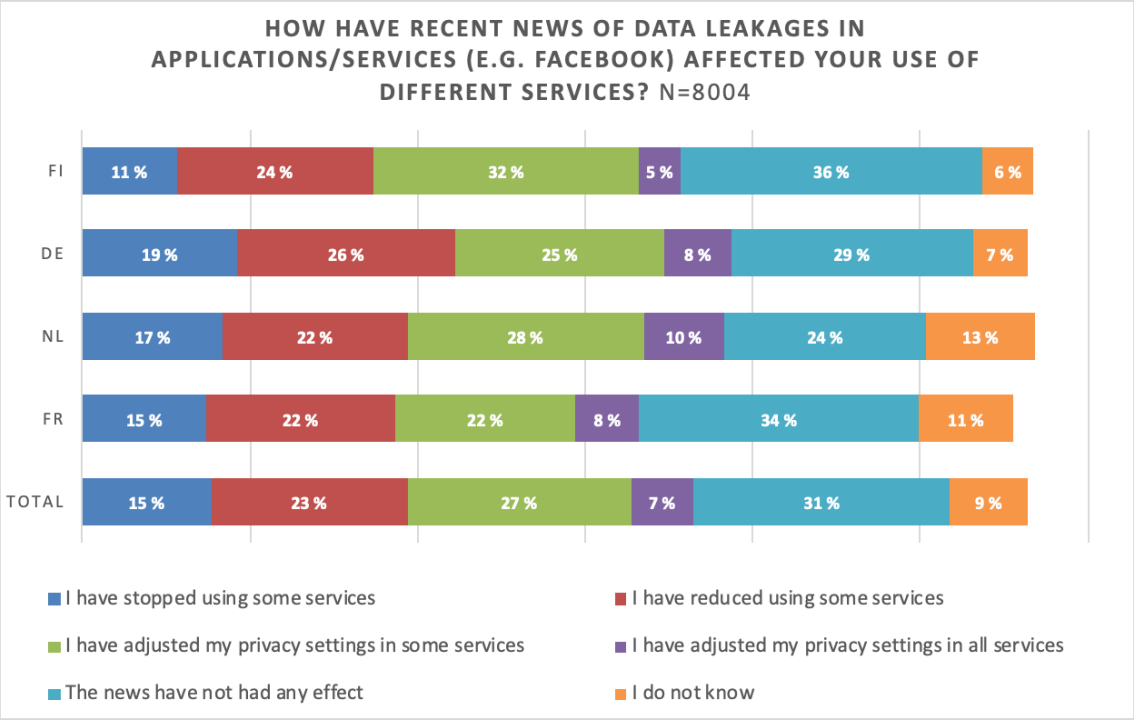


Figure 20. Effect of Data Leakages to Use of Different Services by Country.

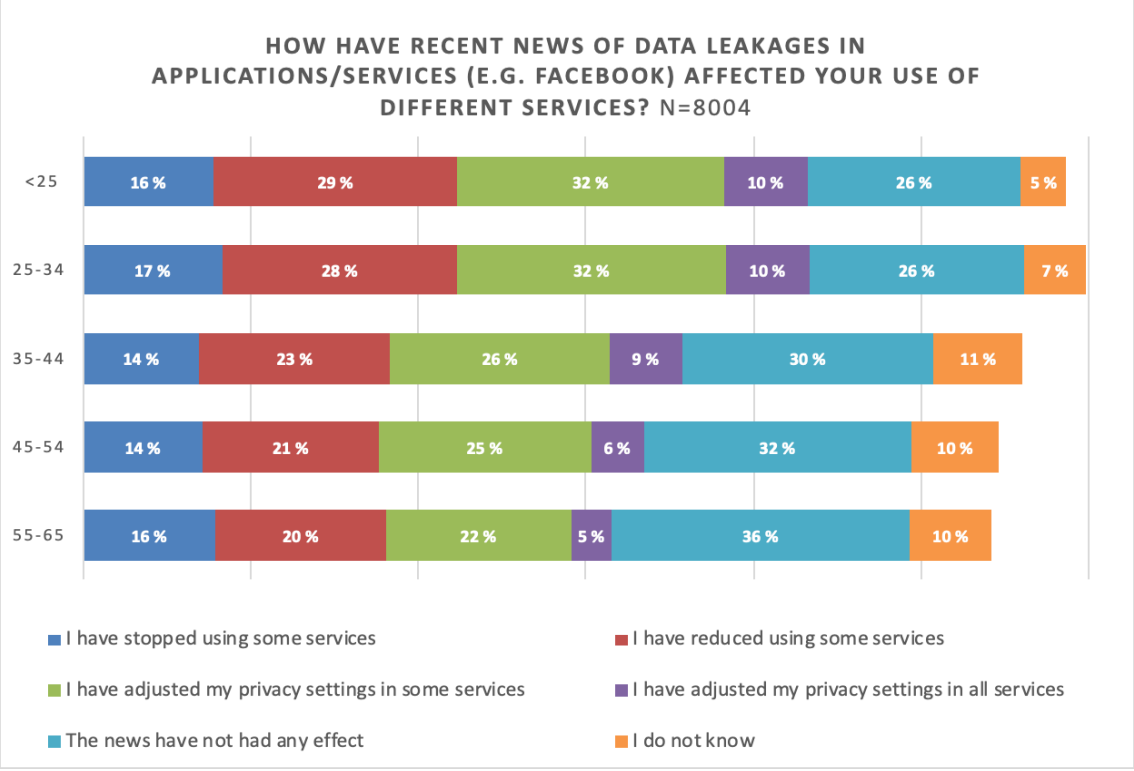


Figure 21. Effect of Data Leakages to Use of Different Services by Age.

The respondents were asked to evaluate if the lack of trust towards service providers prevents them from using digital services. Figure 22 shows that most of the respondents in all the countries either agree or strongly agree that the lack of trust affects the use of digital services. Still, nearly a third of the respondents in Germany, the Netherlands, and France neither agree nor disagree with the statement. In Finland, 20 % of respondents disagree with this statement. If we look at this question from age point-of-view, we can see from Figure 23 that there is not much variation among different age groups. 40% or over of the respondents in all the age groups either agree or strongly agree that the lack of trust towards service providers prevents them from using digital services.

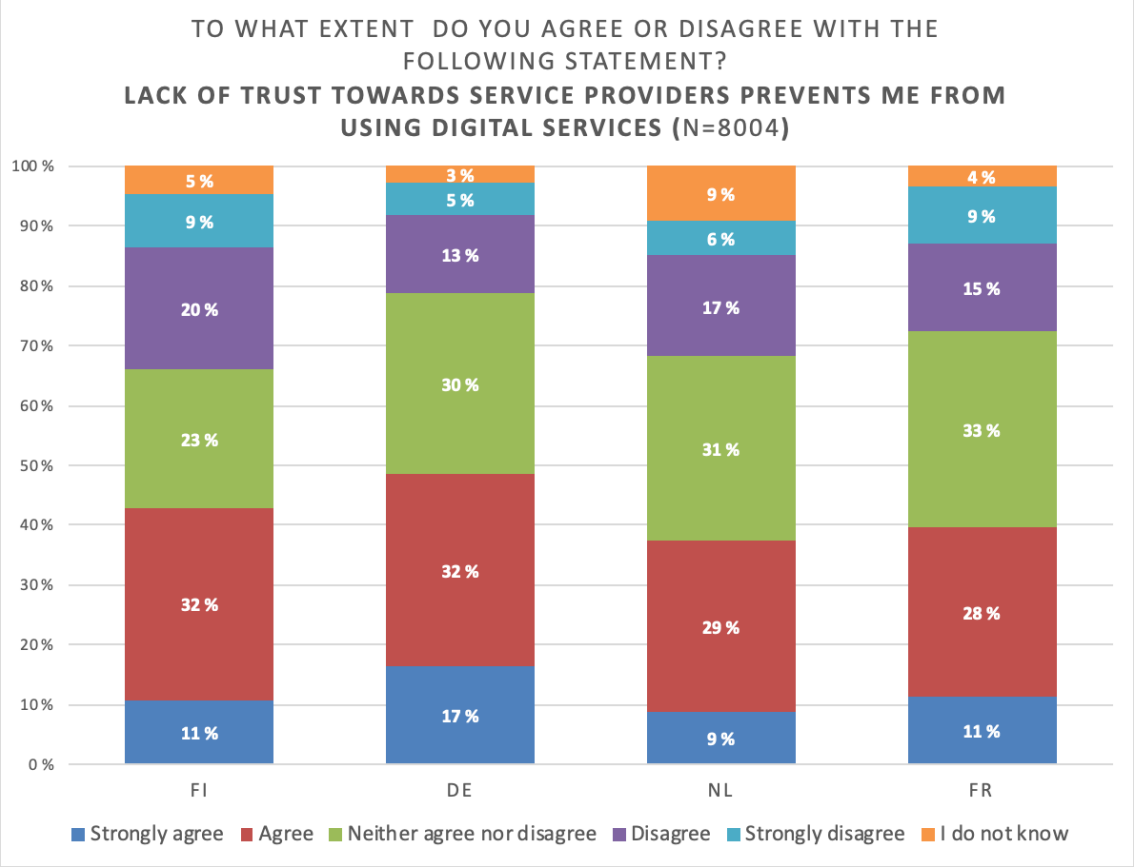


Figure 22. Lack of Trust and the Use of Services by Country.

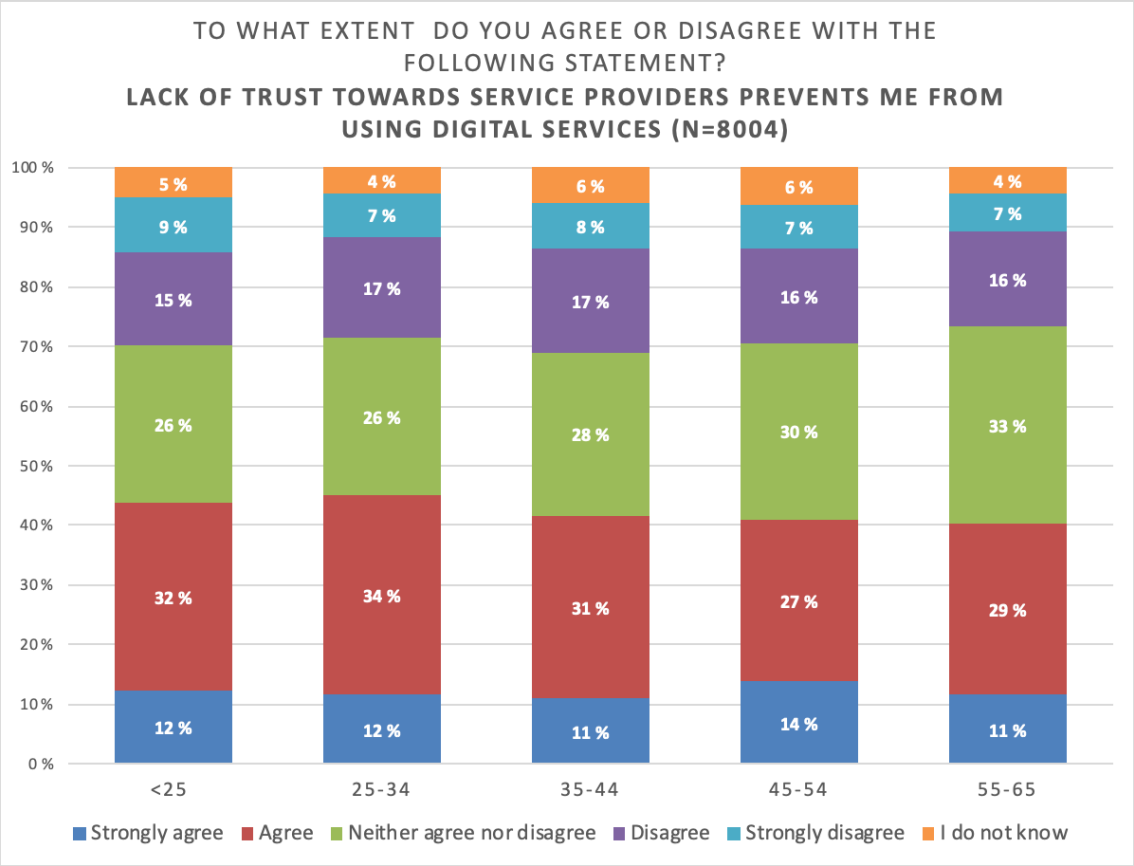
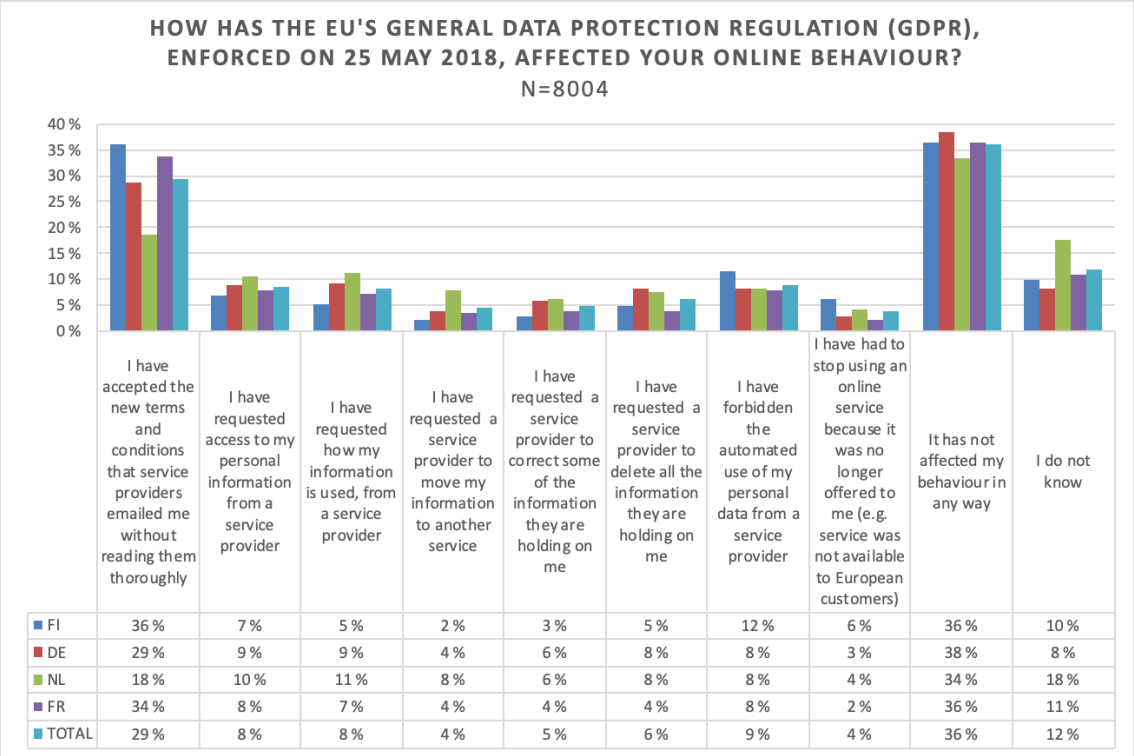


