



Lut School of Business and Management

Bachelor's thesis

International business

Consumer Big data and the influences of General Data Protection Regulation

Kuluttaja Big data ja uuden tietosuoja-asetuksen vaikutukset

17.5.2018

Author: Joonas Heikkilä
Instructor: Anssi Tarkiainen

Table of contents

1. INTRODUCTION	1
1.1 STRUCTURE OF RESEARCH	2
1.2 RESEARCH PROBLEM, OBJECTIVES AND LIMITATIONS	2
1.3 THEORETICAL FRAMEWORK	3
1.4 LITERATURE REVIEW.....	4
1.5 RESEARCH METHODOLOGY AND LITERATURE.....	5
1.6 KEY WORDS.....	6
2. BIG DATA.....	7
2.1 BIG DATA 3 V'S.....	7
2.1.1 Volume.....	8
2.1.2 Variety	8
2.1.3 Velocity	9
2.1.4 Veracity, Value and Variability	9
3. CONSUMER BEHAVIOUR IN DIGITAL ENVIRONMENT	11
3.1 PROBLEM RECOGNITION	11
3.2 SEARCH.....	12
3.3 ALTERNATIVE EVALUATION.....	13
3.4 PURCHASE BEHAVIOUR.....	14
3.5 CONSUMPTION	15
3.6 POST-PURCHASE EVALUATION	16
3.7 POST-PURCHASE ENGAGEMENT	16
3.8 OVERVIEW.....	17
4. CONSUMER DECISION-MAKING PROCESS IN CASE-COMPANIES	19
4.1 COMPANY AND INTERVIEWEE INTRODUCTION.....	19
4.2 CONSUMER DECISION-MAKING PROCESS	20
5. GENERAL DATA PROTECTION REGULATION	23
5.1 INFLUENCES ON COMPANIES AND CONSUMER BEHAVIOR	23
5.2 THREATS AND OPPORTUNITIES.....	25
6. CONCLUSIONS	28
7. REFERENCES	32

Attachments

Attachment 1. Interview questions

Table of figures

Figure 1. The decision-making process of a consumer..... 4

Figure 2. The characteristics of consumer big data..... 8

Figure 3. Overview of consumer’s decision-making process 18

Abstract

Author:	Joonas Heikkilä
Title:	Consumer Big data and influences of GDPR
Faculty:	LUT School of Business and Management
Degree program:	Business Administration/ International Business
Instructor:	Anssi Tarkiainen
Keywords:	Big data, consumer decision-making process, general data protection regulation (GDPR)

The objective of this thesis is to examine consumer behaviour through decision-making process and understand what kind of consumer big data it generates. Moreover, the thesis aims to seek understanding how the new general data protection regulation will influence on the company's data collection as well as consumer behaviour. At the end of the thesis, opportunities and challenges of GDPR will be assessed. The theory of this study is based on previous scientific literature of the subject and the research is carried out as a qualitative research using case companies. The case companies of this thesis are MOI-mobiili and Zeeland Family. In addition to these case companies, an interview has been carried out with an expert from the department of legislation initiatives of Finnish Ministry of Justice to get profound understanding from the new regulation.

The study of consumer behavior has been an investigated field for ages, but is now being changed since the behavior of the consumers is more and more shifting into digital environment, especially into social media. Traditional decision-making process is more and more visible online and in order to retain competitive advantage, companies need to understand the value of consumer data which is generated from it. Now that the cost of storing data has dropped below the cost of deleting it, companies have started to put more effort in gathering consumer data to gain better understanding from consumers. Scholars are talking about an era of big data which means that data are increasingly available everywhere.

The results of the research show that the study of consumer decision making process is a lifeline to every company that wishes to succeed and that the theoretical model can be applied into a practice. Only thing that varies, is that the companies put more weight on different phases of consumer's decision making-process. The new data protection regulation brings new obligations to the companies as well as rights to consumers. Only the time will show how the regulation will influence on these, however some of the influences can be seen already.

Tiivistelmä

Tekijä:	Joonas Heikkilä
Tutkielman nimi:	Kuluttaja Big data ja GDPR:n vaikutukset
Akateeminen yksikkö:	LUT School of Business and Management
Koulutusohjelma:	Kauppätieteet / Kansainvälinen liiketoiminta
Ohjaaja:	Anssi Tarkiainen
Avainsanat:	Big data, kuluttajan päätöksentekoprosessi, general data protection regulation (GDPR)

Tämän tutkimuksen tavoitteena on tutkia kuluttajakäyttäytymistä kuluttajan päätöksentekoprosessin kautta ja saavuttaa ymmärrystä minkälaista kuluttaja big dataa tämä tuottaa. Lisäksi, työ pyrkii löytämään ymmärrystä kuinka uusi henkilötietoihin liittyvä tietosuoja-asetus vaikuttaa yritysten datan keruuseen sekä kuluttajakäyttämiseen. Lopuksi työssä otetaan katsaus tulevaisuuteen, millaisia mahdollisuuksia ja haasteita tietosuoja-asetus asettaa. Tutkielman teoriaosuus perustuu aikaisempaan kirjallisuuteen ja tutkimus on toteutettu laadullisena tutkimuksena. Tutkimuksen kohdeyrityksinä toimii kaksi suomalaista yritystä, puhelinoperaattori MOI-mobiili ja markkinointitoimisto Zeeland Family. Näiden lisäksi Suomen oikeusministeriön lainvalmisteluosaston lainsäädäntöneuvoksen asiantuntemusta on saatu työhön tietosuoja-asetuksen ymmärtämiseksi.

Kuluttajakäyttäytymisen tutkimus on ala, jota on tutkittu vuosikymmenet ja kirjallisuutta aiheesta löytyy loputtomasti. Teknologian ja sosiaalisen median nopean kehityksen myötä on myös kuluttajakäyttäytyminen muuttunut ja siirtynyt enemmän digitaaliseen ympäristöön, etenkin sosiaaliseen mediaan. Nyt, kun tiedon keräämisestä ja säilönnästä on tullut yhä halvempaa, kuin poistaa ylimääräinen tieto, ovat yritykset siirtyneet panostamaan tiedon keruuseen ymmärtääkseen kuluttajia paremmin. Tutkijat puhuvat big datan aikakaudesta.

Tutkimustulokset osoittavat, että kuluttajakäyttäytymisen tutkimus on elinehto yrityksille, jotka haluavat menestyä ja, että teoreettinen malli on yhdistettävissä käytäntöön. Ainut asia, joka eroaa, on kuluttajan päätöksentekoprosessin vaiheiden painotukset eri yritysten välillä. Uuden tietosuoja-asetuksen tullessa voimaan, tuo se uusia vaatimuksia yrityksille, sekä oikeuksia yksityisille henkilöille. Joitakin vaikutuksia on jo nyt nähtävissä, mutta ilmiön todelliset vaikutukset kuluttajakäyttämiseen on vasta nähtävissä ajan myötä

1. INTRODUCTION

Currently, the international population is around 7,2 billion and over 2 billion people of the population are connected online on one way or another. Moreover, over 5 billion individuals own mobile phones which means people are continually connected online and share information at the speed of light. (Khan, N. Yaqoob, I. Abaker, I. Hashem, T. Inayat, Z. Mahmoud Ali, W. K. Alma, M. Shiraz, M. Gani, A. (2014), 1)

Data has been collected since the beginning of modern day business. Not until later - on this decade, the phenomenon called big data has been acknowledged. Nowadays big data is everywhere in every shapes and forms. Along sustainable development and innovation, big data has become one of the biggest business trends on this decade. IBM Tech Trends Report (2011) recognizes business analytics as one of the major technology trends in 2010's based on 4000 information technology professionals (Hsinchun, C. Chiang, R. Sotrey. V. 2012, 1165). Whereas, in prediction of IBM Tech Trends Report (2017) professionals claim that big data will continue to grow as organizations continue to become aware of these enormous data sets and insights that can be extract to create business value (Reddy, T. 2017).

It is hard to name the exact time when big data came to awareness in businesses. Some say that the era of big data began at the point in which the cost of storing data dropped below the cost of deleting it (Hofacker, C. F. Malthouse, E. C. Sultan, F. 2015, 89). Meanwhile, others argue it began in lunch-table conversations at Silicon Graphics Inc in the mid-1990's. Despite the reference to the mid-nineties, the term became widespread as recently as in 2011 when IBM and other leading technology companies started to invest on data-analytics to create niche-markets (Gandomi, Amir & Haider, M. 2014, 137). Generally, the opinion of the starting point lays on rapid growth of social media, as well as online and mobile devices where all actions can be recorded (Hofacker, C. F. et al. 2015, 89).

While the world is transforming more into an information society, companies need to step up their game and master data control through utilizing it, in order to gain and retain competitive advantage. From consumer perspective, gathering information from company products and its alternatives is becoming even easier. For this reason, acknowledging opportunities that big data has to offer is essential in today's business world. Nowadays, every movement of a consumer in Internet is recorded, collected, stored and turned into company advantage. Scholars are talking about big data

revolution, which will lead to new ways of understanding consumer behaviour and new ways of marketing techniques.

