CHALLENGES IN COLLECTING AND USING WEB DATA IN THE BANKING SECTOR

Jyväskylä University
School of Business and Economics

Master's Thesis

2024

Author: Julia Silonsaari
Subject: Challenges in collecting and using web data in the banking sector
Supervisor: Chiara Valentini

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

ABSTRACT

| Author | |
Julia Silonsaari	
Title	
Challenges in collecting and using web data in the banking sector	
Subject	Type of work
Digital Marketing & Corporate Communication	Master's Thesis
Date	Number of pages
April 2024	56

Digitalization has revamped the marketing landscape rapidly and has changed a lot of business actions to online services. It has become increasingly important for companies to be able to seek knowledge on how their customers are performing within their websites and other digital marketing channels. Therefore, web analytics and tracking are important tools to gain understanding and insights for decision-making and the future development. There are, however, still improvements needed for organizations to be fully data-driven and there are still known challenges in web analytics usage. Moreover, there are regulations impacting the data collection within countries in the European Union.

The objective of this study is to focus on researching web analytics as a central component of digital marketing activities in the context of the banking industry. The idea is to increase the awareness and understanding of web data usage and the issues of it within the banking industry, where data involved consists of great amounts of sensitive data. In addition, this study takes place in Finland and as the country is a part of the European Union, it operates under the legislations and regulations within the EU. This study seeks to gain knowledge on the biggest challenges in web data collection and usage, focusing on the banking industry in Finland. Three research questions were created to support the thesis: (1) What is the purpose of using web analytics in banking sector, (2) what the main issues are when collecting web data in a banking sector and, (3) what the main challenges are when analysing web data in a banking sector.

A qualitative approach was chosen to support the explorative nature of the research. Interview data of banking industry professionals were collected and analysed with thematic analysis. The findings are consistent with previous research and indicate that the organizations are not as data driven as they think they are. Even though the industry is highly regulated, data is still collected from customers, however, the tools used have some limitations. Moreover, it seems that key performance indicators are not set to guide the data-driven decision-making indicating the data is not used to guide decisions. The findings suggest that the organizations could improve knowledge within the team units towards data collection and tracking. In addition, there could be improvements in the collaboration between teams.

Key words

GDPR, data protection, web analytics, data collection, web data, web data collection, data-driven decision-making, banking

Place of storage

Jyväskylä University Library

TIIVISTELMÄ

Tekijä Julia Silonsaari	
Työn nimi Haasteet web-datan keruussa ja käytössä pankkialalla	
Oppiaine Digitaalinen markkinointi ja viestintä	Työn laji Pro gradu -tutkielma
Päivämäärä Huhtikuu 2024	Sivumäärä 56

Digitalisaatio on muokannut markkinoinnin kenttää nopeasti ja siirtänyt monia toimenpiteitä myös yritysmarkkinoilla digitaalisiin kanaviin. Yrityksille on muodostunut kasvava tarve saada tietoa siitä, miten heidän asiakkaansa käyttäytyvät heidän nettisivuillaan tai muissa digitaalisissa kanavissa. Tämän vuoksi web-analytiikka sekä seuranta ovat tärkeitä työkaluja tiedon keräämiseen ja kasvattamiseen sekä tukemaan päätöksentekoa ja tulevaisuuden kehityskohteita. Yrityksillä on kuitenkin edelleen kehitettävää pystyäkseen täysin johtamaan päätöksentekoa datalla sekä kehityskohteita web-analytiikan käyttöön liittyen. Lisäksi tietyt regulaatiot Euroopan Unionin maissa vaikuttavat datankeräykseen.