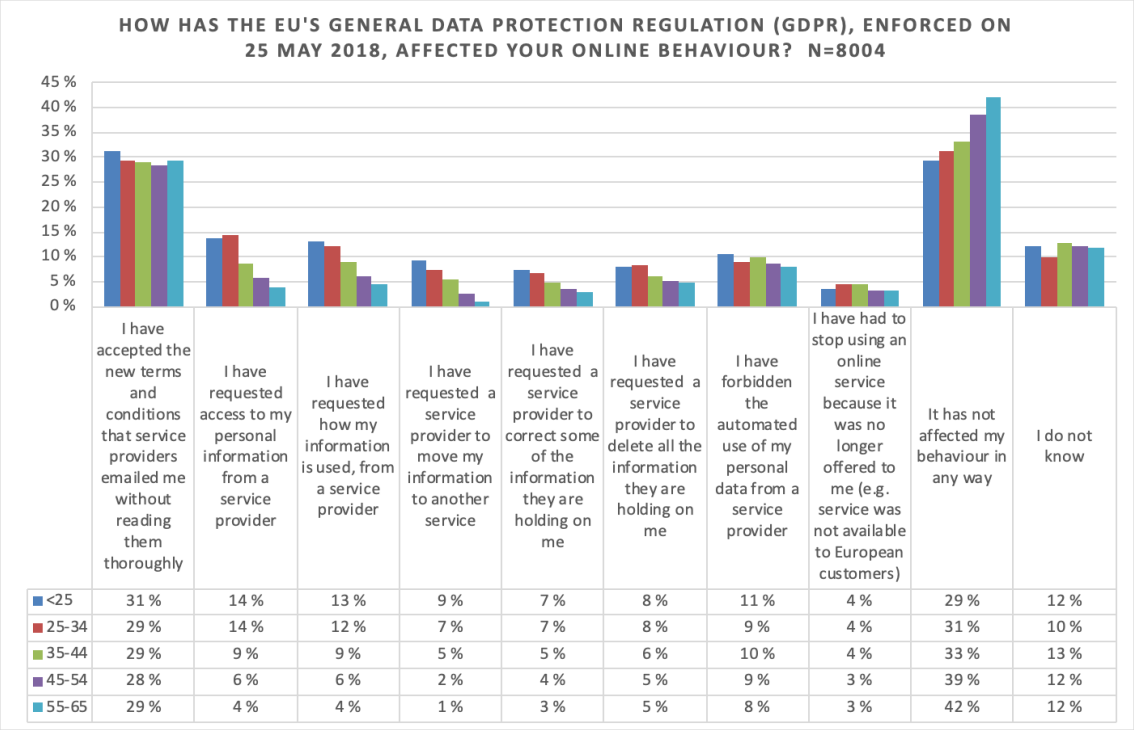
Figure 23. Lack of Trust and the Use of Services by Age.

Figure 24 shows how the GDPR has affected the user's online behavior in different countries. The respondents could choose several options. Most (36%) of all the respondents felt that the GDPR had not affected their behavior in any way. The second popular option was "I have accepted the new terms and conditions that service providers emailed me without reading them thoroughly". Figure 25 shows the responses by age. For the youngest age group under the age of 25, the most popular option was "I have accepted the new terms and conditions that service providers emailed me without reading them thoroughly". For the other age groups, the most common option was that the GDPR had not affected their behavior in anyway. The figure shows also that the younger age groups are more active with requesting information from service providers, and the effect of the GDPR decreases with age.



	I have accepted the new terms and conditions that service providers emailed me without reading them thoroughly	I have requested access to my personal information from a service provider	I have requested how my information is used, from a service provider	I have requested a service provider to move my information to another service	I have requested a service provider to correct some of the information they are holding on me	I have requested a service provider to delete all the information they are holding on me	I have forbidden the automated use of my personal data from a service provider	I have had to stop using an online service because it was no longer offered to me (e.g. service was not available to European customers)	It has not affected my behaviour in any way	I do not know
FI	36 %	7 %	5 %	2 %	3 %	5 %	12 %	6 %	36 %	10 %
DE	29 %	9 %	9 %	4 %	6 %	8 %	8 %	3 %	38 %	8 %
NL	18 %	10 %	11 %	8 %	6 %	8 %	8 %	4 %	34 %	18 %
FR	34 %	8 %	7 %	4 %	4 %	4 %	8 %	2 %	36 %	11 %
TOTAL	29 %	8 %	8 %	4 %	5 %	6 %	9 %	4 %	36 %	12 %

Figure 24. Effect of the GDPR on Online Behavior by Country.



	I have accepted the new terms and conditions that service providers emailed me without reading them thoroughly	I have requested access to my personal information from a service provider	I have requested how my information is used, from a service provider	I have requested a service provider to move my information to another service	I have requested a service provider to correct some of the information they are holding on me	I have requested a service provider to delete all the information they are holding on me	I have forbidden the automated use of my personal data from a service provider	I have had to stop using an online service because it was no longer offered to me (e.g. service was not available to European customers)	It has not affected my behaviour in any way	I do not know
<25	31 %	14 %	13 %	9 %	7 %	8 %	11 %	4 %	29 %	12 %
25-34	29 %	14 %	12 %	7 %	7 %	8 %	9 %	4 %	31 %	10 %
35-44	29 %	9 %	9 %	5 %	5 %	6 %	10 %	4 %	33 %	13 %
45-54	28 %	6 %	6 %	2 %	4 %	5 %	9 %	3 %	39 %	12 %
55-65	29 %	4 %	4 %	1 %	3 %	5 %	8 %	3 %	42 %	12 %

Figure 25. Effect of the GDPR on Online Behavior by Age.

The respondents were asked what kind of factors affect their trust towards service providers when personal data is used for either for service production, development, or targeting or for selling to third parties. We can see from Figure 26, that the ability to delete all personal data and the ability to accept or decline the selling of data to third parties are the most important factors increasing trust towards service providers. In addition, important factors affecting trust are the clarity of terms and conditions, the clarity of data use, and the ability to get information about what kind of data the service provider has about the person. The least significant factors for increasing trust are getting paid or getting extra service in exchange for giving permission to use personal data. Figure 27 shows the factors increasing trust when personal data is sold to third parties. There we can see that the ability to accept or forbid selling personal data to third parties as well as the ability to delete all personal data are the most important factors in trust-building when data is sold to third parties. In accordance with the previous question, getting paid or getting extra service in exchange for giving permission to use personal data are the least important factors in increasing trust when data is sold to third parties.

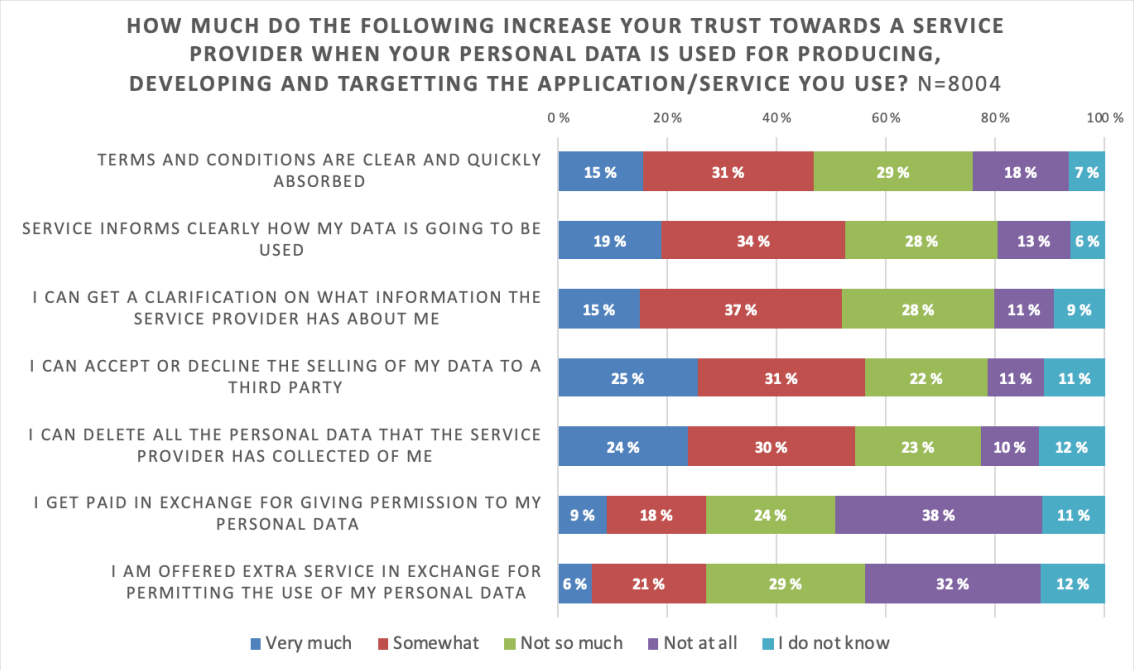


Figure 26. Trust-building Factors in Service Production, Development, and Targeting.

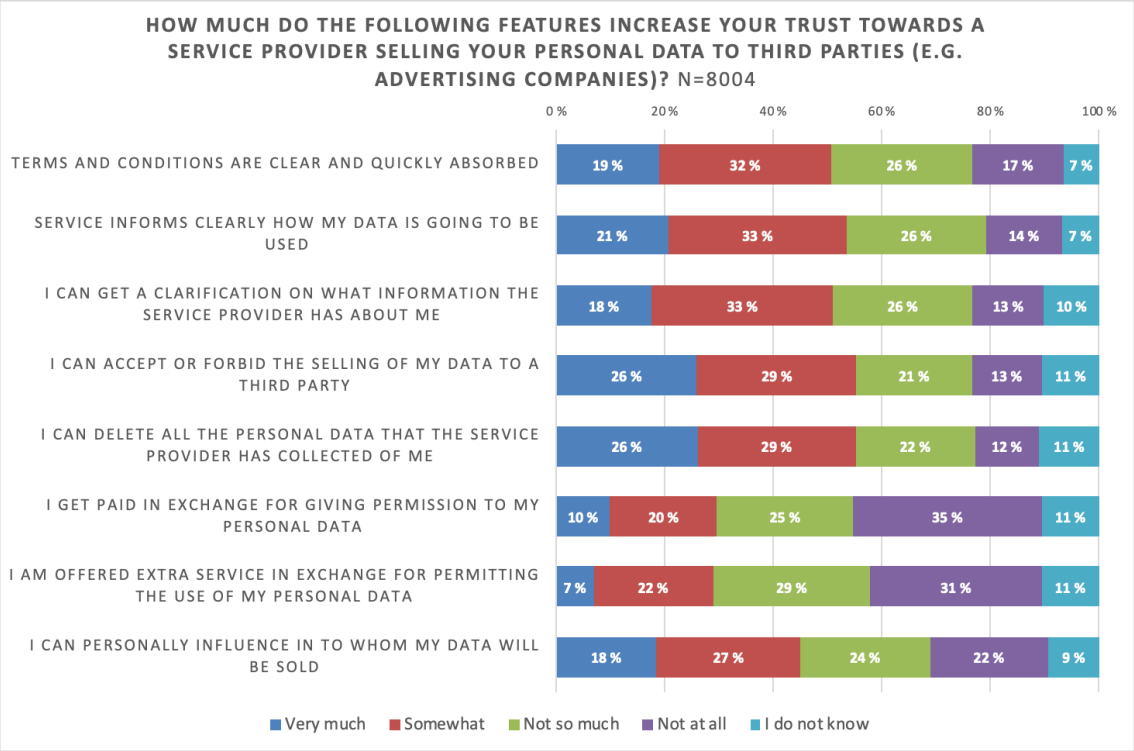


Figure 27. Trust-building Factors in Selling Data to Third Parties.

6.1.3 Management of Information and the Fair Use of Data

The third section of the survey concerns the management of personal information. The respondents were asked under what conditions they would allow access to different kinds of data about themselves. Figure 28 shows that allowing the use of personal data depends greatly on the purpose of data use and the type of information that is being accessed. In general, there seems to be reluctance to allow access to personal information, especially when the information concerns personal wealth and spending. If we look at health or heredity information, 30% of respondents state that they are not willing to provide or allow access to it under any circumstances. 23% would be willing to allow access to this information if it were used for scientific research. What comes to personal data or data about one's personal history, 35% of respondents would not allow access under any circumstances. 16% would allow access to this data if they were paid for it. The same goes for information concerning personal values or beliefs, 28% of respondents would not allow any access, and 18% would allow access if they were paid for it. 29%

of respondents would be willing to allow access to information about their consumption habits or purchases if they were paid for it. For spatial data, 27% would not allow access under any circumstances, and 20% would allow access if they were paid for it. 40% of users would decline all access to information about their wealth and spending, and 18% would be willing to allow access if they were paid for it.