This thesis will introduce big data as a phenomenon and explores consumer behaviour in social media, and finally in the empirical part deals with the fact how the new data protection regulation will influence on the consumer behaviour. The goal is to find out how companies collect big data based on consumer behaviour and what kind of influence GDPR has on this. The research problem is viewed through consumer decision-making process on social media, which will be compared with several interviews with different companies.

1.1 Structure of Research

The structure of the thesis consists five main chapters, which are divided into a theoretical and an empirical part. The first chapter includes the introduction, which presents the phenomenon and gives a vision what the thesis is about. The first chapter also presents the purpose of this study, objectives and limitations. Lastly in the first chapter, theoretical framework is introduced. After chapter one, the thesis moves on to the theoretical part, where big data is defined through the 3 V's model. Additional three v's are presented to describe big data processes as well. In the third chapter the research concentrates on consumer behaviour in digital environment online, which is presented through consumer decision-making process. The fourth and the fifth chapter concerns the empirical part of the thesis, which is built by three interviews with different experts.

1.2 Research Problem, Objectives and Limitations

The research goal is to find out what the influences of GDPR on consumer data collection are. The purpose of this thesis is to examine the consumer behaviour in digital environment, especially the decision-making process, and evaluate how companies collect consumer data. Theoretical framework of this thesis is built around on the consumer decision-making process, which is given by Blackwell, R.D. Miniard, P.W. and Engel, J.F. (2005). Since the outline of the thesis is on digital environment, update for the framework is needed. Conclusively/updated framework is adapted from Hofacker, C.H. et al (2015) and is suitable for social media analysis. To get a better understanding about consumer big data and influence of GDPR, theoretical framework will be compared to several interviews with

case companies. The goal of this thesis, is to describe consumer big data as a phenomenon, and evaluate the effects of GDPR.

The main research question:

- *“How General Data Protection Regulation could influence on consumer behaviour?”*

Other research questions:

- *“What is consumer big data?”*
- *“How GDPR will influence companies’ data collection?”*
- *“What kind of opportunities GDPR might bring in consumer behaviour?”*

The objective of the thesis is to introduce consumer big data as a phenomenon in the light of present day and evaluate the influences of GDPR. Taking into account the extent of subject, the thesis is narrowed down to digital environment, more precisely to social media. In addition to present day examination, the thesis will review future directions, since the effects of GDPR are difficult to be detected immediately. The research has been assembled from various scientific articles, which have been published on this decade.

1.3 Theoretical Framework

The purpose of this thesis is to introduce consumer big data as a phenomenon and study what kind of consumer data companies collect through consumer decision-making process and finally evaluate how General Data Protection Regulation will influence on that. Theoretical framework of the thesis follows the steps of consumer decision-making process. The original structure is given by Blackwell, R.D. Miniard, P.W. and Engel, J.F. (2005), but for the thesis it will be modified to be more suitable for digital environment analysis, hence it follows the steps that are represented in article by Hofacker, C.F. et al., (2015).

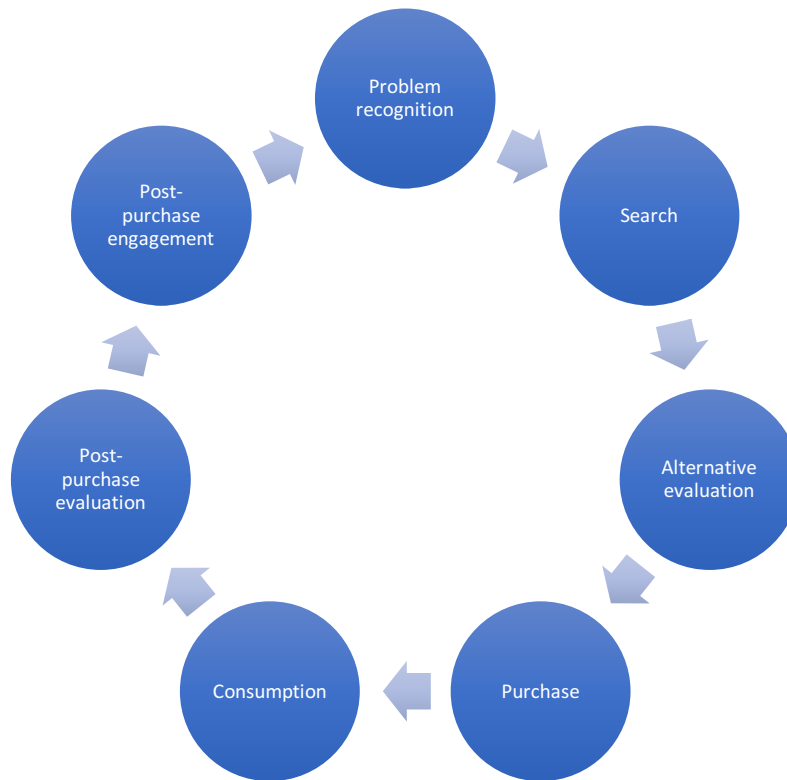


Figure 1. The decision-making process of a consumer

We live in the world, where more and more consumer behaviour, and consumption itself, takes place or is immediately visible in digital environment, especially in social media (Hofacker, C.F. et al., 2015, 90). That is the reason why this thesis concentrates on consumer behaviour in digital environment and its relation to data collection. The steps of consumer decision-making process are problem recognition, search, alternative evaluation, purchase behaviour, consumption, post-purchase evaluation and post-purchase engagement. (Hofacker, C.F. et al. 2015, 90)

In the thesis, Big Data is defined through traditional 3 V's model (volume, variety and velocity). With the 3 V's research gives clear view of Big Data and introduces ways of gathering data in social media. The thesis also describes Big Data through additional three v's (veracity, value and variability).

1.4 Literature Review

The study of consumer behaviour has been investigated for decades. It has its roots in psychology, sociology, anthropology, marketing and economics. In the context of business, the aim is onto understand the actions taken by consumer during a purchasing and what kind of factors determines

these decisions. Dinu, D. Stoika, I. Radu, A. V. 2016, 248) As the big data has raised more attention, the literature is focused more on the behaviour happening online.

Consumer decision-making process has become one of the used framework for study consumer's behaviour although the amount of the phases in the process varies a bit. The classical decision-making process starts from need recognition and continues to search, alternative evaluation, purchase decision and finally to post-purchase evaluation. As the time has passed, new phases have been introduced, such as consumption and post-purchase engagement (Hofacker, C.F. et.al. 2015, 89).

What comes to consumer decision-making and big data related to it, the literature is quite fragmented and there is not a lot about the subject. According to Erevelles, C.F. et.al (2016) only in the recent years, when consumers have become an incessant generator of huge amount of big data, has the research focused more into it. According to Hofacker, C.F et.al. (2015) the research focus shifted towards consumer behaviour and big data within it when information technology and social media users increased rapidly.

1.5 Research Methodology and Literature

This research is conducted by using qualitative methods which enables a descriptive and in-depth analysis. The theoretical part of the thesis is assembled from previous scientific researches of the subject which are carefully assessed to be suitable to this thesis. Moreover, the concentration has been on literature that has been published on this decade.

The empirical part is built based on three interviews with experts from different fields. The aim is to find correlations between practise and theory and to understand the differences between them. Interviews have been carried out by e-mail. Ensuring the reliability of interview answers, clarified questions have been presented and the interviews have been notified that the thesis will be publicly available. The interview questions have been modified a bit, since the interviewees are experts from different fields.

1.6 Key Words

Big data as a term has not been established yet and it has as many definitions as there is research related to it. Gartner It Glossary (2017a) defines big data as a high-volume, high velocity high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight decision making, and process automation. Usually big data is described through 3 V's which are volume, velocity and variety. Volume of the data that is available everywhere and anytime while velocity refers to the fast rate that data is received and acted upon. (Oracle 2016, 4) The third dimension of big data is variety, which refers to the structural diversity in a dataset. (Gandomi, A. & Murtaza, H. 2014, 138)

Consumer decision-making process is usually presented as a process from consumption need to purchase. In this thesis, the decision-making process is presented as a seven-step process and is updated to be more suitable for online environment. The steps are problem recognition, search, alternative evaluation, purchase, consumption, post-purchase evaluation, post-purchase engagement. (Hofacker, C.F. et al., 2015, 90)

General data protection regulation (GDPR) is an upcoming reform to the EU's current data regulation. With the reform, the goal of European Union has been to create an update to the new data regulation and make it strong, cohesive and comprehensive data protection framework for European markets. The regulation will be applied throughout EU, which significantly unifies the personal data process. It concerns every company operating in EU and everything related to individual's personal data. (Tietosuoja, 2018) With the reform of the current regulation, requirements are on more careful data collection and protection. As it puts certain requirements for companies it also gives new rights to consumers. (Elikeinoelämän keskusliitto, 2018)

2. BIG DATA

If companies have always collected data, how does big data differ from regular data which has been collected already before the phenomenon? Firstly, the major difference is the rapid growth of Internet, especially social media and IoT which has led people being connected online continually. Khan, N., et al. (2014) present that over 2 billion people worldwide are connected to the Internet, whereas looking at the number of users in Facebook (around 2 billion), thus it can be deduced that almost everyone connected to the Internet have a Facebook profile (Statista 2017a). It has been estimated that around 50 billion devices are expected to be connected to the Internet by 2020 and through this predicted data production will be 44 times greater than in 2009 (Khan, N. et al. 2014, 1).