Tämän tutkimuksen tarkoituksena on keskittyä tutkimaan web-analytiikkaa digitaalisen markkinoinnin keskeisenä komponenttina keskittyen pankkialaan. Tällä on tarkoitus lisätä tietoisuutta ja ymmärrystä web-datan käyttöön liittyen sekä sen haasteisiin pankkialalla, mikä pitää sisällään isoja määriä sensitiivistä dataa. Lisäksi tutkimus suoritettiin Suomessa, mikä on osa Euroopan Unionia ja täten seuraa EU:n alaisia sääntelyitä. Tämä tutkimus pyrkii selvittämään, mitkä ovat isoimmat haasteet web-datan keräämisessä ja käytössä, keskittyen suomalaiseen pankkisektoriin. Kolme tutkimuskysymystä kehitettiin tukemaan tutkielmaa: (1) Mikä on web-analytiikan käyttötarkoitus pankkialalla, (2) mitkä ovat isoimmat haasteet web-datan keruussa pankkialalla, ja (3) mitkä ovat yleisimmät haasteet web-datan analysoinnissa pankkialalla. Kvalitatiivinen tutkimus valikoitui tutkimustavaksi sen kuvailevan luonteen vuoksi. Aineisto kerättiin haastattelemalla pankkialan ammattilaisia ja analysoimalla tuloksia teema-analyysillä. Tulokset ovat pitkälti yhteneväisiä aiemman tutkimustiedon kanssa ja osoittavat, että yritykset eivät vielä ole niin dataohjautuvaisia, kuin oletetaan. Vaikka ala on kovin säännelty, asiakasdataa kuitenkin kerätään, mutta käytössä olevat työkalut ovat rajoitettuja. Lisäksi vaikuttaa, että KPI-mittareita ei aseteta ohjaamaan päätöksentekoa, mistä voidaan olettaa, että data ei ohjaa päätöksiä. Löydökset osoittavat kehityskohteen yrityksille datan keruun ja seurannan tiedonjaon parantamiseen liittyen sekä yhteistyöhön tiimien välillä.

Asiasanat GDPR, tietosuoja, web-analytiikka, datankeräys, web-data, web-datankeräys, dataohjattu päätöksenteko, pankkiala
Säilytyspaikka Jyväskylän yliopiston kirjasto

6

CONTENTS

LIST OF TABLES AND FIGURES

1   INTRODUCTION

Ever since the turn of 2010s, tracking and analysing web data have been possible due to the emergence of digital analytics (Lahtinen et al., 2023). Organizations have started to utilize multiple different digital marketing tools that allow them to advertise, customize content as well as target different stakeholders online to enhance their business objectives (Provost & Fawcett, 2013). Moreover, with the emergence of these new digital tools for analysis, data collection has become less expensive and data analysis has become easier than before, shifting organizations into a more customer-centric approach (Sleep, Hulland & Gooner, 2019). Simultaneously, with more and more activities taking place online in a digital environment, the amount of data collected has become enormous and businesses can find data about almost any topic of their interest (McAfee & Brynjolfsson, 2012). Even though there is an increasing interest within organizations towards data collection, the European Union's General Data Protection Regulation (GDPR) regulates what data can be collected and therefore also analysed (Siegert et al., 2020). Since the digitalization has evolved and facilitated most of the interaction between users and organizations to digital channels, it has become even more relevant to operate securely and store data about internet users in a safe way (Scheuing & Niininen, 2022).

Digitalization has revamped the marketing landscape rapidly and has changed a lot of business actions to online services, creating competition on the markets. Customers have a lot more options to choose from than before and hence they have all the power to switch providers if they wish. Moreover, it is required to gain deeper understanding on customer online behaviour as well as measure and evaluate the outcomes within the business (Bucklin et al., 2009). It is a competitive advantage on the market to analyse the daily business operations and organizations within every industry are focused on using data to gain the advantage on their market (McAfee & Brynjolfsson, 2012 & Provost & Fawcett, 2013). Research shows that organizations that identify themselves as data-driven, also perform better on financial and operational levels (McAfee & Brynjolfsson, 2012, Kemppainen et al., 2022). Therefore, managers who analyse the business will gain knowledge that can be transformed into decision-making and ultimately also business performance. Unfortunately, many organizations claim to be more data-driven than they actually are (McAfee & Brynjolfsson, 2012). It is as critical to have data-analytic thinking and knowledge, not just for the analyst, but also throughout the whole organization to be able to gain the best results and the leadership role in the market (Provost & Fawcett, 2013).

Within the financial markets, the transformation from manual actions towards the electronic trading in 1986 was the starting point for many studies on data usage within the financial industry (Seddon & Currie, 2016). Moreover, organizations have adopted data-driven decision-making as well as automatic decision-making in different times, the finance and telecommunication industries