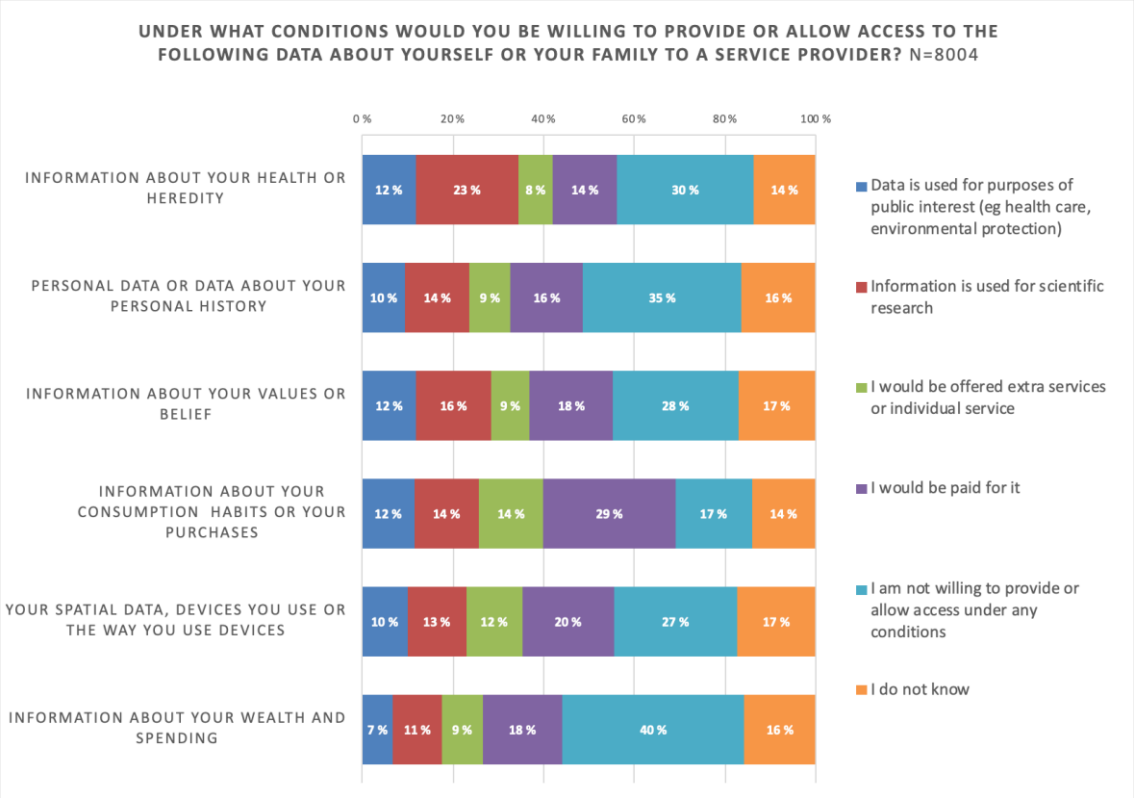


Figure 28. Conditions for Allowing Access to Personal Information.

The respondents were asked what would be the best way for a service provider to collect and handle personal data. As can be seen in Figure 29, most of the respondents in all the countries prefer giving consent separately to each service provider. The next popular option was that the service provider would make the decision on behalf of the individual according to the conditions the individual has given. A large number of respondents are not sure what would be the best way, e.g. in France, 22% of respondents have stated “I do not know”.

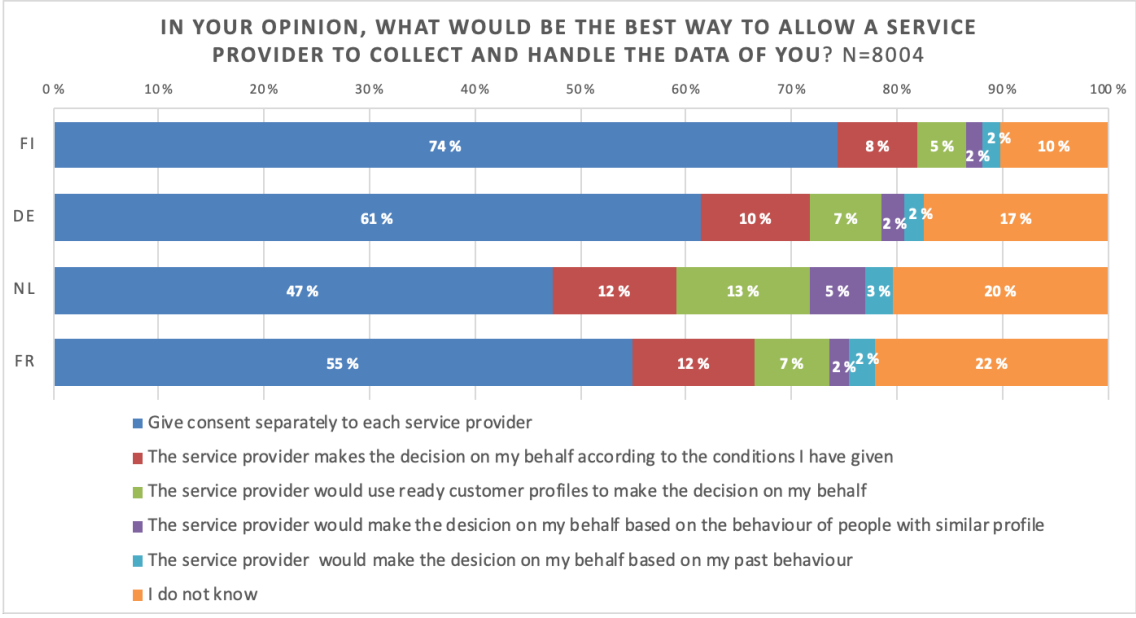


Figure 29. Preferred Ways of Collecting and Handling Data.

The respondents were then asked what kind of features are important for digital services or applications. Figure 30 shows that for the majority of respondents, security of the service is very important. Also, over half of the respondents feel that the reliability of the service provider is very important. Altogether 80% of the respondents feel that it is very or somewhat important that the purpose of the data collection is clearly and transparently reported. In addition, 80% of the respondents thought that ease of use is very or somewhat important. The least important feature was the personalization of services based on previous usage. 11% of all the respondents felt that this feature is not at all important.

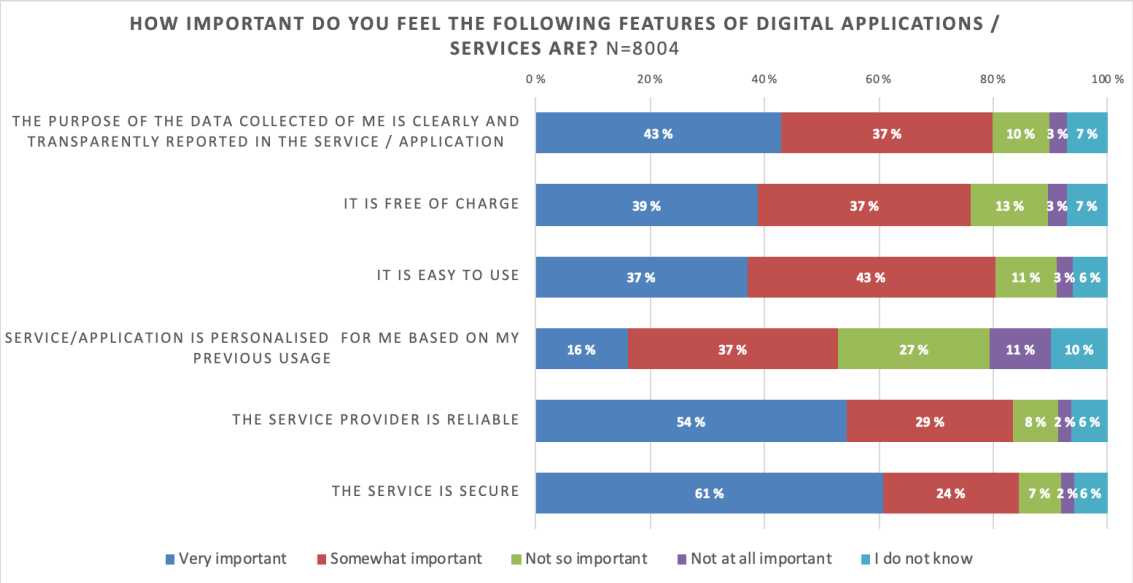


Figure 30. Importance of Features for Digital Services or Applications.

The respondents were further asked how information management and utilization should be developed in comparison to the current situation. In this question, the users could choose only one option from the list. Figure 31 shows the responses by country. The most popular option for the Finnish respondents was that authorities should regulate the use of data more strongly and that organizations should take more responsibility for the ethical use of data. Both of these options was selected by 24% of the respondents. In Germany, 31% of respondents felt that users should take a more active role in managing their data. In the Netherlands, this was also the most popular option, with 23 % of respondents choosing this option. In France, 25 % of the respondents felt that the authorities should regulate data more strongly. If we look at the same question from the age point-of-view, we can see from Figure 32 that the emphasis on the user's own responsibility and the regulation of authorities grows with older ager groups. Respondents under the age of 25 emphasized the organization's role more than in the other age groups.

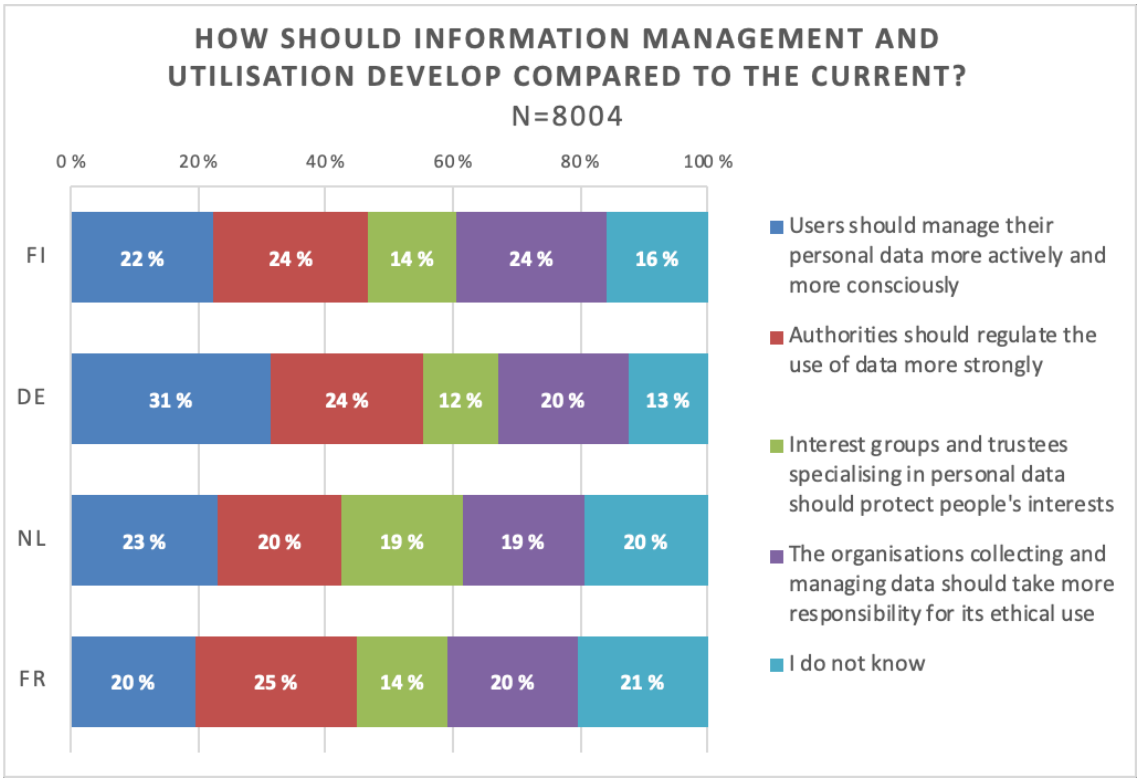


Figure 31. Views on Development of Information Management by Country.

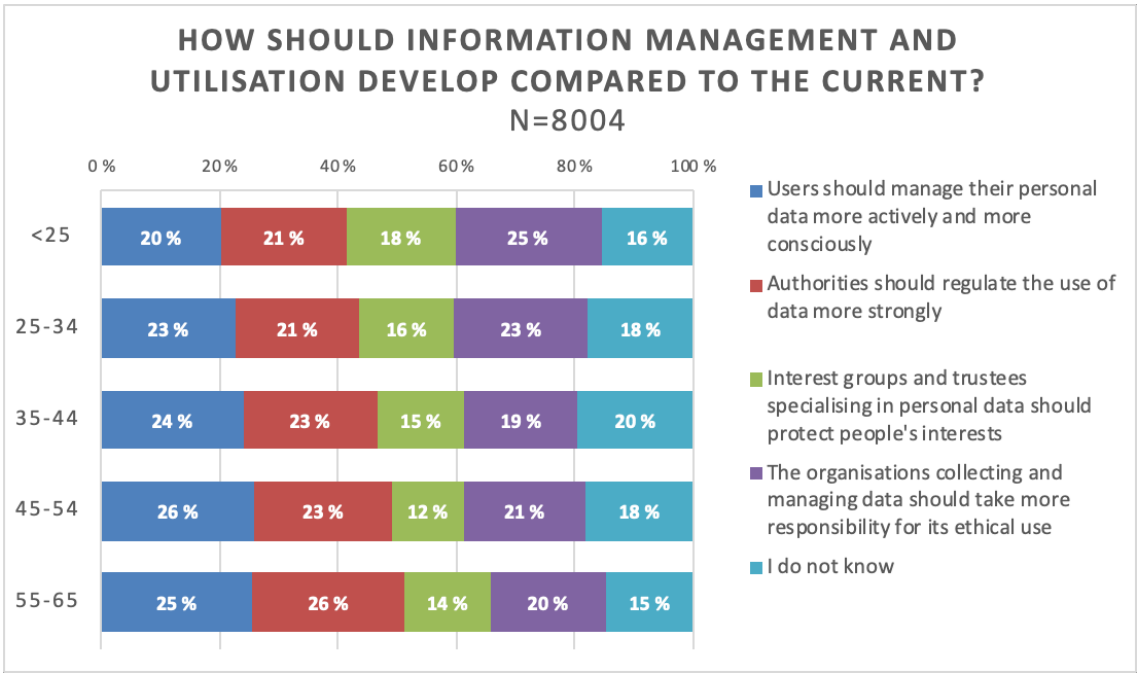


Figure 32. Views on Development of Information Management by Age.