Defining big data is not an easy task, because the phenomenon has not been generally established yet. Commonly, big data can be defined as a huge amount of data that can be analysed to extract patterns, information, insights and trends. From a theoretical perspective, the best way to define big data is through the 3 v's, which refers to volume, velocity and variety of primary data. Gartner It Glossary (2017a) describes big data with high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation.

In this thesis, the 3 V's model is used to describe big data characteristics. Although the 3 v's are highly used to describe big data, additional two v's are necessary to have deeper understanding. These three v's are veracity, variability and value which concentrate in collecting, analyzing and extracting insights from the primary data (Erevelles, S. et al. 2015, 898). These three v's are presented more specifically in the end of chapter two.

With the first three v's, the goal is to give a clear vision from the characteristics of big data, and through the last three v's the purpose is to describe the data processes associated with it.

2.1 Big Data 3 V's

The opinions of big data characteristics are continually changing and new ways of describing big data are raising all the time. Gartner's analyst Doug Laney was one of the first to suggest 3 V's model for

big data management. As introduced in the chapter before, Doug Laney presents that the 3 V's refer to volume, variety and velocity of primary data (Meta Group Inc 2011).

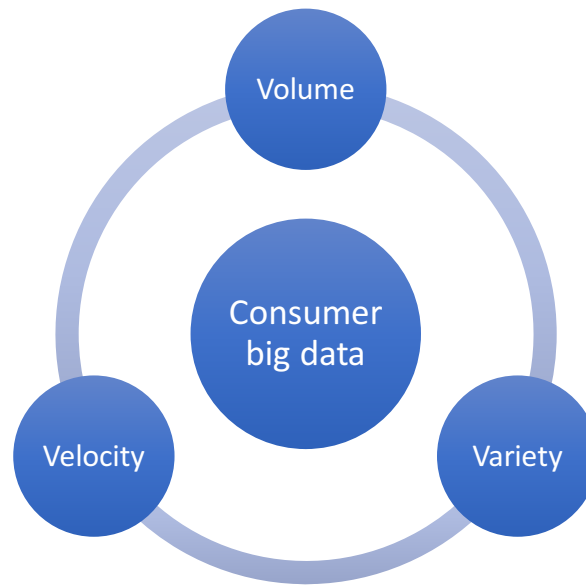


Figure 2. The characteristics of consumer big data

2.1.1 Volume

The first thing that comes up to mind from big data is the size. The size refers to volume of data that is available almost everywhere. It is hard to measure the extent of data worldwide, but the general view is that by 2020 the digital universe is expected to reach 44 zettabytes (Erevelles, S. et al. 2015, 898). Usually, data is measured as in petabytes, exabytes or zettabytes (Erevelles, S. et al. 2015, 898). To clarify terms, one petabyte is equivalent to 20 million traditional filing cabinets of text; putting it in perspective, one of the world's biggest retail chain Walmart is estimated to create 2.5 petabytes of consumer data every hour (Sunil, E., et al. 2015, 898). Whereas one terabyte can contain over 16 million Facebook profile pictures. (Gandomi, A. & Murtaza, H. 2014, 138) Volume can also be quantified by counting records, transactions, tables or files (Russom, P. 2011, 6).

2.1.2 Variety

The second dimension of big data is variety, which refers to the structural diversity in a dataset. Variety is maybe the trickiest one because of the diversity of the data does not always pay high value. With today's advanced technology companies are more capable to collect different types and forms of data, such as structured and unstructured data. Structured data, which constitutes only around 5% of all existing data, comprehends the tabular data in spreadsheets or relational databases. (Gandomi,

A. & Murtaza, H. 2014, 138) Whereas unstructured data or “human information”, which constitutes 80% of the data worldwide and account for 90%. Unstructured data consists for example, photos, e-mail, Facebook data, call-center conversations, website clicks etc. (Nawser, K. et al. 2014, 2)

2.1.3 Velocity

The third dimension of big data is velocity, which refers to the fact that data is constantly being generated at the speed of light. IBM (2017b) defines velocity as an analysis of streaming data. It is the rate at which data arrives at the enterprises and the time that it takes the enterprises to process and understand the data (IBM, 2017b). Whereas Oracle (2016, 4) defines velocity as a fast rate that data is received and perhaps acted upon. With current and real time given evidence, marketers are able to contemplate and create more accurate moves considering marketing activities. Imaging marketing executive who knows at any given time how many items are purchased, which colors customers prefer, which links are clicked on a website or how many likes product has received on Facebook etc.

2.1.4 Veracity, Value and Variability

Additional three v’s are essential to introduce in order to get a deeper understanding of big data characteristics. These three v’s are important considering data collecting, analyzing and extracting insights from Big Data (Ervelles, S. et al. 2015, 898).

Veracity is presented as the fourth dimension of big data to better understand the uncertainty of data world. IBM (2012) demonstrates the uncertainty of data by using weather forecasts as an example, despite having all the data, weather is still uncertain and still company forecasts it. As the world keeps producing data in massive volumes, it is important to resolve which data are reliable and how to extract it. One way to accomplish in this, is to create context around the wanted data, IBM (2012) is speaking about data fusion, where multiple less reliable sources are combined to create a more accurate and useful data point, such as social comments appended to geospatial location information.

Oracle (2016) points out the value of big data’s intrinsic value which is necessary to be discovered. The data masses usually pose low value density, but an organization’s capability to recognize which of the data are valuable and relevant plays a big role. As the amount of the data keeps increasing, it raises challenges how to derive the right data from it.

SAS introduced the fourth dimension for big data, which is variability. Variability describes how a data flows can be highly inconsistent with periodic peaks. (SAS 2017a) Basically, it means that the data quality can vary across seasons which can lead to difficulties in data managing. Especially, trends in social media can create peaks in data that could influence on overall view of the data.

3. CONSUMER BEHAVIOUR IN DIGITAL ENVIRONMENT

Consumer behaviour is a field investigated for decades. It has its roots in psychology, sociology, anthropology, marketing and economics and has always been aimed to understand customer decision-making during a purchasing process, and sort out what kind of choices are associated within it (Dinu, D. et.al. 2016, 248).

The value in the data comes from the ability to harness the data for company usage. As corporates have access to unlimited amount of customer data, they are capable to create more efficient marketing actions for consumers. For example, after buying flight tickets to Buenos Aires online, next time you sign in Facebook you might see AirBnB advertisement for accommodations in Buenos Aires. Another classic example is, where a teenager's parents found out that their daughter is pregnant from coupons she received based on her behaviour in Internet. While writing the thesis, after searching articles about big data, I suddenly started to receive loads of advertisement from Oracle and IBM, which are well-known big data software providers.

These all are based on the fact how consumer behaves in Internet. As consumers are slowly shifting into digital environment and social media - where all the consumption takes place - it is important to harness the data control and figure out how to extract the relevant data from mass.

In this section, consumer behaviour in digital environment is described through consumer decision-making process. The goal is to examine consumer behaviour through decision-making process and assess what kind of consumer big data is derived from it. Hofacker C.F. et al. (2015) framework for social media decision-making process is adapted into this thesis as a framework. The thesis follows the steps of consumer decision-making process, which are problem recognition, search, alternative evaluation, purchase behaviour, consumption, post purchase evaluation and post-purchase engagement. With these steps the thesis strives to describe how consumer behaviour derives valuable data.

3.1 Problem recognition

The first step of the decision-making process is problem recognition. Hofacker, C.F. et al (2015) describes problem solving process as a gap between what the individual has, or has experienced, and

what the individual desires. Khang, K.Z.K. and Benyoucef, M. (2016) think that consumer's decision-making process starts from the need recognition stage, which indicates that the consumer establish consumption needs or becomes aware of certain products or services. Since the goal of consumers is to fulfil their needs, the role of marketing executives is to fulfil the need or on the other hand be a creator of the need.

Past decades, companies have relied on continual stream of new business ideas by internal staff and experts or inventors (Bayus, B.L. 2013, 226). Traditional inside-outside thinking is more and more shifting towards outside-inside monitoring. Recently, there have been a much discussion about social media as a platform of consumer's need recognition as well as company's environment to identify new ideas for products and their improvements. For instance, Bayus, B.L. (2013) speaks about crowdsourcing, how new ideas for products and services can be collected from a large crowd of consumers. Crowdsourcing is a web-enabled system to monitor community and extract new ideas for products and services (Bayus, B.L. 2013, 226). Moreover, Chan K.W et.al. (2015) discuss how customer peer-to-peer interactions in social network with others influence the likelihood of subsequent idea generation.