being the forerunners for it. Already in the 1990s, the automated decision-making as well as data-driven control systems with fraud control changed the banking industry not to forget the online banking that emerged fully in the 1990s together with the Internet (Provost & Fawcett, 2013 & Beattie, 2023). Especially, the era of Internet Banking has emerged and affected the industry, where banking services are offered through the Internet. It has opened the markets for banks for more competition, since the technologies can be accessed anywhere instead of the old physical location (Yiu et al., 2007). The banking industry has during the recent years undergone changes due to financial deregulation where government rules towards banks have been lessened, and globalization (Bakar et al., 2016, Seddon & Currie, 2016). Furthermore, the COVID19 and Russia invading Ukraine have both affected the industry. The latter and more recent one affected in increase of costs in the energy markets, surge in inflation, decrease in housing markets and the decrease in foreign trade within the industry (Kettunen, 2022). In addition, the rapid technological development and customer demand for banking services online have increased (Seddon & Currie, 2016). Even though the premium services are still only offered in banking offices, most of the banking services have shifted both to the bank's online service and mobile apps. Therefore, there is an increased need to understand customer behaviour within the banking industry, especially since the industry is highly competitive and many retail banks offer the same services and develop their services fast on both internet banking and mobile banking platforms (Juhaida et al., 2016 & Bakar et al., 2016).

1.1 Research problem and questions

This study focuses on studying web analytics as a central component of digital marketing activities in the context of banking to increase the awareness and understanding of web data usage and the issues of it within the banking industry, where data involved consists of great amounts of sensitive data. In addition, this study takes place in Finland and as the country is a part of the European Union, it operates under the legislations and regulations within the EU.

The biggest changes within the financial and banking markets have been the computerization that has shifted for instance digital money transfers, payments and portfolio analysis to the web platforms, which is still scarce in the research areas (Seddon & Currie, 2016). A lot of focus and research has been given to customer loyalty on banking (Bakar et al., 2016 & Christanto & Santoso, 2022 & Nguyen et al., 2022). However, data analytics is understudied in business research on the financial industry as it covers regulation, compliance, global trading strategies and infrastructure, transactions between institutions, networks and firms, risk management and trading algorithms (Seddon & Currie, 2016). Furthermore, when looking into research focusing on the Finnish banking industry, there are only a few scientific articles found either from the 1990s or the beginning of 2000s.

Moreover, research on digital marketing performance using web analytics is limitedly studied (Järvinen & Karjaluoto, 2015). Moreover, typically web analytics as a method is used solely on ad-hoc basis instead of using it to strategic decision-making, which makes the actual benefits unclear (Järvinen & Karjaluoto, 2015) Therefore, this opens up a research focus to understand the usage better. In addition, previous studies state that more research is needed for web analytics outside the US market as well as focusing on more organizations in different sectors. Furthermore, more research is needed on how organizations utilize web analytics. Overall, there is a lack in analysing the usage and challenges of web analytics within the banking industry. (Järvinen & Karjaluoto, 2015 & Germann et al., 2012).

The objective of this research is to get insights on the major challenges with web data implementation process and the usage. Moreover, this master's thesis seeks to gain knowledge on the privacy aspect on the market as well as understanding how to address GDPR, which is the main legal framework in Europe and thus in Finland regulating people's personal data usage by organizations. Three research questions are attempted to be answered to reach the goal of the thesis. These following research questions were created to support the study of challenges in collecting and using web data in the banking sector:

1. What is the purpose of using web analytics in the banking sector?
2. What are the main issues when collecting web data in the banking sector?
3. What are the main challenges when analysing web data in the banking sector?

Qualitative approach was taken to address the research questions, since it allows to gain a deeper understanding and explanation on the issues studied (Eriksson & Kovalainen, 2008). Web analytics professionals on the banking industry were interviewed to gain understanding to the research questions. More in-depth overview of the methodology is presented later in chapter 3.

1.2   Research structure

This master's thesis is structured as follows. The contents of the five chapters are as illustrated below in figure 1. First, the research topic is introduced, and the research problem and research questions are presented. The master's thesis seeks answers to how web data is utilized within the banking industry in Finland and tries to pinpoint the main issues involving the processes.

The second chapter presents the literature review and theory supporting the subject of the thesis, which focuses on subjects such as privacy, data-driven decision-making, web analytics and GDPR.

The methodology is presented in chapter three. Qualitative research with semi-structured interviews were selected as the research method in this study to

gain knowledge and understanding on why and how things are done in web data collection and usage within the banking industry and what are the main issues (Arsel, 2017). Qualitative semi-structured interviews made it possible to gain knowledge on the specific issues in web data collection within the industry of banking. The semi-structured interviews took place between January and February of 2024. The interviewees were web analysts, implementation specialists and data analysts working within the banking sector and focusing on web data, both within B2C and B2B fields. The data collected was analysed with thematic analysis, which is a data analysing method for identifying and analysing patterns within the collected data (O'Gorman et al., 2015).