Respondents were then asked if they would use a single application for managing all data collected from different services. Figure 33 shows the results by country and Figure 34 by age. Responses vary by country. 60% of Finnish respondents, but only 37% of French respondents would use a single application for data management. If we look at the responses by age, the willingness to use a single application decreases with age, though respondents between the ages of 55 and 65 seem to be more willing to use a single application than those aged between 44 to 54.

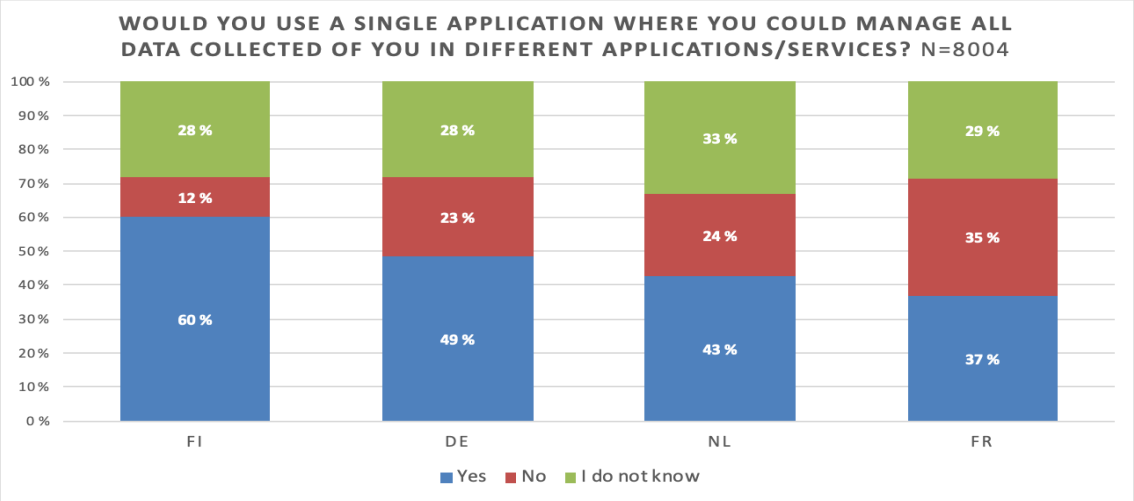


Figure 33. Willingness to Use a Single Application for All Data Management by Country.

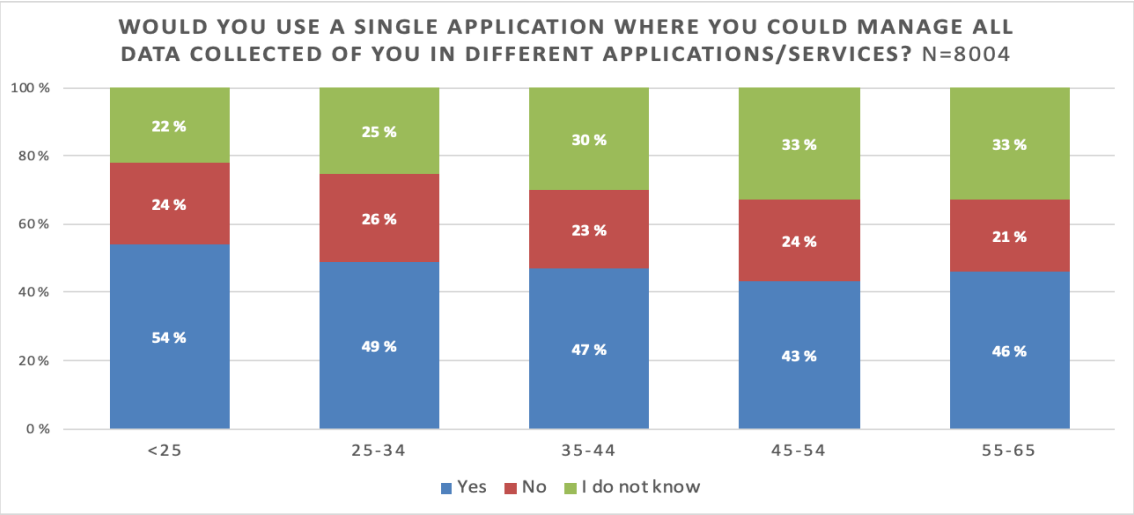


Figure 34. Willingness to Use a Single Application for All Data Management by Age.

Those who answered that they would use a single application for data management, were further asked what kind of features they would expect from that application. The respondents could choose more than one option. The expectations for a single data management application by country can be seen in Figure 35. In Finland, Germany, and the Netherlands, the most important features are safety and reliability. In France, the second most important feature after safety was that the application would be free of charge. For all age groups, safety and reliability are the most important features, as shown in Figure 36.

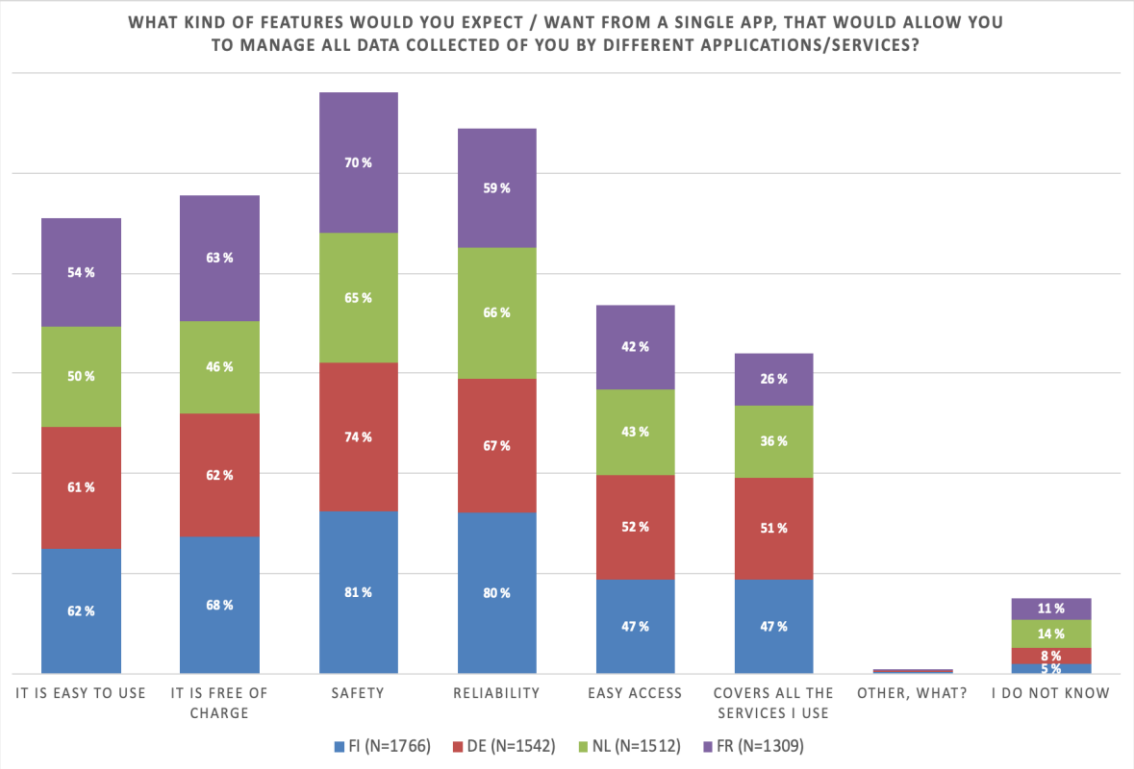


Figure 35. Expectations for Single Data Management Application by Country.

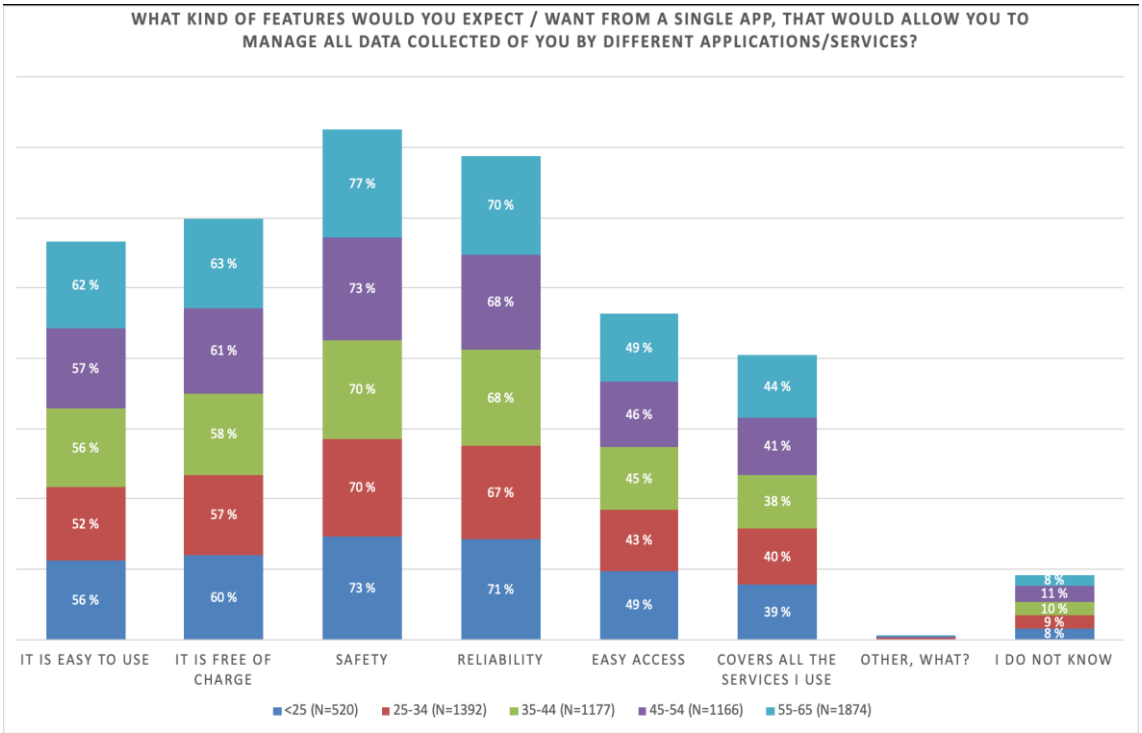


Figure 36. Expectations for Single Data Management Application by Age.

Thirty-six of the respondents who chose the option “other” gave an open-ended answer. The responses varied, but many answered that users should be able to manage and delete data by themselves. Some respondents also mentioned that data management should be done by neutral and independent operators.

6.1.4 Fair Use of Data

The final part of the survey is concerned with the fair use of data. The first question in this section was how service providers should manage personal data for people to feel that it is fair for them. This was an open-ended question, and the number of respondents was 2,386. The responses were analyzed with Voyant Tools, and Figure 37 shows the most common terms in a word cloud. From the figure, we can see that security, transparency, and consent are among the most common terms. Table 1 shows the number of the most frequent terms.

Figure 37. Important Factors for Management of Fair Data as a Word Cloud.

TERM	NUMBER OF MENTIONS
“data”	723
“use”, “used”	573
“security”, “secure”, “secured”, “securely”	282
“transparent”, “transparency”, “transparently”	205
“consent”	166
“collected”	136
“service”	118
“information”	111
“parties”	103

Table 1. The Most Frequent Terms Regarding the Management of Fair Data.

Figure 38 shows the importance of a fair data label for services similar to the Fair Trade Label in consumer goods. A majority of the respondents in all the countries thought that is very or somewhat important that data services would have a fair data label. Figure 39 shows that between age groups there is only little variation in the responses. The majority of the respondents in all the age groups felt that a fair data label would be very or somewhat important.

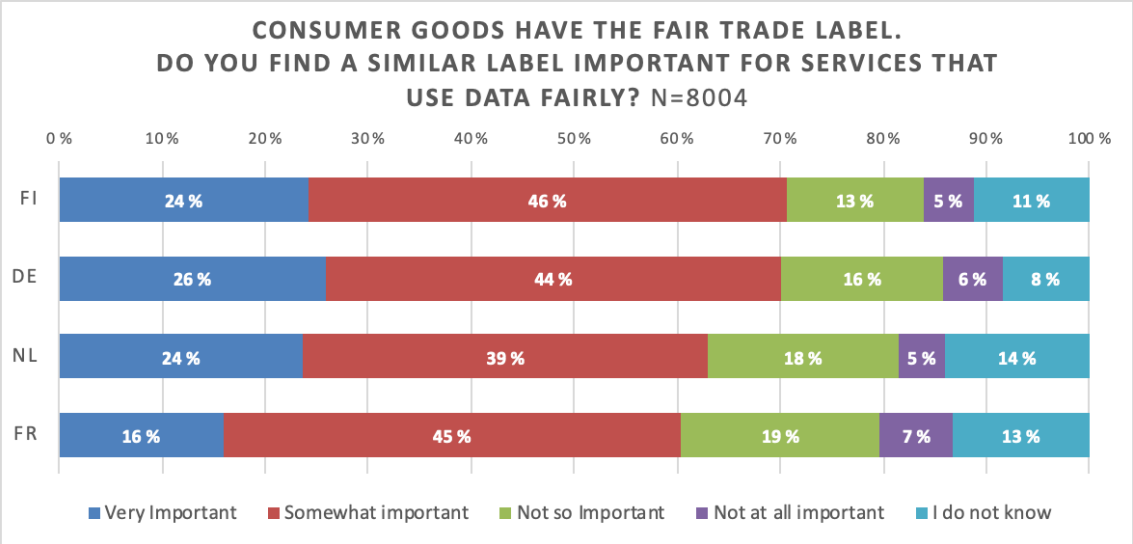


Figure 38. Importance of Fair Data Label by Country.