Since today's social communities appear in social media, it is important for companies to recognize these transaction patten signals among consumers. For example, Apple has only in Facebook and Instagram over 15 million followers (Facebook & Instagram 2018). Previously mentioned social media channels, provide platforms for consumers to connect with each other, discover, share, recommend, rate and like products or services appearing in social medial.

3.2 Search

Consumers have become more rational and less tolerant to traditional marketing communication on a products and services, so they have accessed forums and social networks to search user reviews and video demonstrations on what they desire (Dinu, D. et.al. 2016, 250). So, after the consumer has identified the need, the search process starts the. Whang K.Z.K. et.al. (2016) defines the search stage as a consumer's information searching process for making informed choices. In this phase, consumer searches information about the products or services and finally continues to alternative evaluation. Before the consumer gets there, he or she needs to find out what kind of possibilities there are. Of course, the process of purchase can end in search process, if consumer decides impulsively to buy the

product or service without any extended search. Whether the process ends or continues, company can obtain valuable information of the consumer's behaviour.

Before Internet, consumers had to spend a very long time to searching information from alternatives, while for the companies, data gathering was limited down to what customers really purchased. Nowadays, a company can record all of the search activity on the company's website and its mobile app that will or will not lead to purchase. Such activity on websites and apps as, searched items, clicks, items that have been added on shopping cart or wish list, abandoned and purchased items are extremely valuable for companies. (Hofacker, C.F. et al. 2015, 91)

Serial of clicks in social network are usually closely associated in search phase. The web usage behaviour, or other words clickstream, is defined as the path a consumer takes through one or more websites. It includes within-site information such as the pages visited, time spent on each page, clicks on a page, and between-site information such as the websites visited (Senecal, S. Kalcynski P.J. Nantel, J. 2005, 1601). The list goes on and on, but the game changer thing is to benefit from these actions. Clickstream analysis gives an opportunity to study consumer navigation through Web sites and which produces valuable data. For example, the site where the consumer sees an advertisement of a product is the path he or she takes to get to the advertiser's site and this can generate valuable clickstream data for a company usage.

Powers, T. et.al. (2012) point out rapid growth of mobile phone users and its role in consumer decision-making process. They remarked that shoppers are using mobile devices (phones and tablets) for online activities at many phases in the decision process. Being connected to social media was more common via mobile, whereas access to brand and company's web sites or review sites was more likely to be done by PC's. Taking these notes into consideration, marketing executives should pay attention that the online searching and shopping will most likely to be accessed via mobile devices.

3.3 Alternative Evaluation

The evaluation phase might be the most critical phase from the company view, especially for small company which pleases to compete among big ones. After consumer has searched enough information about alternatives, consumer forms a consideration sets and inferred rules based on navigation sequences (Hofacker, C.F. et al. 2015, 91). Lindsey-Mulkin, J. & Borin, N. (2017) present that rather than passively gathering information provided at company-controlled touch points,

consumers actively seek information from other consumers, and this way their increasingly active information searches tend to lengthen the evaluation stage.

In the evaluation phase, recommendations of others play a big role. Senecal, S. Kalczynski, P.J. and Nantel, J. (2005) emphasize the importance of other's recommendations in consumer's decision-making process and their online shopping behaviour. According to Senecal, S. et.al. (2015), whether a product or service recommendations are available to the consumer, he or she has two options, consult and follow it, or consult and not to follow it. In case the consumer chooses not to consult, he or she relies only on his or her prior knowledge or experience about the product or service and makes the decision based on it. Assuming the product or service is something new, consumers are more willing to consult recommendations and follow based on them. Consumers have multiple sources available to get recommendations, and usually these are known social media platforms that have huge communities built around specific product or service. Recommendations are a great data source known as social media data. Chan, H.K. et.al. (2017) define social media data as a secondary data which is openly accessible and contains a large set of information for no specific research purpose. Taking travellers for example, it is convenient for them to use Facebook or blogs to evaluate travelling information, rather than reliance information of travel agencies and hence, reduces agencies marketing activities (Chan, H.K. 2017, 5028) Moreover, Hofacker, C.F. et.al. (2016) mention shopping cart abandonment, which can signal that the customer might make price comparisons at other sites.

3.4 Purchase Behaviour

After a consumer has assessed the alternatives, the consumer makes the decision to buy or not to buy. Again, from a company's view, it is essential to make the purchase process as easy as possible for the consumer. To achieve in this, companies need to investigate consumer behavior and harness the data within it.

Data sources include online browsing, adding products to shopping carts and wish lists, purchase responses on direct marketing and transactions (Hofacker, C.F. et.al. (2016), 92). In literature, consumer purchase behavior does not attract huge attention as search and alternative evaluation, but is still a prominent factor.

3.5 Consumption

Consuming might be the most interesting part of the process, since consuming is more and more invisible in Internet, especially in social media. Although, we are used to digital consumption, such as Netflix, iTunes and Spotify, the offline consumption is also slowly shifting towards a digital one. For example, an individual that decides to have a dinner in a nice restaurant, checks alternatives with a mobile phone by using Google maps, and before being seated in a restaurant's table, shares his location in Foursquare and adds a notification with a hashtag of the name of the beer he or she is drinking, and eventually at the table takes a picture of the food he or she is having, and shares it in Instagram or Facebook. After having the dinner, the individual tweets about the restaurant, again using hashtags, for friends to see. These are all perfect to illustrate the variety of big data formats, including transactional, location-based and pictorial (Hofacker, C.F. et al., 2015, 92). According to Hofacker, C.F. et al. (2015) the latter one poses major challenges to managers.

Nowadays more and more people are connected to online, in the form of sharing pictures and videos, for example sharing pictures in Instagram of wearing a new pair of Nike's. Hofacker, C.F. et al. (2015) are speaking about the "quantified self" and the "measured life" referring to increasing actions to share daily things in social media. Lindsey-Mullikin, J. & Borin, N. (2017) state that Instagram has currently 400 million active monthly users who share more than 60 million photos every day.

One of the recognized future technology trends is Internet of Things (IoT). It is a new technology paradigm envisaged as a global network of machines and devices capable of interacting with each other (Lee, I. & Lee, K. 2015, 431). IoT brings an unlimited amount of possibilities to the business world, especially into the studies of consumer behavior. Interacting devices already exist but the total potential of the technology has not been utilized yet. IoT has been used for a long time in healthcare, delivery and supply chains, for example, with RFID-tags attached to cargo, suppliers and buyers are capable to track cargo and are more efficient in placing orders (Lee, I. & Lee, K. 2015, 433). What comes to consumption behavior, the possibilities are endless as well. For example, Netflix is known to use big data. Instead of relying on managing director's instinct to create hit TV shows and movies, Netflix predicts preferences of its 34 million viewers by analyzing subscribers records of streaming activities (Ervelles, S. et.al. 2016, 900).

3.6 Post-Purchase Evaluation

Referring to the problem solving phase, consumers begin to evaluate the gap between their expectations and consumption experience during and after the consumption. It is hard to find any consumer big data from the post-purchase evaluation, since consumers tend to do the evaluation process in private. However, positive or negative experiences can occur in social media as reviews, tweets or even recommendations. (Hofacker, C.F. et al. 2015, 92)

As social media have become an integral part of our daily life, there have been emerged new ways to evaluate products. Posting content on social media is a new type of word-of-mouth (WOM) behavior, which is defined as an oral, informal, person-to-person communication – between a perceived noncommercial storyteller and a receiver regarding a brand, a product, an organization or a service (Duan, J. & Dholakia, R.R. 2017, 404). Giving an example, according to O'Connor (2014) Nike is the most popular brand on Instagram with over 22 million posts with the hashtag Nike.

3.7 Post-Purchase Engagement

The final step of consumer's decision-making process is post-purchase engagement. The interactive environment of new digital technologies enables both the company and the customer to have a continuous dialogue and share information with one another and offers multiple opportunities to engage with each other (Kunz, W. et.al. 2017, 163). The big data obtained from each consumer decision-making step's in digital environment provides a fruitful context in which behavioral aspects of customer engagement can be captured and examined.

Prototypical consumer big data exemplars are product or service reviews, comments, likes and recommendations, which takes us full circle back to the problem recognition phase. In turn, people's willingness to show their engagement to the product creates comments and likes which influences on other's alternative evaluation, one consumer's behavior becomes the antecedent for another's. (Hofacker, C.F. et al., 2015, 92) In the article, "Customer engagement in a Big Data World", Kunz, W. et.al. (2017) present three distinct levels where company can benefit from customer reviews and discussion of their experiences on the product. First, at a company level, company can gather valuable market insight for managing their reputation, complaints and intelligence for improving processes. Secondly, at a market level, customer can become a strong brand advocator or an e-marketer of the

brand. Finally, at a customer level, customers can enhance their self-brand connection and brand usage intent, their subsequent brand loyalty as well as enrich and personalize their experiences.