The results are presented in the fourth chapter in this thesis. Finally, the fifth chapter discusses the implications of this research and presents the conclusions and potential ideas for future research.

No content generated by AI technologies has been used in this assessment.
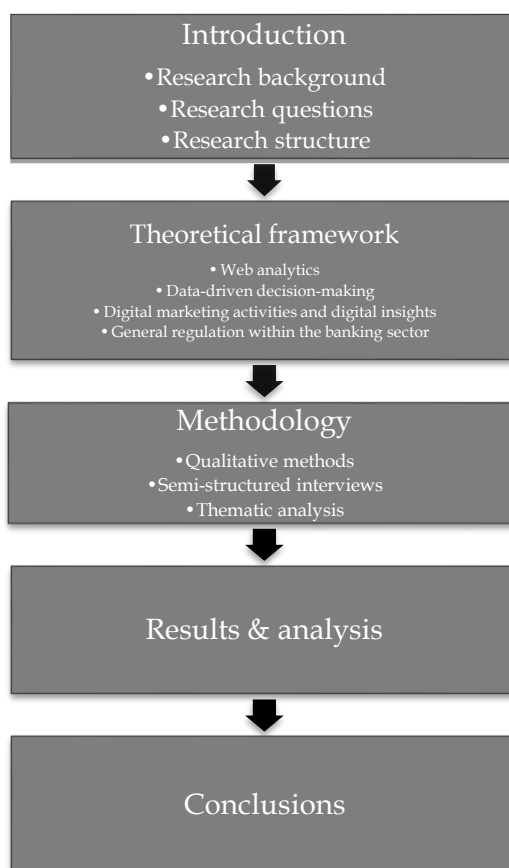

Figure 1: Research structure

2   LITERATURE REVIEW

This chapter explores the key concepts for this research that are web analytics, data-driven decision-making, digital marketing activities and insights and general regulation within the banking industry. Moreover, a presentation of the Finnish banking market is given to gain insights on the industry within the chosen research context.

2.1   Web analytics

Digitalization has created a need for companies to have their own websites. With an increasing number of websites also comes the growing competition to outdo each other on the market. To be able to receive a competitive advantage, businesses have to be able to track their websites to gain this knowledge (Kumar & Ogunmola, 2020). Web analytics as a term consists of both the tools used for web data analysis as well as the practice of analysing the collected web analytics data (Beasley, 2013). Web analytics is used to investigate and learn about user interactions on web platforms such as websites and mobile apps by picking up user behaviour as different dimensions that can be paired up with different data numbers, metrics, that will be able to be analysed for business insights (Beasley, 2013). The growing use of web analytics and its topicality stems from the easy accessibility of the tools in the market as well as the associated time and cost savings for utilizing it (Boegershausen et al., 2022).

Essentially any user action performed on a website can be picked up with a web analytics tracking code or a small modification in the page code, changed into data, and turned into customer insights. Within the European Union, however, tracking overall is regulated and will be discussed in the chapter below. The most essential information provided by web analytics are the visited webpage views, when those are viewed and in which order those are viewed (Beasley, 2013). This will collect a good overview on customer paths, how users have entered a website and where they will drop off. Moreover, devices used, screen resolution or for instance browsers used can be used for further analysis. In addition, demographics of the users such as gender, location or how often they are visiting can be used to guide the organization business to the right direction. Web analytics divides into web analytics analysis and web analytics implementation. Some metrics and dimensions are provided automatically by the web analytics tool, but any custom metrics, such as clicks or scroll depths or analysis for an ecommerce purchasing path need to be implemented separately. Beasley calls this click analytics (Beasley, 2013). The web analytics process, presented by Jensen and colleagues (2008), divides into four steps that showcases the process from data to insights:

1. Data collection
2. Processing the information
3. Developing the key performance indicators (KPIs)
4. Formulating an online strategy.

Data-driven decision-making will be discussed in more details in the next chapter.