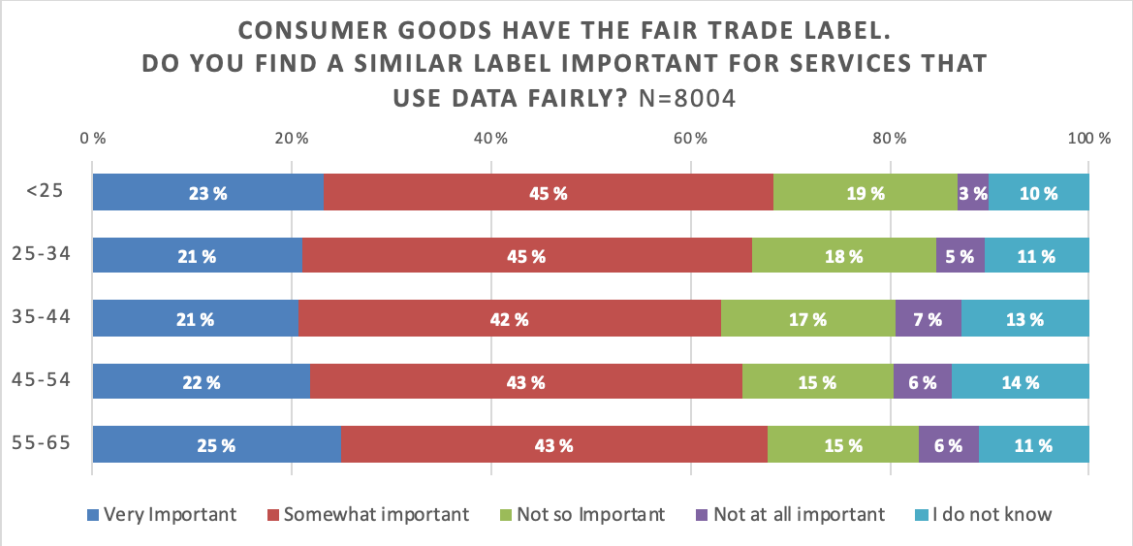


Figure 39. Importance of Fair Data Label by Age.

Finally, the respondents were asked what would be the minimum requirement for a "fair data label" for services that use personal data. This was an open-ended question, and the total number of responses from all countries was 2,406. The responses were gathered together and analyzed with Voyant Tools. Figure 40 shows all the most frequently used terms in the dataset as a word cloud. Table 2 shows the number of the most fre-

quent terms. Some of the terms for the term count were combined, e.g. the terms “trustworthy” and “trustworthiness”. It can be seen from these results, that people value security, transparency, fairness, trustworthiness, and consent in regard to fair data label. There are also several mentions for reliability, openness, permission, honesty, and confidentiality.

Figure 40. Requirements for Fair Data Label as a Word Cloud.

TERM	NUMBER OF MENTIONS
“data”	1063
“security”, “secure”, “secured”, “securely”	510
“transparency”, “transparent”, “transparently”	277
“use”, “used”	276
“fair”, “fairly”, “fairness”	230
“trustworthiness”, “trustworthy”	159
“consent”, “consents”, consented”, “consenting”	137
“parties”	97
“service”	97
“information”	91

Table 2. The Most Frequent Terms Regarding the Requirements for Fair Data Label.

6.2 Content Analysis of #dataehics in Twitter

For the content analysis, the tweets were categorized into themes. Figure 41 shows the themes of the tweets as well as the number and percentage of the tweets in each category. There were 12 of distinct categories that arose from the dataset. There was also a miscellaneous category, which included tweets that could not be categorized in any other category or were otherwise unclear. A big portion of the tweets were about events, either promotion for events or tweets from ongoing events. The next largest categories were critique either to named companies or general critique regarding data use.

I wanted to examine the company positive and company critical themes a little further. The tweets in those categories were categorized into smaller groups. The themes were gathered from literature, i.e. the themes were privacy, identity, transparency, confidentiality, and power. Some of the tweets were assigned to more than one category. Only a few (4) of the tweets did not fit into any of the above themes. Table 3 shows that a large proportion of the company negative tweets were concerned with privacy. Also, confidentiality issues were raised in almost half (45 %) of the tweets. Table 4 shows the themes of positive comments towards companies. Over half (59%) of the tweets were concerned with transparency. Privacy was the theme in almost half (47%) of the tweets.

Category	Definition	Example	Tweet count	Percentage % of all tweets
Events	Promotion of an event, feedback on events, comments about events while they are happening Seminars, conferences, live podcasts	If you are active in #BigData research, policy or industry and are wondering how best to deal with #ethical challenges, also in view of the rise of #AI, this event is for you! Join us in Brussels on Nov 14! Register FREE: https://t.co/gZVs8ISxd2 #beyondprivacy #dataethics https://t.co/ci9Hiv7xek	170	29 %
Company - Critical	Unethical practices of named companies, giving critical feedback on companies	To have that undermined by Google is not just a breach of the rights of the people visiting our site from a Google search - it is also defamatory and potentially damaging to our brand. I am not ok with this and need to get to the bottom of it. #privacy #gdpr #dataethics	74	13 %
General Critique	General criticism about data use not aimed towards a specific company	Your smartphone tracks you even more than you think - Twelve Million Phones, One Dataset, Zero Privacy https://t.co/5gfyHSAMnM #Ethics4EU #DigitalEthics #Ethics #DataEthics	57	10 %
Promotion	Promotion and information of tools and services related to data ethics and data management.	Check this out: a very useful data ethics canvas! #smartcity #opendata #datascience #dataethics #odihq https://t.co/Djx3iDh8tI	52	9 %
Career and Education	Open vacancies, announcements of promotions, career opportunities and education possibilities	The UK Information Commissioners Office appoints its first data ethics adviser: https://t.co/UJKMrIk1qa via @ICOnews #dataethics #personaldata	31	5 %
Regulation	Discussion about regulation and legislative initiatives	A.I. Regulation Is Coming Soon. Here's What the Future May Hold https://t.co/L7r2L93x5a #ai #dataethics	30	5 %
Scientific Reseach	Tweets about scientific research regarding data ethics. Links to research articles.	The myth of anonymisation. New study demonstrates near-perfect re-identification of individuals in any dataset using 15 demographic variables. #Privacy #DataEthics #GDPR https://t.co/T2XGAlEFcI https://t.co/T2XGAlEFcI	27	5 %
Company- Positive	Postitive feedback on organizations' practices regarding data	Fascinating, as a marketer, to hear how the @NSPCC are managing how they restrict how users are tracked via Google products and social media. They're thinking beyond compliance to ethics. Brilliant role models. #ODIFridays #marketing #datamarketing #digitalmarketing #dataethics https://t.co/BuQ4PcxITv	17	3 %
General Positivism	Positive news and views on data ethics, not specific to any company. Innovations and advancements.	Pleased to see moves towards a digital ethics charter for the public service. Culture shift will be key. #DataEthics https://t.co/2hF9GbUQM8	12	2 %
Networking	Invitation to join a network, announcing networks, asking for advice or recommendations	Now there's no excuse. We are here. Honoured to be with these awesome women https://t.co/z3Mpq5B2HF #aiethics #dataethics #DataEthics #PropTech https://t.co/klb1Bdxysi	10	2 %
Good Practices	Instructions, guidelines, and good practices regarding data ethics	10 Ways to Apply Ethics to AI #Ethics4EU #DigitalEthics #Ethics #DataEthics #datascience #MachineLearning #artificialintelligence Part 1- https://t.co/zGBYWb0v2P Part 2-https://t.co/JtZ3J9rpzY	9	2 %
Future Propspects	Future prospects for data ethics and data management.	Predictions #2020 #Privacy and #data ethics bring #marketing to the boardroom #dataprivacy #personaldata #dataethics #business #law #legal https://t.co/R8A2eZpj63	9	2 %
Miscellaneous	Tweets that cannot be categorized in any of the above categories.	Clinical coding humour (with a serious point). #ICD10 #Billing #DataEthics #DigitalHealth #Ontologies #HIT https://t.co/7A4uH4LhEs	81	14 %
TOTAL TWEET COUNT			**579**	**100 %**

Figure 41. Distribution of Tweets by Themes.

Company Critique	Frequency	Percentage of total tweets in theme (n=74)
Privacy	58	78 %
Confidentiality	33	45 %
Transparency	22	30 %
Identity	16	22 %
Power	14	19 %

Table 3. Themes of Company Critique Tweets.

Company Positive	Frequency	Percentage of total tweets in theme (n=17)
Transparency	10	59 %
Privacy	8	47 %
Confidentiality	5	29 %
Power	4	24 %
Identity	2	12 %

Table 4. Themes of Positive Company Tweets.

I also wanted to find out what kind of recurring issues are discussed and for this purpose I created a word cloud with the help of Voyant Tools. The word cloud includes tweet content as well as the hashtags used in the tweets. The tool automatically removes stop-words from the data set. Stopwords are words that do not bring much meaning to the text, e.g. articles, prepositions and pronouns, and removing them reduces the noisiness of data (Saif et al, 2014). I also included URL's to the stopwords list and excluded the original hashtag #dataethics from the list, because it appears in every tweet. Figure 42 shows the most frequent terms used in the tweets as a word cloud and table 5 shows the number of the most frequent terms. We can see that AI and privacy are some of the most frequent terms. GDPR is also among the 15 most frequent terms.

Figure 42. The Most Frequent Terms in Tweets in a Word Cloud.

TERM	NUMBER OF MENTIONS
“data”	275
“ai”	176
“ethics”	133
“privacy”	101
“aiethics”	57
“ethical”	50
“great”	47
“new”	44
“use”	38
“dataprotection”	37
“gdpr”	36
“dataprivacy”	35
“datascience”	33
“digital”	33
“public”	33

Table 5. The Number of Most Frequent Terms in Tweets.

6.3 Findings

This chapter examines the conclusions that can be drawn from the research results. First, I will explore the results of the "Use of Digital Services" survey. The research showed that, in general, individuals are not very well aware of their rights regarding privacy and data use. The knowledge about data rights also varies in different countries. What comes to terms and conditions, there is variation on how the terms and conditions are understood in different countries. A large amount of people (25%) do not understand the terms very well or not well at all. People also do not read the terms and conditions of services very thoroughly. If we look at the changing of settings according to the user's needs, younger people are generally more eager to change the settings of services or applications according to their own needs. The reasons for not changing the settings vary with age. Younger people feel that changing settings is not important, and they do not want to spend time on it. Older people believe more that changing the settings does not have any effect, or they are unaware of how it is done.

If we look at trust towards service providers, we can see that it varies in different countries. Roughly 40% of people feel that the lack of trust prevents them from using digital services, whereas 25-30 % disagree with this statement. The news about data leakages has not affected the behavior of over a third of the respondents. Still, nearly 40 % state that they have either stopped or reduced using some services. In terms of age, the effect of data leakages was more significant for younger people, whereas older people felt more that the news haven't had any effect. If we look at the factors that affect trust, the ability to delete all personal data and the ability to accept or decline the selling of data to third parties are the most important factors that increase trust. Getting paid or getting extra service in exchange for giving permission to use personal data are the least important factors in increasing trust. The GDPR has not affected the behavior of over a third of the individuals, especially for the older age groups.

When it comes to data management, individuals are generally reluctant to give access to personal information under any circumstances. Still, people would be willing to allow

access to information about consumption habits and previous purchases if they were paid for it. People are also willing to give access to personal information for scientific research. The reliability and security of digital services or applications are very important to the majority of people. In addition, transparency, ease of use, and chargeless services are important. Personalization was quite surprisingly the least important factor.

Opinions about data management and utilization in the future vary in different countries. In Finland, the emphasis was on authorities and organizations, whereas in Germany the emphasis was more on the user's own activity. The emphasis on the user's own responsibility and the regulation of authorities grows with older ager groups. Respondents under the age of 25 emphasized the organization's role more than in the other age groups. The acceptance of a single application for all data management varies in different countries, Finns being the most eager to use a single application, and the French the least. Safety and reliability were the most important features for a single data management application in all the countries and age groups. The respondents felt that the fair use of data includes security and transparency and bases on user consent. A majority of respondents in all the countries and age groups would welcome a fair data label for services. The requirements for a fair data label would include security, transparency, and trustworthiness. Again, consent was also at the top of mentions.

Next, let's take a look at the results of content analysis. It is no surprise that events were such a significant proportion of twitter conversations about data ethics. Twitter is an excellent platform for spreading the news as it employs speedy communication, and it does not involve very profound communication due to its nature. Conversations and topics are quickly replaced by new ones. Still, Twitter is used quite a lot for reporting on the unethical behavior of organizations. Critical tweets were more common than positive ones. The most intriguing categories for this research were the tweets that contained criticism or praise towards companies. When looking at the critique towards companies, the majority of the tweets were concerned with privacy. Privacy also goes hand in hand with confidentiality, as a breach of confidentiality means that personal information is

leaked or sold. In the positive tweets, organizations were praised for transparent actions regarding privacy and confidentiality. The uses of AI raised a lot of concern. It is natural, that it creates a lot of worry among consumers, as it is being utilized increasingly. Also, AI and its implementations are still somewhat a mystery to the general public, so this might be the reason it causes so much concern. Not surprisingly, privacy was among the top most frequent terms.