3.8 Overview

Consumer's decision-making process is a complex process and the length of it can vary from impulsive buying to a year-long decision making. The process can also end at any phase, for example at the search phase if consumer can not acquire enough information from the product or there is too much of it. Whether the process ends before purchase or continues after purchase, companies have a great chance to collect valuable data from the consumer's actions. Figure 2 presents the decision-making process of the consumer and illustrates what kind of data each phase generates.

Consumer's decision-making process starts from problem recognition stage where the consumer becomes aware of certain products or services. As mentioned previously, digital environment, especially social media's communities, have a big impact on the consumer's awareness of products. Monitoring or crowdsourcing what is going on in these communities, what they are talking about how the products are reviewed and what recommended generates valuable consumer data. Realizing the tremendous impact that the communities in social media have on consumer's need recognition can create new product ideas as well as improve effectiveness of marketing activities. The key is to scope which platform has the biggest impact to consumer need recognition.

Searching information is a stage towards alternative evaluation. In the search phase, a consumer goes through different information sources such as, company's websites, different social media platforms, forums and blogs. Clickstream or in other words the Internet path to the product generates valuable data about how consumer ends up choosing a specific product into the alternative evaluation. Analysing clicks includes within-site information such as the pages visited, time spent on each page, clicks on a page, and between-site information such as the websites visited. Moreover, rapid growth of mobile phone users has led to the point where people are more often using mobile devices for information search. What makes it interesting is that mobile phone users are increasingly searching information about the products from social media

The products which eventually get into the evaluation phase more and more tend to be the ones that other people have recommended. That is the reason why recommendations play a big role in alternative stage. For companies, realizing that consumers have become less tolerant to traditional

marketing activities but are nowadays more relied on other's opinions of products is the key. This social media data can be a product of "likes" in Instagram, reviews in Twitter or even VLOGs in YouTube.

Purchase phase itself does not generate much big data, but to some extent the channels that a consumer uses to buy the product, does he buy it from company's web site or from e-retailers, for example Amazon, can be defined as big data. Consumption instead generates maybe the most data from the whole process since it is increasingly present online. Huge amount of data flows continually as consumers post pictures and videos about the products that they have bought and in addition the rapid growth of IoT which allows to monitor consuming habits in real time.

Post-purchase evaluation and followed possible engagement takes back to the beginning of process as a consumer evaluates the gap between expectations and consumption experience. Product or service evaluation can be presented online as reviews, recommendations or non-recommendations hence the amount of consumer's engagement depends on the positive or negative consumption experience. Recommendations are a significant big data source as they are visible in social media communities and this way can influence other's decision-making process, so one's behaviour becomes antecedent for another's

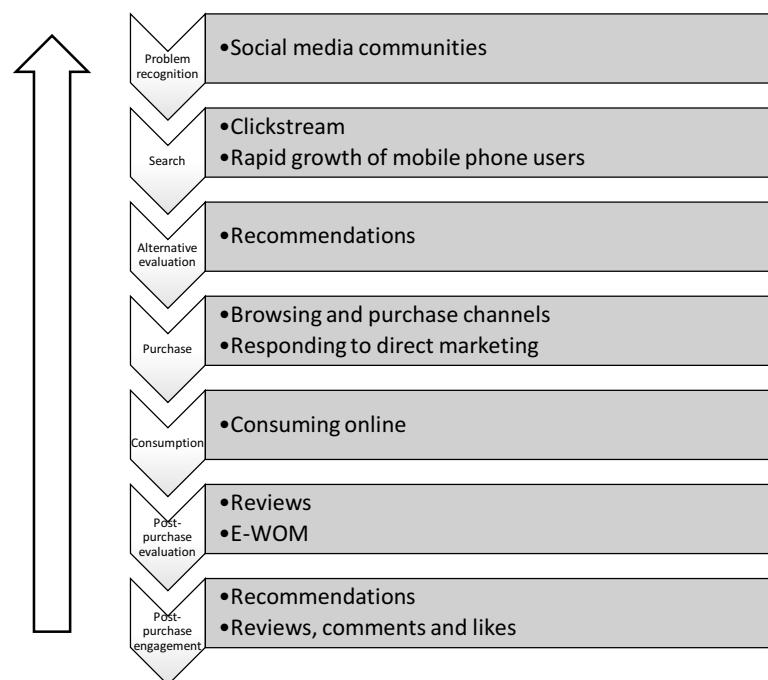


Figure 3. Overview of consumer's decision-making process

4. CONSUMER DECISION-MAKING PROCESS IN CASE-COMPANIES

The goal of the empirical part is to analyze the interview answers and compare them to the given theory and find similarities or even differences between them. The empirical part is built on three interviews from different experts from various fields. The interviews have been carried out through e-mail interviews. Making sure that the interview answers were accurate, some amplified questions were needed. The interviewees responded frankly to all questions and there was no need to skip the questions because of classified information. The interview answers are transcribed and some of the answers are presented as direct quotes.

4.1 Company and Interviewee Introduction

The first interview is carried out with a B2C company MOI-mobiili, which is a Finnish phone service operator established in 2016. It is completely owned by Finnish investors and is based in Helsinki region with six employees. MOI-mobiili is the only company in the field of phone service operators which fully operates and conducts its marketing and selling in digital environment. The interviewee is done with MOI-mobiili's founder and current partner Juha Koivuniemi. MOI-mobiili was chosen to the thesis, because it fully operates in digital environment and this way gives versatile and relevant information to research.

The second interview is conducted with Zeeland Family's marketing architect and media strategist Antti Leino who possess 20 years of experience in digital media and marketing from different design companies. Zeeland Family is an all-around digital marketing agency concentrated in B2B customers. Zeeland Family is a significant company and has produced different marketing solutions for Finland's biggest companies such as, Hartwall, Metsä Group, YIT and Elisa. The interview was conducted with Zeeland Family to study consumer behavior in business-to-business environment and compare the result with B2C.

The third interview is carried out with Tanja Jaatinen who works in the department of legislation initiatives of Finnish Ministry of Justice and has been working closely with the new data protection regulation. She also has been participating in the international executive task force which has prepared the new regulation. Moreover, she works in the board of data protection as a secretary. Tanja wants to notify that the following statements present her personal views of the matter and have

nothing to do with the organizations that she is part of. Jaatinen is a great addition to the thesis and her involvement gives an inside information about the goals of the new regulation.

4.2 Consumer Decision-Making Process

The purpose of this chapter is to view the relation between data and consumer behavior and examine how this differs in MOI-mobiili and Zeeland Family. The goal is to find correlations between theoretical and empirical parts and understand the differences with them.

“Big data is a massive data set which requires a software to process the data into an understandable form. Usually, almost always based on a quantitative, unconscious data collection. In a practice, Kantar TNS Atlas database is an embodiment of big data” – Antti Leino

“It is a mass data available from various sources, which has not yet been processed into anything particular. Analyzing big data helps to extract the relevant data, which can be used for a company’s benefit. Of course, increased computing power, the cost of storing data and the development of data analytics will change the data interpretation.” -Juha Koivuniemi

The role of big data varies between companies. Data is everywhere and available from secondary sources, such as government’s and national institute’s registers and collecting it from external sources or even from their own database. With the data, the purpose is to expand the view from markets, enhance company’s internal and external operations, and study consumer behavior. As a phone service operator, MOI-mobiili gathers and stores enormous amounts of data, which will be turned into products and services. As a phone service operator, the company has a chance to collect, process and compound the data as long it meets the general regulations of personal data protection. *“The role of the data is huge to us and the regulation of the industry is a very strict, so we must be extra careful how can we use the data from consumer behavior.” – Juha Koivuniemi*

In case of Zeeland Family, the role of big data is not so big yet but constantly growing. *“Basically, all the decisions of media planning are based on data and done by using some software for example, AdWords and Facebook Business Manager etc. Keyword optimizing and social media advertisement targeting decision are all based on consumer big data.” – Antti Leino*

Some important big data analytic tools and softwares to collect, analyze and store data, which surfaced during the interviews, were Google Analytics, AdWord, Facebook Business Manager, LinkedIn Campaign Manager, Twitter Ads, Adform, Doubleclick and addition these some data are bought from Kantar TNS Atlas, Futusome and Ihminen 360.

Consumer's decisions are based on emotional and psychological reasoning and the decision-making process can vary from impulsive to year-long decision-making. In B2B markets consumers usually require customized products, whereas in business-to-customer markets, consumers buy standardized products and the threshold of changing a vendor is low. In the era of social media, the decision-making process is affected more by blogs, chat forums, word of mouth and social media communities. Since people are more connected with digital environment, there is also more data available and better opportunities to create efficient dialogue with the target audience. Enriching the understanding of consumer behavior, what are the trends and motivational factors within the purchase process, helps to create new advertising and communication concepts. The goal is to continually improve and build better contacting strategies in digital environment.

Based on the interviews it is fair to say that the study of consumer behavior is essential for every company that wishes to succeed. Despite of the differences in B2C and B2B markets, a lifeline is to know current customers and to recognize the potential ones. Due to the world is transiting more to the online environment, data collection is becoming easier. From the very first step the consumer takes in the search process, data starts to flow and information from consumers are continually obtained. In case of MOI-mobiili, where the all actions – selling and marketing – happens through network, knowing the consumer is everything.