The aim of the mixed method research is to examine a subject from different perspectives and analyze the similarities and differences brought forth by the different researches. Undoubtedly, it is hard to find similarities between the two pieces of research, because they are so different in nature. However, both of the researches showed that privacy and transparency are critical issues for consumers. Both of the research results are consistent with the literature, and the same kind of themes arose from literature as well as from the empirical research, i.e. the issues of privacy, transparency, confidentiality, identity, and power. Confidentiality was seen as an essential issue in the tweets, and many of the critical tweets were concerned with third-party access to personal data. Mergers and acquisitions were also seen as a threat to confidentiality. In addition, it was shown in the survey that people want to have the power to decline the selling of their personal data to third parties and the power to delete or adjust their personal data. Mergers and acquisitions are also linked to power asymmetry, as in the tweets they were seen as a manifestation of increased power of organizations and a threat to identity. In the survey, the power asymmetry came apparent in that many people feel that changing the settings on services and applications has no effect. Also, we can speculate whether data access is seen as a threat to identity, as so many individuals were reluctant to give access to their information in any circumstances.

7 Conclusions and Discussion

The overall aim of this research was to investigate what kind of role ethics play in data-driven marketing and what kind of ethical challenges marketers are faced within the current business environment.

In the context of data-driven marketing, the specific research objectives aimed to answer the following questions:

1. How Big Data affects the business and consumer environments?
2. What kind of challenges are involved in the use of personal data?
3. What is the role of ethics in current marketing?
4. What kind of implications does the concept of privacy pose to marketers?
5. How data can be used in an ethical way in marketing?

This chapter will summarize the findings regarding each research objective and present conclusions based on those findings. In addition, recommendations and managerial implications are discussed, as well as recommendations for future research. Also, the limitations of the research are contemplated in this chapter.

7.1.1 The Effect of Big Data for the Business and Consumer Environments

The business and consumer markets have undergone a series of changes in the past decades. The biggest influencer in this change has admittedly been technology and the rise of the Internet and social media, and especially the rise of Big Data. The technological changes, as well as globalized competition, have made it difficult for organizations to differentiate from the competition.

In the modern world, consumers are faced with endless amounts of information available in seconds, and this has created the need to filter and personalize online content to suit the individual needs of customers. In addition, while consumers operate in the online world, they leave trails of data in the form of e.g. online transactions, social media

communications, and web searches. Big Data enables organizations to gather and combine data about consumers from various sources. As the data is combined and analyzed, organizations can get insights about their current and potential customers, which they can use to plan and implement in marketing. Hence, Big Data can be seen as the cornerstone of personalization. Big Data is so powerful compared with traditional data because organizations are able to gain extensive and detailed knowledge about consumers and their behavior. The information composed this way may have huge societal impacts, from threats to identity to institutional surveillance. This is why we need safeguards to ensure that consumer data is not used for wrong purposes. Regulations regarding data use are one way of ensuring consumers' rights. Still, as established in this thesis, they are not sufficient, and marketers and organizations have a significant role in this picture, ensuring that consumers do not lose their trust and continue using services.

7.1.2 Challenges in the Use of Personal Data

Though the use of personal data brings many benefits for organizations, it does not come without challenges. Undoubtedly, the most crucial challenge for organizations in the use of personal data is how to build and maintain trust. Consumers' use of personalized services is dependent on whether the organization is perceived to be trustworthy or not. The empirical research and literature have shown that the lack of trust reduces the use of digital services. Transparency and clear privacy policies can then be seen as a key ingredient in trust building.

For consumers, the biggest challenge in managing data is that consumers are often unaware of how their data is being used. Also, individuals may not understand the terms and conditions they agree to. The challenge with personalization is the paradox between personalization and privacy. Though consumers want services to be personalized to their needs, they also value their privacy. If personalization is seen as posing a threat to privacy, consumers usually reduce or stop using the service. The biggest challenge with personalization is that filtering of content often occurs silently. This, too, underlines the need

for transparency. Segmentation based on data may paint a too simplistic image of consumers. Hence, organizations should remember not to rely too heavily on data and algorithms, as it can lead to ignoring the human aspects of communication. To gain competitive advantage, organizations should always consider the consumers' interest in the long run.

7.1.3 The Role of Ethics in Current Marketing

As organizations are part of society, they need to abide by the commonly accepted norms of behavior. Though there are regulations protecting the personal data of consumers, regulations can only be seen as minimum requirements for behavior. The general view on acceptable behavior usually goes far beyond regulations. This is why organizations need to accept that ethics must be taken into account in their operations. In fact, the need for ethics is even more apparent in marketing than ever before, because of the broad access to personal data. The power of Big Data does not come from the raw data itself, but the ability to combine and merge data, thus creating detailed insights about consumers. There also lies the foundation for ethical dilemmas.

The basis of ethics is trust; consumers need to trust that organizations behave in a socially responsible manner and handle consumers' data in a trustworthy manner. The content analysis showed that transparent organizations receive praise from consumers. Those organizations, which jeopardize the privacy and confidentiality of information, receive heavy criticism. In addition, the survey showed that reliability and security are essential for consumers using digital services. Consumers are becoming more aware of ethical issues in marketing, and in the future, ensuring ethics in all business operations is a way to ensure success in the long-run.

7.1.4 The Importance of Privacy

Privacy as a concept is somewhat problematic, as the need and the perception of privacy is not a fixed one, but it rather fluctuates depending on the context and the situation.

Still, in data-driven marketing, privacy is a key issue. Consumers are concerned about the privacy of their data, and there is a growing distrust between consumers and organizations. Organizations are also not trusted as much as fellow consumers. There has been a lot of talk about users' consent to data gathering. Do consumers really know what kind of permission they are giving for the use of their data, how the data is used, and who is using their data? Individuals should be given the ability to control their personal data and also the ability to see exactly how their data is used. That way, they can make better-informed decisions regarding allowing access to personal data. More control for individuals increases trust, which in turn makes individuals share their data more eagerly.

According to the survey, individuals want authorities and organizations to have a bigger role in data management. Still, individuals themselves need to take more responsibility over sharing their personal data. If we think that data ownership changes when the data is released, then we should be more careful when giving out our personal data in the first place. This view puts more emphasis on the individual's own actions. One of the main problems with this is that individuals do not necessarily understand how their data is being used. In addition, individuals don't have much means to fight over the big data brokers. Therefore, we need marketers and organizations to take the stand for privacy. Caring for the customer's privacy also indicates that organizations value customers' needs, and this can be a big differentiating factor, increasing loyalty and trust.

7.1.5 Ethical use of Data in Marketing

Being ethical in the use of data is essential for maintaining trust between organizations and consumers. The same rules apply to data use in marketing, as to other uses of data. Data, and especially Big Data, can create complex ethical dilemmas. Ethical challenges arise especially from combining consumer data from different sources and from selling data to third parties. There are some ways that organizations can minimize the ethical dilemmas brought by Big Data. First, organizations need to ensure that the privacy of individuals is not jeopardized. In order to achieve this, clear privacy policies should be established, and they should be communicated in an understandable and transparent

way. Second, the confidentiality of personal data should be protected. This requires sufficient data protection methods. In addition, if the data is sold to third parties, it should be stated in an understandable way. Consumers also want to feel that they have the power to manage personal data. Giving the consumers the ability to opt-out from data use and data selling without it affecting the use of services can create an enormous competitive advantage in the future. Though this may lead to decreased profits in the short run, in the long term, it may increase the loyalty of consumers.

The most effective way to embrace ethics is to integrate it with the operations throughout the organization instead of being an add-on. Having an ethics-based organizational culture ensures that ethical norms are also followed, not merely stated. Consumers should be viewed as collaborators than mere targets. This changes the whole perspective of the organization and its role in society.

7.2 Summary of Findings and Conclusions

All in all, we can see from the survey results that the people's perception of the use of data varies in different countries, and there is not a consistent European view on these matters. There was more consistency with the results when looking from the age point of view, younger people being more active e.g. with changing the settings on applications and services. The results of the survey are still somewhat contradictory. It can be stated that, in principle, individuals value privacy and security very high. However, in practice, data leaks do not affect their behavior in any way for some of the people. In addition, people in principle are not willing to give access to their personal information, but they are also not very eager to change the settings in services or applications.

What all of the above then mean for data-driven marketing? First, if data is used for marketing, the same assumptions apply to it as in other data uses. Consumers fear for their privacy and do not trust organizations with their personal data. Ensuring that the control over data remains within the consumers is a way to increase trust. Transparency is also very important for consumers, and it is an excellent way for organizations to stand

out from the competition. Surprisingly, personalization was not seen as an important factor for a digital service in the survey. Do consumers really understand how personalization works, and would they be willing to lose it? Personalization is such an integral part of modern marketing, and it also affects the usability of services. If the content is not filtered and personalized, it might result in an overload of irrelevant information. The services would not be then as easy to use, one factor that was viewed as important in the survey.

Organizations often utilize external data for marketing, e.g. through Facebook's or Google's marketing tools. Though the organization itself is not gathering data, it is buying it. Often organizations do not think about how ethical their partners are. However, the requirements for ethics should be extended to cover any additional parties. If an organization is committed to acting ethically, it should require ethical conduct also from external parties. Still, in practice, this may be very difficult to execute. However, ethical behavior should be the starting point of operation, rather than being an add-on. This means that genuinely ethical organizations think about how ethics can be employed in every aspect of operations. It should be the core value of an organization and part of the organizational culture.

If we think about consumers' power on data use, they can be seen as influential players in the market, e.g. in setting up boycotts and informing their peers about unethical conduct by companies, as seen by the content analysis of tweets. The question lies, how influential boycotts are for individuals if the harm is already done, and personal information jeopardized. Of course, boycotts do have impacts on organizations' future revenue and brand image. In addition, boycotts may draw attention to ethical issues so that more organizations start to think about ethical issues proactively. Trust takes a long time to build but can be lost in seconds. Responsible marketers take ethics seriously and understand that ethical practices bring benefits. It is advisable to override short-term gains with long-term vision. Of course, organizations need to balance between profitability and customer satisfaction. Still, these two do not need to be mutually exclusive. If an

organization is able to plan operations with long-term objectives in mind, it is possible to ensure both ethical conduct and profitability. Customers reward those kinds of organizations with trust and loyalty.

7.3 Limitations

This chapter discusses the limitations and potential problems related to the chosen research strategy and methods. Mixed method study is often used when the researcher wants to get a holistic view of the issue being researched. However, Silverman (2013, pp. 138) warns that the researcher needs to be very careful with this research method as there is a risk that one or other of the datasets is under-analyzed. He also states that trying to achieve the whole picture is an illusion and cannot result in an overall truth. Creswell (2005, pp. 535) also states that the use of mixed method research requires understanding both quantitative and qualitative research.

The empirical part of this research uses secondary data. One of the main problems with using secondary data is that the data is collected for another study with different research objectives and hence may not be the best fit for the research problem at hand. Moreover, the researcher is always responsible for the accuracy of data. Secondary sources cannot be blamed for inaccuracy. Therefore, the researcher has to check the accuracy of the original source of data and to make sure that the original research is of high quality. (Ghauri et al., 1995, pp. 56.) Though this research provided valuable insights on people's perception of data use and trust towards service providers, it did not answer directly about people's opinions of data use specifically in marketing.

Content analysis was used as a research method as well and it also has some drawbacks. The most crucial challenge with conducting a social media content analysis is the ethical issues. Though Twitter is a public social media channel, many users may still find their content to be private, and though they agree on the terms and service when they sign up, it may come as a surprise that the content can be used for research. However, Thelwall (2014, pp. 85) suggests that because tweets are generally public and available

for everyone to read, they should be treated as documents instead of human-related data, and therefore they do not have the same ethical or privacy requirements as e.g. an interview or questionnaire data. Also, this research does not reveal the individuals behind tweets, so their privacy is not jeopardized.

The sample may not also be representative of the population, as not all people use Twitter. Only a small amount of people are active online, and often like-minded people form groups where they share opinions about a subject. Other drawback is the representation of tweets that are gathered. It has to be remembered that not all Twitter users post anything but merely read other's posts. Also, those who post their opinions online may have more negative or stronger views on matters than people on average. In addition, the search engine in Twitter does not give out a complete archive of posted tweets but produces a representation of the tweets. Therefore the study of tweets may overemphasize more central and active users. (Gafney & Puschmann, 2014, pp. 64-65; Hakala & Vesa, 2013, pp. 224.) With manual processing of the data, the analysis is also subject to human error. When categorizing tweets, there is a possibility of error, and the categorization may be biased by the researcher's own opinions or attitudes.

Qualitative content analysis aims at making perceptions about a phenomenon, and it does not aim at producing generalizable facts. When studying online conversations, there is no certainty of the demographics of the people in those conversations. Though the research does not produce quantifiable results that can be generalized, it can produce indicative notions and provide the researcher with an idea of what is significant in a phenomenon. (Hakala & Vesa, 2013, pp. 222, 224.)

7.4 Credibility and Validity

The methods for analyzing research's credibility and validity are different for quantitative and qualitative studies. As mixed methods research combines both quantitative and qualitative research, often validation needs a mixed process as well. (Hesse-Biber, 2010, pp. 87). In this chapter, I will evaluate the validity of the two studies independently.

Quantitative research needs to be objective, reliable, and valid, and it needs to produce generalizable results. Objectivity means that the data and findings are the same irrespective of who does the research and reliability means that the data has to be error-free. Validity means that the instruments for data collection and analysis measure what they were meant to measure. (Schreier, 2012, pp. 26-27.) For secondary data analysis, it is important to evaluate the original research in terms of trustworthiness. The quality of the data is essential, and the researcher needs to have additional metadata about the research available to evaluate the validity and reliability of the original research. The metadata can be e.g. the purpose of the research, data collection details, and entities being researched. (Hox & Boeije, 2005.) The validity of the secondary is increased by the fact that the research reports as well as the raw data of the survey are publicly available. The original survey stated that the results are representative of the whole population of the countries involved in the ages of 18 to 65 (Hyry, 2019). In this sense, it can be said that the results of the survey can be generalized to the whole population. In the analysis phase, the research results were rechecked at the end of the research to ensure research validity. There were minor differences in some of the research results compared with the primary research, probably resulting from using raw data instead of the weighed data. In addition, some of the percentages may be rounded differently, creating slight differences between the primary and secondary analysis of the survey. However, the differences were so small that they do not affect the overall conclusions that can be made from the results.