“Because our customer acquisition happens in digital environment, the whole process is run by data. With the help of different methods and softwares, the strive is to extract the relevant data from datasets.” - Koivuniemi

“The goal is to create a process driven by our consumers' choices, which continually breeds both the messages, media choices as well as products” – Koivuniemi

To study consumer behavior, consumer decision-making process is a general model and can be adapted almost everywhere. The only thing that varies is the length of each step. The same path can be found everywhere, from a chocolate bar to television and the length of it ranges from impulsive

buying to even years. *“The most important thing is to recognize where the brand x falls off from decision-making process and find the reasons why, and finally act upon them”* – Leino. It is hard to point out which of the steps are the most important ones, since it can depend on so many things, such as market conditions and presentation of alternatives. *“Of course, without a need (problem recognition) there is no business.”* – Leino

For companies that operate in digital environment consumer data collection usually takes place on websites through forms with some incentives or competition. Some companies, like electric providers, purchase data from Posti. Despite the fact, that MOI-mobiili operates in B2C -markets and Zeeland Family B2B –markets, both agree that each step of consumer decision-making process have equal importance and that they generate valuable information for different usage. Being a relatively young company, MOI-mobiili puts a lot of weight on the alternative evaluation. *“MOI-mobiili is a relatively new brand (under 2 years old), which means that brand awareness continues to be emphasized. Being a new comer in the field of phone service operators, the most important thing is to get into alternative evaluation with the biggest competitors, such as Elisa, Telia and DNA. Of course, I can not highlight the purchase process enough. Continuous improvement and making the purchase process as easy as possible for customers and make the benefits of it more concrete.”* – Koivuniemi

5. GENERAL DATA PROTECTION REGULATION

On May 25th, the new General Data Protection Regulation comes into effect which brings new challenges and obligations to companies around Europe. While the regulation requires companies to follow data protection regulation more specifically, it influences on companies as well as consumer behavior and forces the dialogue between company and consumer to be more transparent.

The current data protect regulation is over 20 years old which has forced European Union to take actions and reorganize it. EU's current data protection regulation, which will be now repealed with the new regulation, has been implemented in the Member States in different ways and thus made difficulties to the information flow. With the reorganization, the goal of European Union has been to create an up-to-date, strong, cohesive and comprehensive data protection framework for European markets. The regulation will be applied throughout EU, which unifies the processing of personal data significantly. *“The reorganization has been necessary especially due to the rapid development of information technology and inconsistent data protection regulations between Member States. The goal has been to unify data protection regulations between countries and facilitate the information flow in Europe.”* – Jaatinen

5.1 Influences on Companies and Consumer Behavior

The topic has widely generated conversations about what kind of impact it has on business as well as on consumer behavior. Generally, the topic has raised uncertainty among companies and drifted the conversation towards the imagined threats. GDPR means big changes to companies in how they need to protect their customers, employees and others personal information. Sanctions are huge; either 4% from annual sales or 20 million euros depending on the company. The regulation forces each company that conduct business with personal data in the EU to keep a better record of all information that they have on the individuals. Basically, it means that a company needs to put more weight on to getting a permission from consumer to collect data.

Starting at end of May, individuals have more rights concerning the information that companies have on them, including online account data, location data and biometric data. Basically, the new regulation obliges companies to provide a statement about individual's personal data if he or she demands it. Addition to this, individuals are entitled to have the data deleted or transferred to

elsewhere if wanted. Companies also are obligated to prove these actions. With these reorganizations, companies are forced to reform their data control systems as well as do some re-arrangements in organization levels.

Companies have had a two year transition period which has made the adaptation easier. In Zeeland Family, the company's internal working group has drawn general instructions and the responsible person has been appointed. Whereas in MOI-mobiili, where personal data collection is greater, from collecting to storing and analyzing processes related to personal data have been overviewed and renewed. The biggest change is ensuring that the data-related processes are secured including with the subcontractors, must be carefully designed and the whole process needs to be transparent to the individual. The transformation has not forced the case companies to outsource, but the previous contracts with subcontractors have been checked.

“The transformation into GDPR has not created a huge conversation in our company.” –Antti Leino

“Our industry has already been strictly regulated for a long time, so the change is not as significant as it can be in some other industries. Of course, the transformation has brought a little more work.”
–Juha Koivuniemi

With the new regulation, the emphasis is on getting the permission to collect personal data from the consumer. If a company is not justified by the law to collect personal data, the company needs to present a clear request to collect personal data from the individual. GDPR requires that the company's presentation to gather personal data is clear, separated from the rest and the form of it is easily to understand. Moreover, the company is obligated to prove the granted permission. Basically, the easiest way to present this is through a form. For example, making a reservation from a restaurant online, the restaurant is required to ask a permission to register personal data before sending newsletters to email. For some, the act does not bring any difficulties on the table, like in case of MOI-mobiili. However, it seems that the biggest influence will be on marketing design agencies and on their campaigns. One of the most important subject is how data protection should be communicated behalf of B2B –marketers who especially use consumer big data to obtain leads and scoring-models. That is the why agencies should be as transparent as they can.

“Campaigns will be more detailed and customers are more informed that the information will be analyzed and used in marketing activities. Also, informing that opening a digital newsletters and other clicks are registered and possible used for targeting advertisement.” –Leino

One of the most important subject is how this should be communicated behalf of B2B –marketers who especially use consumer big data to obtain leads and scoring-models. That is the why agencies should be as transparent as they can.

For the consumers, a change will be the right to have all the data that company might possess of them and get it destroyed if they want so. It is hard to predict how the entitlement will manifests itself among consumers. The individuals already have several legal privileges, for example to review the information that company possess of them. Companies that hold personal data records, have already had requests to review the consumer’s personal information. With the new regulation, it is most likely possible that the individuals will use more this privilege, especially now when the awareness of data protection regulation is increasing in EU.

“At the first, the tension might be in how quickly the company can react, if consumer uses the right.” –Antti Leino.

GDPR seeks to build trust among consumers, in order for the digital economy to develop in EU. An individual must be able to follow what and where their personal information is used as well as trust that the information is handled properly throughout its life cycle no matter in which Member State the information is used. It is hard to predict how GDPR will change consumer behavior nor competition. For instance, it is possible that from now on consumers will avoid companies that are not operating transparently enough. It may also be that a foreign company can gain more trust from consumers for being more transparent.

5.2 Threats and Opportunities

The new regulation forces companies to take data protection seriously, not only an additional custom in a marketing campaign. Of course, some companies perform better than others for example, companies in retail and service sectors such as, S-Ryhmä and Kesko as well as companies with smart CRM processes, are pioneers at data processing in Finland. To most companies’ data control over different softwares has been challenging for a while and the data has been fragmented, and handled and saved across organization levels. From now, if an individual asks to see the information on him

or her, company needs to respond properly regardless how and where to information is stored. To some extent, the data collection will become more difficult because the collection requires permissions from consumers.

“Companies that have been juggled with Excel for a long time, are now facing challenges and investments.” – Leino

One of the biggest threat is to fail in data protection. For example in recent days there have been a movement, where a lot of Facebook users have piled onto the hashtag #DeleteFacebook, threatening to desert their Facebook accounts as a protest to Facebook for mishandling their personal information (Guynn, J, 2018).

“There have been boycotts and protesting among consumers, for example Facebook Delete – movement.” - Jaatinen

As well as the data protection can be a challenge but properly implemented it can generate cost savings in the long run, thus better products and services, more satisfied customers, recommendations and eventually new customers. Of course, with the companies that has not yet had effective data processes are now forced to master their data control. With a good control over the consumer data, company is capable to extract the relevant data and to delete the irrelevant or expired, hence streamline its business and achieve cost savings.

Since the regulation will be the same in each Member State, opportunities arise especially for companies who operate business more than in several countries. From consumer's view, the fact that he or she can trust the personal information in the hands of unfamiliar company, can influence his or hers' purchase process hence, taking special care of company's data protection can turn up to competitive advantage. Consumers might favor companies that mind their data protection and thus are more transparent than others.

“I hope that in a long run, GDPR builds more belief and trust among consumers’. At least, MOI strives to operate as transparently as possible and to tell straightforward about our operating models and data processing.” -Koivuniemi

To companies, that have not been able to process data appropriately before, the compulsion to take actions in this field can present unexpected advantages. The better consumer data structure is build,

the better data is understood and interpreted. We are talking about customer 360 picture. Knowing consumer's needs and interests, marketers are more able to create effective communication.

"I rather receive ten advertisement targeted to my needs than a hundred mixed." – Koivuniemi

"Nothing new, the better and smarter you operate, the better you will achieve in marketing and competition." – Leino

6. CONCLUSIONS

The purpose of this study has been to study consumer behavior in digital environment and to find out what kind of consumer big data it generates and how the new data protection regulation (GDPR) could influence on that. This has been done by answering to the research questions of the thesis. The theoretical part is built on previous researches of the topic and the aim of it has been in creating a broad picture of the phenomenon. With the empirical part, the objective has been to find out how companies study consumer behavior and what kind of data they collect. The study of consumer behavior in companies, especially in the data collection, and in finding hidden patterns behind the behavior, are essentials. The previous data protection regulation has not forced companies to take seriously the questions of data protection but starting on end of May, the new general data protection regulation comes into effect which will change the way companies collect personal information from consumers. It is hard to predict how the new regulation will influence consumer behavior and whether the influences appear right away or after few years.

The research questions of the thesis are:

“How General Data Protection Regulation could influence on consumer behaviour?”

“What is consumer big data?”

“How GDPR will influence companies’ data collection?”

“What kind of challenges and opportunities GDPR might bring into consumer behaviour?”

With the help of sub-questions, thesis strives to answer to the main research question.

“What is consumer big data?”

Big data grows bigger and bigger every day and the consumer behaviour on social media represent a major driver for the phenomenon. By defining consumer big data, the theoretical part of the thesis seeks a comprehensive explanation for the phenomenon. It can be stated that there are as many definitions or understandings of consumer big data as there are people studying it. However, the

exponential growth of social media usage by the consumer and the growth of IoT has resulted the three general characteristics of consumer big data, volume, velocity and variety (Hofacker, C.F. et.al. (2016), 95) Moreover, consumer big data, or big data generally, has other additional characteristics as well. The value, variability and veracity are the concepts made to understand the uncertainty and complexity of the data.

For the first sub-question, the thesis aimed to seek answers from the theory and compare it to the answers of interviews. It can be stated that with case companies, big data is understood a bit differently that it is presented in the theoretical part. Gartner It Glossary (2017a) describes big data through high-volume, high-velocity and high-variety information assets. However, in the case companies, consumer big data is understood as a massive data and the emphasis of velocity and variety lacks not to mention the additional three v's. Analyzing consumer big data helps to extract the relevant data, which can be used for a company's benefit. With the consumer data, the purpose is to expand the view from markets, enhance company's internal and external operations, and to create marketing communication to the consumers.

“How GDPR will influence on companies' data collection?”

GDPR forces companies to put serious thoughts on the consumers' personal data protection, not only additional manner in marketing campaigns. It means big changes on how companies protect their consumers', employee's and other people's personal information. Companies need to understand what, where and how to deal with consumer's personal data. It puts a lot weight on getting a permission to collect the consumer data, which will slower the whole process of data collection.

With the new data protection regulation, the study of consumer decision-making process and the data capturing from it is becoming even harder. Hofacker, C.F. et.al. (2015) in the article “Big Data and consumer behavior: imminent opportunities” speak about how consumer big data analytics has significantly transformed the way marketing is conducted today and how consumer decision-making process generates huge amount of consumer big data. Since the upcoming regulations requires companies to ask a permission to collect personal data from the consumer, will the consumer big data collection become harder and in some extent slow down. The emphasis will be in transparency and how companies are able have a two-way dialogue with the consumer.

A good and an efficient data protection and transparency of data collection builds trust among consumers. This way they are more willing to be in a contact with the company as well as giving more accurate personal data about their interests and needs. Investing on a better and more secured data processes will lead into a better data control and thus creates cost savings within a data collection. The biggest influence concerns B2B –marketers, the companies who especially use consumer big data to obtain leads and scoring-models. Whereas for MOI-mobiili the new regulation does not bring any big changes since the industry has already been strictly regulated in personal data collection.

“What kind of challenges and opportunities GDPR might bring into consumer behaviour?”

Generally, the companies have taken the transition to the new regulation lightly and the overall understanding of securing data is good in Finland. The biggest threat is on failing. The new regulation forces companies to take data protection seriously, not only an additional custom in a marketing campaign. Giving an extreme example from possible failure, how Facebook mishandled personal information of user and how it resulted to Facebook Delete –movement. The better companies are able to succeed in the data protection, the more consumers are willing to trust and give more personal information about them.

The world is changing all the time and more and more younger generation are becoming consumers, for whom digital environment and shopping online is as familiar as shopping in a regular store. Therefore, the growth of online shops and the increase of globalization keeps going on. Language translation providers will progress, which may result that even Chinese language would not restrain purchasing online. Algorithms that facilitate more and more our everyday life such as, where to invest, who do we date and where do we get our inspirations, will increase all the time. Moreover, the unstable weather conditions can lead to transformation of shopping malls – shopping malls becoming hang-around-places. It is interesting if they could be places with an entrance fee, where the consumers can have experiences, spend their holidays at beaches and other things that are restrained by our weather conditions.

Customer profiling will continue and develop further, although the negative sides of it have been addressed and tried to control. A good example of it is widely reported company that is focused to detect manipulation of elections. Since the digital world is constantly evolving, it is going to be hard for legislation to keep up the pace of technology development.

“I’m sure that in five to ten years every household has voice command devices, such as Alexa and Siri, for different tasks.” – Leino

“Nothing is anymore remembered, but we can remember where to find it.” –Leino

Considering the mentioned guidelines of the future, only the time will show how GDPR will restrain future’s consumer behavior and the collection of consumer big data.

7. REFERENCES

Bayus, L.B. (2013) Crowdsourcing new product ideas over time: Analysis of the Dell Idea Storm community. *Management Science*, 50(1), 226-244.

Blackwell, R.D. Miniard, P.W. and Engel, J.F. (2005) Consumer Behavior. 10th Edition. South-Western College Publications.

Chan, H.K. Lacka, E. Yee, R.W.Y. & Lim, M.K. (2017) The role of social media data in operations and production management. *International Journal of Production Research*, 55(17), 5027-5036.

Chan, K.W. Li, S.T. Zhu, J.J. (2015) Fostering customer ideation in crowdsourcing community: The role of peer-to-peer and peer-to-firm interactions. *Journal of Interactive Marketing*, 31, 42-46

Dinu, D. Stoica, I. Radu, A.V. (2016) Studying the consumer behavior through Big Data. *Quality-Access to Success Journal*, 17(3), 246-254.

Doug, L. (2001). "Application Delivery Strategies". *Meta Delta*.)" [Online]. [Accessed 20.11.2017] Available at: <https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

Elinkeinoelämän keskusliitto, (2018), Tietopaketti yrityksille: On tärkeätä valmistautua EU:n tietosuoja-asetukseen. [Online]. [Accessed 9.5.2018] Available at: <https://ek.fi/mita-teemme/yrityslainsaadanto/tietosuojalainsaadanto/tietopaketti-yrityksille-on-aika-valmistautua-eun-yleiseen-tietosuoja-asetukseen/>

Duan, F. Dholakia, R.R. (2017) Posting purchases in social media increases happiness: the mediating roles of purchases' impact on self and interpersonal relationships. *Journal of consumer marketing*. 34/5, 404-413

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

Gartner (2017). “Big Data” [Online]. [Accessed 10.10.2017] Available at: <https://www.gartner.com/it-glossary/big-data/>

Guynn, J. (2018) Delete Facebook? It’s a lot more complicated than that. USA Today [Online]. [Accessed 2.4.2018] Available at: <https://www.usatoday.com/story/tech/news/2018/03/28/people-really-deleting-their-facebook-accounts-its-complicated/464109002/>

Hanssens, D. M. & Pauwels, K. H. (2016). Demonstrating the value of marketing. *Journal of Marketing*, 80(6), 173-179.

Hofacker, C. Malthouse, E. C. Sultan, F. (2016). Big Data and consumer behavior: imminent opportunities. *The Journal of Consumer*, 33(2), 89-97.

IBM (2017a). “Tech Trends – predictions for 2017” [Online]. [Accessed 15.11.2017] Available at: <https://www.ibm.com/blogs/watson/2017/01/seen-future-top-5-predictions-2017-tech-trends-report/>

IBM (2017b). “What is big data? More than volume, velocity and variety...” [Online]. [Accessed 15.11.2017] Available at: <https://developer.ibm.com/dwblog/2017/what-is-big-data-insight/>

IBM (2017c). “What is Big Data Analytics” [Online]. [Accessed 15.11.2017] Available at: <https://www.ibm.com/analytics/us/en/big-data/>

Khan, N. Yaqoob, I. Ibrahim, A. T. H. Zakira, I. Waleed, K. M. A. (2014) Big Data Survey, Technologies, Opportunities and Challenges. *The Scientific World Journal*.

Kunz, W. Aksoy, L. Bart, Y. Heinonen, K. Kabadayi, S. Ordene, F.V. Sigala, M. Diaz, D. Theodoulidis, B. (2017) Customer engagement in a Big Data world. *The journal of service marketing*, 31(2), 161-171.

Lindsey-Mullikin, J. & Borin, N. (2017). Why strategy is the key for sales in social media. *Business Horizons*, 60(4), 473-482.