Whereas quantitative studies can be evaluated through objectivity, reliability, and validity, these concepts are not so clear-cut in qualitative research. One way to ensure the reliability of qualitative research is to have several researchers working on the material. (Schreier, 2012, pp. 19, 26.) The content analysis in this research was not implemented by another researcher, so there is a possibility that another researcher would analyze the dataset differently. However, qualitative studies are always based on interpretation as they deal with symbolic material and are dependent on the context. Therefore, it is

possible to make different interpretations of the same material, both of them still being equally valid. Still, the results of qualitative research need to be trustworthy. In content analysis, objectivity and reliability have a bigger role than in other qualitative studies. (Schreier, 2012, pp. 20, 34.)

An essential aspect of evaluating qualitative research is validity, which means that the entire study and its results and findings need to be valid and of good quality. To achieve validity, the research needs to be conducted in a systematic way, and the procedure and reasoning need to be transparent. Also, the research design and method need to suit the research problem. Double-coding by another researcher or by the researcher himself adds to the reliability of the results. Validity also comes from the content analysis process, as it is data-driven, and categories are adapted in the analysis process. (Schreier, 2012, pp. 27, 34-35.) As explained in the Research Methods chapter, the process of content analysis included a phase where the initial codes were rechecked in order to confirm that the tweets were categorized correctly. The procedure for the data analysis was also carried out systematically and explained in detail. These procedures enhance the reliability and validity of the results.

7.5 Recommendations for Practitioners

Organizations are faced with the dilemma of making profits, and on the other hand, accepting that they are part of society and, therefore, subject to its norms. Privacy, transparency, and consent are themes that continuously arise when discussing data ethics and ethics in marketing. Therefore, marketers need to realize their role in the battle for privacy. They simply cannot hide away from the fact the personal information can be used irresponsibly. Instead of just talking about ethics, real action needs to be taken. Organizations should be proactive in developing and deploying marketing ethics strategies and practices to meet the high demands of consumers, stay competitive, and embrace their roles as essential players in society.

Organizations need to decide how ethical issues will be dealt within the organization. All the ethical theories introduced in this thesis have both benefits and drawbacks. The utilitarian approach may not be sufficient in the modern world anymore. If the chosen strategy is utilitarian, then the approach is more unorganized and reactive as ethical dilemmas are contemplated as they appear. Using a normative strategy provides the organization with a more effective capability to address ethical dilemmas as ethical codes are decided in advance. However, real-life situations are often more complicated than what was imagined beforehand. If the chosen approach is virtue ethics, then the organization bases its actions on commonly accepted virtues. The risk with this approach is that organizations may often simply state e.g. that "We value your privacy". What these statements mean in practice may be left vague. It is then left to the consumer's own judgment to decide whether these statements are trustworthy or not. Therefore organizations need transparency and concreteness in their statements for ethics. It is not enough that an organization merely states that it values ethics; it needs to have proof of it as well. Having an ethics framework or guidelines creates consistency and builds trustworthiness.

If organizations and marketers want to embrace ethics, they should also require ethical behavior from the providers of tools and technologies they are using. This is quite an idealistic and heavy demand, and balancing ethics with practice can be daunting. At the very least, organizations should acknowledge the potential ethical issues related to the tools and technologies they use. As a result, organizations can make informed decisions on whether or not to continue using these tools. The use of tools that e.g. pose threats to privacy should be made aware to the customers as well, thereby increasing transparency. On the practical level, it is also vital that consumers have the option to opt-out from personalization without it affecting the use of the service. This way, consumers can feel more in control of their personal data. In addition, organizations should have clear data privacy and security policies and they should be communicated in an understandable way.

7.6 Suggestions for Future Research

This research has examined the concept of data ethics in marketing from a general point of view. This research established that there is a need for organizations to think about data ethics when using data in marketing, and ethics is linked to almost every aspect of marketing. The emphasis of the empirical research was on consumer perception of ethical data use. In the future, it would be interesting to study how organizations in Finland and in Europe implement data ethics in practice, and what kind of data ethics strategies or policies have been established by organizations. It would be valuable to find case study organizations to study what kind of impacts ethical data use has had on organizations and their customers. In addition, the effects of the GDPR could be further examined; what are the consumers' views on the effectiveness of the regulation. Though the secondary analysis of the "Use of Digital Services" survey touched on this matter, the GDPR was still new at the time the survey was made. Therefore, it could be useful to study the effects of GDPR after more time has passed. In addition, it would be valuable to study the consumer views related to the use of data specifically in marketing.

References

Aguirre, E., Roggeveen, A. L., Grewal, D. & Wetzels, M. (2016). *The Personalization-Privacy Paradox: Implications for New Media.* Journal of Consumer Marketing, 33(2), pp. 98-110. doi:10.1108/JCM-06-2015-1458

Al-Khouri, A. M. (2012). *Data Ownership: Who Owns "My Data".* International Journal of Management & Information Technology, 2(1), 1-8.

Amnesty International (2019). *Surveillance Giants: How the Business Model of Google and Facebook Threatens Human Rights.* Retrieved April 29, 2020, from: https://www.amnesty.org/en/documents/pol30/1404/2019/en/

Arthur, L. (2013). *Big Data Marketing: Engage Your Customers More Effectively and Drive Value*. Wiley.

Batrinca, B., & Treleaven, P. C. (2015). *Social Media Analytics: A Survey of Techniques, Tools and Platforms.* Ai & Society, 30(1), 89-116.

Bengtsson, M. (2016). *How to Plan and Perform a Qualitative Study Using Content Analysis.* NursingPlus Open, 2(C), pp. 8-14. doi:10.1016/j.npls.2016.01.001

Bivins, T. (2009). *Mixed Media: Moral Distinctions in Advertising, Public Relations, and Journalism* (2. ed.). New York: Routledge.

Bozdag, V. & Timmermans, J. (2001). *Values in the Filter Bubble: Ethics of Personalization Algorithms in Cloud Computing*. 1st International Workshop on Values in Design – Building Bridges between RE, HCI and Ethics, Lisbon, Portugal, 6 September 2011.

Brown, B., Chul, M. & Manyika, J. (2011). *Are You Ready for the Era of 'Big Data'?* McKinsey Quarterly, 4, pp. 24-35.

Cavoukian, A. (1999). The Promise of Privacy-Enhancing Technologies: Applications in Health Information Networks. In Bennett, C. J. & Grant, R. A. (Eds.), *Visions of Privacy* (pp. 116-128). University of Toronto Press.

Cetinã, I., Munthiu, M., & Radulescu, V. (2012). *Psychological and Social Factors that Influence Online Consumer Behavior.* Procedia - Social and Behavioral Sciences. 62. 184–188. 10.1016/j.sbspro.2012.09.029.

Chellappa, R. & Sin, R. (2005). *Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma.* Information Technology and Management, 6(2-3), pp. 181-202. doi:10.1007/s10799-005-5879-y

Chiu, S. & Tavella, D. (2008). *Data Mining and Market Intelligence for Optimal Marketing Returns.* Butterworth-Heinemann.

Computer Weekly (2019). *State of GDPR in Mainland Europe.* Retrieved May 1, 2020 from: https://www.computerweekly.com/ehandbook/State-of-GDPR-in-mainland-Europe

Court, D., Elzinga, D., Mulder, S. & Vetvik, O. (2009). *The Consumer Decision Journey.* The McKinsey Quarterly, 3, p. 96.

Creswell, J. W. (2005). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (2nd ed.). Upper Saddle River, NJ: Merrill.

Cui, G. & Choudhury, P. (2003). *Consumer Interests and the Ethical Implications of Marketing: A Contingency Framework.* Journal of Consumer Affairs, 37(2), pp. 364-387. doi:10.1111/j.1745-6606.2003.tb00458.x

Cukier, K., & Mayer-Schoenberger, V. (2013). *The Rise of big data: How it's Changing the Way We Think About the World*. Foreign Aff., 92, 28.

European Economic and Social Committee (2017). *The Ethics of Big Data: Balancing Economic Benefits and Ethical Questions of Big Data in the EU Policy Context – Study*. doi:10.2864/97426

Fan, S., Lau, R. Y. & Zhao, J. L. (2015). *Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix.* Big Data Research, 2(1), pp. 28-32. doi:10.1016/j.bdr.2015.02.006

Firat, A., Kutucuoglu, K.Y., Saltik, I. & Tuncel, O. (2013). *Consumption, Consumer Culture and Consumer Society*. Journal of Community Positive Practices, issue 1, p. 182-203.

Flaherty, D. (1999). Visions of Privacy: Past, Present and Future. In Bennett, C. J. & Grant, R. A. (Eds.), *Visions of Privacy* (pp. 19-38). University of Toronto Press.

Floridi, L. & Taddeo, M. (2016). *What is data ethics? Philosophical transactions.* Series A, Mathematical, physical, and engineering sciences, 374(2083), doi:10.1098/rsta.2016.0360

Garcia-Rivadulla, S. (2016). *Personalization vs. Privacy: An Inevitable Trade-off?* IFLA Journal, 42(3), pp. 227-238. doi:10.1177/0340035216662890

GDPR.Eu (2018). *What is GDPR, the EU's New Data Protection Law?* Retrieved November 26, 2019 from: https://gdpr.eu/what-is-gdpr/

Gafney, D. & Puschmann, C. (2014). Data Collection on Twitter. In Weller, K., Bruns, A., Burgess, J., Puschmann, C., Mahrt, M. (Eds.), *Twitter and Society* (pp. 55-67). New York: P. Lang

Ghauri, P. N., Grønhaug, K., Kristianslund, I. & Kristiandslund, I. (1995). *Research Methods in Business Studies: A Practical Guide*. New York: Prentice Hall.

Grandhi, B., Patwa, N., & Saleem, K. (2017). *Data-driven Marketing for Growth and Profitability*. 10th Annual Conference of the EuroMed Academy of Business, 675-694.

Hakala, S. & Vesa, J. (2013). Verkkokeskustelut ja Sisällön erittely. In Laaksonen, S., Matikainen, J. & Tikka, M. (Eds.), *Otteita verkosta: Verkon ja sosiaalisen median tutkimusmenetelmät* (pp. 216-244). Tampere: Vastapaino.

Hesse-Biber, S. N. (2010). *Mixed Methods Research: Merging Theory with Practice.* New York: Guilford Press.

Hox, J. J., & Boeije, H. R. (2005). *Data Collection, Primary Versus Secondary.*

Hyry, J. (2019). *The Use of Digital Services. Total Level Report.* Retrieved April 29, 2020 from: https://www.sitra.fi/en/publications/use-digital-services/

Jackson, J. (2018). *A Perfect Fit: Personalization Versus Privacy.* Journal of Organizational Psychology, 18(4), pp. 33-42. doi:10.33423/jop.v18i4.83

Jamnik, A. (2011). *The Question of Ethical Decision in Marketing and Ethics.* Cultura económica, 80, pp. 41-53.

Janeček, V. (2018). *Ownership of Personal Data in the Internet of Things.* Computer law & security review, 34(5), 1039-1052.

Jiménez-Zarco, A. I., Rospigliosi, A., Martínez-Ruiz, M. P., & Izquierdo-Yusta, A. (2019). Marketing 4.0: Enhancing Consumer-Brand Engagement Through Big Data Analysis. In *Web Services: Concepts, Methodologies, Tools, and Applications* (pp. 2172-2195). IGI Global.

Kotler, P., Armstrong, G. & Parment, A. (2016). *Principles of Marketing* (Scandinavian edition. 2nd edition.). Harlow, England: Pearson.

Kotler, P., Kartajaya, H., & Setiawan, I. (2017). *Marketing 4.0: Moving from Traditional to Digital.* Wiley.

Kotler, P. & Keller, K. L. (2011). *Marketing Management*. (14th ed.) Prentice Hall.

Kumar, V., Chattaraman, V., Neghina, C., Skiera, B., Aksoy, L., Buoye, A. & Henseler, J. (2013). *Data-driven Services Marketing in a Connected World.* Journal of Service Management, 24(3), pp. 330-352. doi:10.1108/09564231311327021

König, M. (2012). *An Investigation on Contemporary Consumer Resistance: How Web 2.0 Makes Consumers Powerful*. Diplomica Verlag, 2012.

Laczniak, G. R. & Murphy, P. E. (2019). *The Role of Normative Marketing Ethics.* Journal of Business Research, 95, pp. 401-407. doi:10.1016/j.jbusres.2018.07.036

Laaksonen, S. & Matikainen, J. (2013). Tutkimuskohteena Vuorovaikutus ja Keskustelu Verkossa. In Laaksonen, S., Matikainen, J. & Tikka, M. (Eds.), *Otteita verkosta: Verkon ja sosiaalisen median tutkimusmenetelmät* (pp. 193-215). Tampere: Vastapaino.

Lee, J. & Jin, C. (2019). *The Role of Ethical Marketing Issues in Consumer-Brand Relationship.* Sustainability, 11(23), p. 6536. doi:10.3390/su11236536

Marr, B. (2016). *Big Data in Practice*. Wiley.

Martin, K. & Murphy, P. (2017). *The Role of Data Privacy in Marketing.* Journal of the Academy of Marketing Science, 45(2), pp. 135-155. doi:10.1007/ s11747-016-0495-4

Mazur, E. (2010). Collecting Data from Social Networking Sites and Blogs. In Gosling, S. & Johnson, J. A. (Eds.), *Advanced methods for Conducting Online Behavioral Research.* Washington: American Psychological Association.

McDonald, M. (2011). Market Segmentation. In Kourdi, J. (ed), CIM Kaihatsu Kenkyukai Staff, Kourdi, J. & The CIM. *Marketing Century: How Marketing Drives Business and Shapes Society* (pp.27-50). Wiley.

Murphy, P. E. (2017). *Research in Marketing Ethics: Continuing and Emerging Themes.* Recherche et Applications en Marketing (English Edition), 32(3), pp. 84-89. doi:10.1177/2051570717701414

Murphy, P. E., Laczniak, G. R. & Wood, G. (2007). *An Ethical Basis for Relationship Marketing: A Virtue Ethics Perspective.* European Journal of Marketing, 41(1/2), pp. 37-57. doi:10.1108/03090560710718102

Nadler, A. & Mcguigan, L. (2018). *An Impulse to Exploit: The Behavioral Turn in Data-driven Marketing.* Critical Studies in Media Communication, 35(2), p. 151. doi:10.1080/15295036.2017.1387279

Nantel, J. & Weeks, W. A. (1996). *Marketing Ethics: Is There More to it Than the Utilitarian Approach?* European Journal of Marketing, 30(5), pp. 9-19. doi:10.1108/03090569610118713

Nill, A. & Aalberts, R. J. (2014). *Legal and Ethical Challenges of Online Behavioral Targeting in Advertising.* Journal of Current Issues & Research in Advertising, 35(2), pp. 126-146. doi:10.1080/10641734.2014.899529

Nupponen, M. (2019). *Sitra's Recent Survey Probes the Secret Life of Data Collected About Us.* Retrieved February 27, 2020 from: https://www.sitra.fi/en/news/ sitras-recent-survey-probes-the-secret-life-of-data-collected-about-us/

Omnicore (2020). *Twitter by the Numbers: Stats, Demographics & Fun Facts.* Retrieved May 16, 2020 from: https://www.omnicoreagency.com/twitter-statistics/

Painter-Morland, M. (2004). *Business Ethics as Practice.*

Paliwal, M. (2006). *Business Ethics.* New Age International.

Parıltı , N., Külter Demirgüneş , B. & Özsaçmacı , B. (2015). *Ethical Issues in Marketing: An Application for Understanding Ethical Decision Making.* İktisadi ve İdari Bilimler Dergisi

Peppers, D. & Rogers, M. (2011). Strategic Marketing. In Kourdi, J. (ed), CIM Kaihatsu Kenkyukai Staff, Kourdi, J. & The CIM. *Marketing Century: How Marketing Drives Business and Shapes Society* (pp.19-26). Wiley.

Petersen, L. B., Nash, C. & Person, R. (2014). *Connect: How to Use Data and Experience Marketing to Create Lifetime Customers.* Wiley.

Reinecke, J. & Ansari, S. (2015). *What Is a "Fair" Price? Ethics as Sensemaking.* Organization Science, 26(3), p. 867. doi:10.1287/orsc.2015.0968

Richards, N.M. & King, J.H. (2013). *Three Paradoxes of Big Data.* 66 Stanford Law Review Online 41 (2013).

Richards, N. M. & King, J. H. (2014). *Big Data Ethics.* Wake Forest Law Review, 49, pp. 393-1545.

Richterich, A. (2018). *The Big Data Agenda: Data Ethics and Critical Data Studies.* University of Westminster Press.

Saif, H., Fernández, M., He, Y., & Alani, H. (2014). *On Stopwords, Filtering and Data Sparsity for Sentiment Analysis of Twitter.*

Sagiroglu, S., & Sinanc, D. (2013). *Big data: A Review*. 2013 International Conference on Collaboration Technologies and Systems (CTS), 42-47.

Schneider, M. J., Jagpal, S., Gupta, S., Li, S. & Yu, Y. (2017). *Protecting Customer Privacy When Marketing with Second-party Data.* International Journal of Research in Marketing, 34(3), pp. 593-603. doi:10.1016/j.ijresmar.2017.02.003

Schreier, M. (2012). *Qualitative Content Analysis in Practice*. Los Angeles: Sage.

Sihem, B. (2013). *Marketing Mix- An Area of Unethical Practices?* British Journal of Marketing Studies Vol.1, No.4, pp. 20-28, December 2013.

Silverman, D. (2013). *Doing Qualitative Research* (Fourth edition). Los Angeles, CA: Sage.

Singh, A. & Mishra, N. (2018). *Ethical Theory & Business. A study based on Utilitarianism and Kantianism*. 10.30546/2523-4331.2018.2.1.97.

Smith, A., Ayanian, J., Covinsky, K., Landon, B., McCarthy, E., Wee, C. & Steinman, M. (2011). *Conducting High-Value Secondary Dataset Analysis: An Introductory Guide and Resources.* Journal of General Internal Medicine, 26(8), pp. 920-929. doi:10.1007/s11606-010-1621-5

Solomon, M. R., Bamossy, G. J., Hogg, M. K. & Askegaard, S. (2016). *Consumer Behaviour: A European Perspective* (Sixth edition.). Harlow: Pearson Education Limited.

Stieglitz, S., Mirbabaie, M., Ross, B. & Neuberger, C. (2018). *Social Media Analytics – Challenges in Topic Discovery, Data collection, and Data Preparation.* International Journal of Information Management, 39, pp. 156-168. doi:10.1016/ j.ijinfomgt.2017.12.002

Strycharz, J. & Smit, E. (2019). *Contrasting Perspectives – Practitioner's Viewpoint on Personalised Marketing Communication.* European Journal of Marketing, 53(4), pp. 635-660. doi:10.1108/EJM-11-2017-0896

Tene, O., & Polonetsky, J. (2012). *Big data for All: Privacy and User Control in the Age of Analytics.* Nw. J. Tech. & Intell. Prop., 11, xxvii.

Thelwall, M. (2014). *Sentiment Analysis and Time Series with Twitter.* In Weller, K., Bruns, A., Burgess, J., Puschmann, C., Mahrt, M. & . (Eds.), *Twitter and society* (pp. 83-95). New York: P. Lang

Thompson, C. B. (2009). *Descriptive Data Analysis*. Air medical journal, 28(2), 56-59.

Turow, J. (2017). Personalization. In Ouellette, L. G. & Gray, J. (Eds.), *Keywords for Media Studies* (pp. 135-137). New York: NYU Press.

Van Bommel, E., Edelman, D., & Ungerman, K. (2014). *Digitizing the Consumer Decision Journey*. McKinsey Quarterly, (June). Retrieved May 1, 2020 from: https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/digitizing-the-consumer-decision-journey

Walker, K. L. & Moran, N. (2019). *Consumer Information for Data-Driven Decision Making: Teaching Socially Responsible Use of Data.* Journal of Marketing Education, 41(2), pp. 109-126. doi:10.1177/0273475318813176

Westin, A. F. (2003). *Social and Political Dimensions of Privacy.* Journal of Social Issues, 59(2), pp. 431-453. doi:10.1111/1540-4560.00072

Yarrow, K. (2014). *Decoding the New Consumer Mind: How and Why We Shop and Buy*. Jossey-Bass.

Young, L. (2011). *The Marketer's Handbook: Reassessing Marketing Techniques for Modern Business*. John Wiley & Sons (UK).

Appendices

Appendix 1. Questions for the “Use of Digital Services” Survey

How old are you?

Are you...?

- Man
- Woman
- Other
- Do not wish to say

In which region do you live?

"This survey is about use of digital services and attitude towards data privacy. We wish to find out how people feel about the potential use of their personal data from data protection and privacy's point of view.

The results of the survey will be utilized to accelerate the development of human-driven data economy and to give input into public debate. In addition, we will provide businesses with information on the values of individuals, as well as, on citizens' attitude towards the data economy."

What rights do you believe you have concerning the personal data the service provider has gathered of you?

You can choose several options.

- Right to have access to any personal data the service provider has collected of me
- Right to have any personal data collected of me corrected or updated
- Right to have any personal data collected of me deleted

- Right to move any personal data from one system to an other
- Right to know how and for what purpose my personal data is used
- Right to get a notification from the service provider when my personal data is being sold/handed over to a third party
- Right to restrict handling of my personal data, especially where automated decision making is conserned
- Right to be informed when there is a service provider data breach involving my data
- I do not believe I have any rights
- I do not know

How do you feel about the terms and conditions of different applications/services?

Choose the alternative that best suits your opinion

- I always read them and then accept
- I usually read them and then accept
- I sometimes read them and then accept
- I accept them without reading
- I only use applications/services whose terms and conditions I can accept
- I don't use applications/services that require accepting terms and conditions
- None of the above

How well do you think you understand the terms and conditions of different applications/services?

- Very well
- Fairly well
- Somewhat well
- Not very well
- Not well at all
- I do not know
-

Have you changed the settings for different applications / services to meet your usage needs (eg location, privacy)?

- I have changed the privacy settings on all the applications/services I use
- I have changed the settings on most applications/services I use
- I have changed the settings on some applications/services I use
- I have not changed any settings
- I did not know it is possible to change the settings
- I do not know

Why haven´t you changed privacy settings on applications/services?

Choose the most important reason for you.

- I do not want to spend time on it
- It is difficult to change the settings
- I do not think it is important
- I do not know how to change the settings
- I do not believe it has any effect
- Some other reason
- I do not know

How have recent news of data leakages in applications/services (e.g. Facebook) affected your use of different services?

You can choose several options.

- I have stopped using some services
- I have reduced using some services
- I have adjusted my privacy settings in some services
- I have adjusted my privacy settings in all services
- The news have not had any effect
- I do not know

"To what extent do you agree or disagree with the following statement?

Lack of trust towards service providers prevents me from using digital services"

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree
- I do not know

How has the EU's General Data Protection Regulation (GDPR), enforced on 25 May 2018, affected your online behaviour?

You can choose several options.

- I have accepted the new terms and conditions that service providers emailed me without reading them thoroughly
- I have requested access to my personal information from a service provider
- I have requested how my information is used, from a service provider
- I have requested a service provider to move my information to another service
- I have requested a service provider to correct some of the information they are holding on me
- I have requested a service provider to delete all the information they are holding on me
- I have forbidden the automated use of my personal data from a service provider
- I have had to stop using an online service because it was no longer offered to me
- (e.g. service was not available to European customers)
- It has not affected my behaviour in any way
- I do not know

"Accumulated consumer data is used for multiple business purposes.

A service provider can for example use accumulated data for:

- for producing, developing and targetting their service

- selling /handling over user data to third parties (e.g. advertisers)

How much do the following increase your trust towards a service provider when your personal data is used for producing, developing and targetting the application/service you use?"

Terms and conditions are clear and quickly absorbed

Service informs clearly how my data is going to be used

I can get a clarification on what information the service provider has about me

I can accept or decline the selling of my data to a third party

I can delete all the personal data that the service provider has collected of me

I get paid in exchange for giving permission to my personal data

I am offered extra service in exchange for permitting the use of my personal data

- Very much
- Somewhat
- Not so much
- Not at all
- I do not know

How much do the following features increase your trust towards a service provider selling your personal data to third parties (e.g. advertising companies)?

Terms and conditions are clear and quickly absorbed

Service informs clearly how my data is going to be used

I can get a clarification on what information the service provider has about me

I can accept or forbid the selling of my data to a third party

I can delete all the personal data that the service provider has collected of me

I get paid in exchange for giving permission to my personal data

I am offered extra service in exchange for permitting the use of my personal data

I can personally influence in to whom my data will be sold

- Very much
- Somewhat
- Not so much
- Not at all
- I do not know

Under what conditions would you be willing to provide or allow access to the following data about yourself or your family to a service provider?

Information about your health or heredity

Personal data or data about your personal history

Information about your values or belief

Information about your consumption habits or your purchases

Your spatial data, devices you use or the way you use devices

Information about your wealth and spending

Data is used for purposes of public interest (eg health care, environmental protection)

- Information is used for scientific research
- I would be offered extra services or individual service
- I would be paid for it
- I am not willing to provide or allow access under any conditions
- I do not know

In your opinion, what would be the best way to allow a service provider to collect and handle the data of you?

- Give consent separately to each service provider
- The service provider makes the decision on my behalf according to the conditions I have given
- The service provider would use ready customer profiles to make the decision on my behalf

- The service provider would make the desicion on my behalf based on the behaviour of people with similar profile
- The service provider would make the desicion on my behalf based on my past behaviour
- I do not know

How important do you feel the following features of digital applications / services are?

The purpose of the data collected of me is clearly and transparently reported in the service / application

It is free of charge

It is easy to use

Service/application is personalised for me based on my previous usage

The service provider is reliable

The service is secure

- Very important
- Somewhat important
- Not so important
- Not at all important
- I do not know

How should information management and utilisation develop compared to the current?

Choose the option that you feel is the best.

- Users should manage their personal data more actively and more consciously
- Authorities should regulate the use of data more strongly
- Interest groups and trustees specialising in personal data should protect people's interests
- The organisations collecting and managing data should take more responsibility for its ethical use
- I do not know

Would you use a single application where you could manage all data collected of you in different applications/services?

- Yes
- No
- I do not know

What kind of features would you expect / want from a single app, that would allow you to manage all data collected of you by different applications/services?

- It is easy to use
- It is free of charge
- Safety
- Reliability
- Easy access
- Covers all the services I use
- Other, what?
- I do not know

Service providers collect a lot of data of you. In your opinion, how should this data be managed for you to feel that it is fair for you?

- I do not know

Consumer goods have the Fair Trade label. Do you find a similar label important for services that use data fairly?

- Very important
- Somewhat important
- Not so important
- Not at all important
- I do not know

If services that use personal data would have a "fair data" label, what would be the minimum requirement for it?

List below words that best describe a "fair data" service

- I do not know

Do you live in...?

- City
- Town/Urban area
- Countryside
- I do not know

What is your highest level of education completed?

- Compulsory education
- Vocational qualification, vocational college
- Matriculation/A-levels
- Higher education level vocational qualification
- Advanced higher education level vocational qualification, bachelor’s degree
- Masters degree or higher
- Other
- I do not know

Which of the following best corresponds to your occupational group or situation?

- Managerial position working for someone else
- Senior white collar
- Junior white collar
- Worker
- Self-employed or sole trader
- Farmer
- Unemployed
- At school or student
- Pensioner

- At-home mother, at-home father
- Other
- I do not know