O'Connor, C. Feb 13 (2014) Starbucks And Nike Are Winning Instagram (And Your Photos Are Helping). Forbes [Online]. [Accessed 16.11.2017] Available at: <https://www.forbes.com/sites/clareoconnor/2014/02/13/starbucks-and-nike-are-winning-instagram-and-your-photos-are-helping/#13cc1f4b4894>

Oracle (2017a). "An Enterprise Architect's Guide to Big Data" [Online]. [Accessed 16.11.2017] Available at: <http://www.oracle.com/technetwork/topics/entarch/articles/oea-big-data-guide-1522052.pdf>

Oracle (2017b) "Oracle Privacy Policy" [Online]. [Accessed 16.11.2017] Available at: <https://www.oracle.com/legal/privacy/privacy-policy.html#cookies>

Powers, T. Advincula, D. Austin, M.S. Graiko, S. Snyder, J. (2012) Digital and social media in purchase decisions process. A special report from the advertising research foundation. *Journal of advertising research*. 479-489

Senecal, S. Kalczynski, P.J. Nantel, J. (2005). Consumers' decision-making process and their online shopping behavior: a clickstream analysis. *Journal of Business Research*, 58(11), 1599-1608.

Senecal, S. & Nantel, J. (2004) The influence of online recommendation on consumers' online choices. University of Montreal. *Journal of Retailing*, 80, 159-169.

Statista (2018). "Number of monthly active Facebook users worldwide as of 4th quarter 2017 (in millions)" [Online]. [Accessed 2.12.2017] Available at: <https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/>

Tietosuoja (2018), EU:n tietosuojaudistus [Online]. [Accessed 9.5.2018] Available at: <http://www.tietosuoja.fi/fi/index/euntietosuojaudistus.html>

Wang, G. Zhang, X. Tang, S. Zheng, H. Zhao, B.Y. (2017). Clickstream User Behavior Models. *ACM Transactions on the Web*, 11(4, 21).

Wang, Y. & Yu, C. (2017). Social interaction-based consumer decision-making model in social commerce: The role of word of mouth and observational learning. *International Journal of Information Management*, 37(3), 179-189.

Werner, K. Aksoy, L. Bart, Y. Heinonen, K. Kabadayi, S. Ordenes, F.V. Sigala, M. Diaz, D. Theodoulidis, B. (2017) Customer engagement in a Big Data world. *Journal of service marketing* 31/2, 161-171

Wolny, J. & Charoensuksai, N. (2014) Mapping customer journeys in multichannel decision making. *Journal of direct, data and digital marketing practices*, 15(4), 317-326.

Yaqooba, I. Hashema, I.A.T. Gania, A. Mokhtara, S. Ahmed, E. Anuara, N.B. Vasilakosb, A.V. (2016). Big data from: beginning to the future. *International Journal of Information Management*, 36(6), 1231- 1247.

Zhang, K.K.Z. & Benyoucef, M. (2016) Consumer behavior in social commerce: A literature view. *Decision Support System*, 95-108.

Zhang, H. Zhao, L. Gupta, S. (2018). The role of online product recommendations on customer decision making and loyalty in social shopping communities. *International Journal of Information Management*, 38(1), 150-166.

Attachments

Interview 1:

Background

- Mikä on haastateltavan henkilön asema? (Liittyen GDPR:ään + voi myös vapaasti kertoa omista työkokemuksista)?

The first part

- Mikä on ollut suurin syy uuden lakiasetuksen kehittämiseen/syntyyn?
 - Mitä lakialoitteella tavoitellaan?
- Kuluttajilta on asetuksen myötä pyydettävä erikseen suostumus henkilötiedon keräämiseen. Tuleeko tämä konkretisoitumaan ja kuinka näette tämän vaikuttavan tiedon keruuseen?
- Uskotteko kuluttajien olevan tietoisia uudesta tietosuoja-asetuksesta?
 - Nuorempi sukupolvi on asioista entistä tietoisempia, luuletteko heidän olevan enemmän tietoisia asioista ja onko heillä enemmän halukkuutta pyytämään yrityksiltä tietonsa?
 - Mikäli ei ole, niin luuletko asian muuttuvan lähitulevaisuudessa?
 - Luuletteko, että kuluttajat tulevat kysymään tietojaan yrityksiltä ja pyytämään poistamaan ne? (Nykyhetki tai tulevaisuudessa?)
- Mitkä ovat suurimmat haasteet ja mahdollisuudet mitä tietosuoja-asetus tuo tullessaan?
 - Yrityksille?
 - Investoinnit? (esim. Suuryritykset vs. pienet / vanhat vs. alle 2v vanhat yritykset)
 - Kuluttajadatan hallinta?
 - Kustannussäästöjä?
 - Customer 360?
 - Kuluttajille?
- Tietosuoja-asetuksen myötä tuntuu, että kaikki lähtevät samalta viivalta mitä tulee asiakastiedon keruuseen. Kuinka itse koette kilpailun/markkinoiden muuttuvan asetuksen myötä? (Tuleeko kuluttajista entistä luottavaisempia läpinäkyvyyden lisääntyessä, että ovat valmiimpia antamaan enemmän/kattavampaa tietoa itsestään?)

- Kuinka näette kuluttajakäyttäytymisen muuttuvan seuraavan 5vuoden/10vuoden aikana asetuksen pohjalta? Tai kuinka yritysten toiminta voisi muuttua?

Interview 2:

Background

- Mikä on haastateltavan henkilön asema yrityksessä (+voi myös vapaasti kertoa omista kokemuksista markkinoinnin alalta)?
- Big Datan määritelmiä on yhtä paljon kuin on sitä tutkivia. Kuinka te kuvailisitte mitä Big Data on?
 - Kuluttaja Big Data/Kuluttaja 360?
 - Minkälainen rooli Big Datalla on yrityksessänne?
 - Kuinka suuri merkitys on henkilötietojen keruulla yrityksessänne? (asiakasyrityksellenne)

The first part

- Millaista on yrityksenne kuluttajatiedon hankinta?
- Kuinka suuri osa yrityksenne liiketoiminnasta perustuu kuluttajakäyttäytymisen tutkimiseen?
 - Datan keruu ja hallinta?
 - Kuluttajadatan jalostamiseen insighteiksi?
 - Millaisia tavoitteita teillä on kuluttajakäyttäytymisen tutkimisella verkossa?
- Kuvaan työssäni kuluttajan päätöksentekoprosessia kuudella vaiheella (ongelman tunnistaminen, tiedon haku, vaihtoehtojen arviointi, ostokäyttäytyminen, kuluttaminen, jälkiarviointi). Näiden avulla minun on tarkoitus selvittää, millaista kuluttajadataa jokainen vaihe tuottaa.
 - Miten vaiheista näette itse kuluttajan päätöksentekoprosessin muodostuvan?
 - Mitä prosessin vaiheita pidätte tärkeimpinä yrityksenne näkökulmasta ja miksi?
- Millaisia työkaluja yrityksellänne on datan keruuseen, säilöntään ja analysointiin?
- Mitkä kanavat ovat tärkeimmät kuluttajakäyttäytymisen tutkimiseen?

The second part

- Direktiivin siirtymäaika on ollut kaksi vuotta. Kuinka yrityksessänne on valmistauduttu toukokuussa astuvaan tietosuojasetukseen?

- Millaisia valmiuksia on jouduttu tekemään (organisaatio, ohjelmistot, vastuuhenkilöt)?
- Onko uuden tietosuoja-asetuksen myötä jouduttu jotain toimintoja ulkoistamaan?
- Kuinka asia on koettu organisaatiossa?
- Minkä koette olevan uuden tietosuoja-asetuksen suurin muutos verrattuna aikaisempaan asetukseen?
- Kuluttajilta on asetuksen myötä pyydettävä erikseen suostumus henkilötiedon keräämiseen. Kuinka näette tämän vaikuttavan tiedon keruuseen?
- Uskotteko tavallisen kuluttajan olevan tietoinen uudesta tietosuoja-asetuksesta?
 - Miten luulette asetuksen vaikuttavan kuluttajan halukkuuteen antaa henkilötietoja?
 - Ovatko kuluttajat entistä halukkaimpia tietämään, mitä tietoa yrityksellä on hänestä?
 - Halukkaita, että tämä tieto hävitetään?
 - Läpinäkyvyys lisääntyy asetuksen myötä. Luuletteko tämän edistävän yrityksen ja kuluttajan välistä kommunikaatiota?
 - Ovatko kuluttajat entistä halukkaimpia antamaan tietoa itsestään?
- Mitkä ovat suurimmat haasteet mitä tietosuoja-asetus tuo tullessaan?
 - Mahdollisuudet?
 - Yrityskuvan muuttuminen (vastuullisuus)?
 - Customer 360 (tiedon hallinta)?
 - Kustannussäästöt?
- Tietosuoja-asetuksen myötä tuntuu, että kaikki lähtevät samalta viivalta mitä tulee tiedon keruuseen. Kuinka itse koette kilpailun tai markkinoinnin muuttuvan asetuksen myötä?
- Kuinka näette kuluttajakäyttäytymisen muuttuvan seuraavan 5vuoden/10vuoden aikana?