The business research regarding web analytics stems all the way back of examining data in the technological changes from the first mainframe computers to personal computers, and all the way to the Internet (Seddon & Currie, 2016). However, even though web data nowadays is collected and used widely, the challenges related to the collection process has limited attention by researchers (Boegershausen et al., 2022). Various studies show the positive impact of web analysis towards organizational performance (Järvinen & Karjaluoto, 2015). Understanding the fundamental concepts of web analytics as well as having some data-analytic thinking will have positive impacts such as give insights on any data-oriented competitive threats or envision any opportunities (Provost & Fawcett, 2013). Research states that to be able to gain positive effects on web analytics, such as being able to monitor and optimize the site performance, strong skill level and good tools for web analytics within an organization are needed (Kumar & Ogunmola, 2020). Being able to measure data effectively will give decision-makers and managers radically more business insights and that knowledge can be transformed into improved decision-making and performance (McAfee & Brynjolfsson, 2012). Furthermore, efficient usage of web analytics is a crucial management tool for analysing online performance today but also in the future (Phippen et al., 2004). Research says that monitoring performance on the banking industry needs to be consistent since it provides knowledge for managers about improvement ideas, ideas to maintain services and highlights the focus areas (Bakar et al., 2016). However, even though banks offer majority of their products online nowadays, the majority of the customers are utilizing less than half of the provided products (Sandhu & Arora, 2022). Research states that at least within the Indian market, public banking customers are slower to integrate to new services online whereas private banking customers are moderately using the services (Sandhu & Arora, 2022). Therefore, monitoring and knowing how and which customer segments are utilizing the digital services is important within the banking sector through the use of web analytics and also on other markets. However, many practitioners are remaining sceptical towards the benefits measuring the performance and rely on intuition in decision-making rather than data, which is stated to needing to be researched more (Järvinen & Karjaluoto, 2015).

2.2   Data-driven decision-making

Decision-making is one of the most important factors influencing performance within an organization. Due to the evolving digitalization, companies now have access to large amounts of data to be utilized in their decision-making (Kemppainen et al., 2022). Moreover, using data analytics efficiently is an important management tool for organizations (Phippen, Sheppard, & Furnell, 2004). Provost and Fawcett (2013, p. 53) call data-driven decision-making (DDD) as: "the practice of basing decisions on the analysis of data rather than purely on intuition". On an ideal situation, digital marketers would select advertisements based on data, rather than based on the experience in the industry (Provost & Fawcett, 2013).

The data collected from customers are usually online customer behaviour or other important information about the users online that can be called as data analytics. Data analytics encases methods for measurement, analysis, predictions and managing any type of organizational performance indicators (Kemppainen et al., 2022). Data-driven decision-making is used to be able to make better and faster decisions on the organizational level (Kemppainen et al., 2022). Moreover, if the collected data is real-time and is combined with decision-making on a strategic level, it can make it possible to become more agile compared to competitors (McAfee & Brynjolfsson, 2012).

However, studies show that there are still some issues in being data driven on the organizational level. For instance, organizations pretend to be a lot more data driven than they are (McAfee & Brynjolfsson, 2012). Additionally, it has been observed that managers are still not fully trusting data for insights and as a tool for decision-making (Zaitsava et al., 2022). Even though there is a lot of emphasize on data-driven decision-making nowadays, it might be a challenge to use web analytics effectively. Kemppainen and colleagues (2022) list the biggest challenges relating to becoming successful in data-driven decision-making for companies as the lack of participation from the managers, lack of resources for analysing the data, budgets and skill levels, low quality data and poor web analytics tools. Moreover, a study highlights that the data also needs to be trustworthy, and if it has been implemented by humans, there is a risk factor in the trustworthiness (Zaitsava et al., 2022). However, data should be able to be still turned to decision-making if the skillset to do so is present.

It is not enough to only use web analytics to measure digital marketing performance, but to turn the analysis into action within the organizational management level (Järvinen & Karjaluoto, 2015). One of the most important element is to turn the data-analytic thinking from solely being behind the analyst to the whole organizational level, including managers (Provost & Fawcett, 2013). Organizations do not succeed simply since they have better data collected, but since they have management level teams that set clear goals and ask the right questions for the analysis (McAfee & Brynjolfsson, 2012). Moreover, the people

who are able to understand the problems should be brought together with the correct data as well as with the analysts who can help the management level with the analysis process (McAfee & Brynjolfsson, 2012). In addition, cross-functional cooperation is highlighted, meaning that it is important for the analyst to be able to discuss business language together with the leaders to enable them to reform the challenges towards a data aspect as well as to be able to help visualize the business data in a preferred way (McAfee & Brynjolfsson, 2012). Therefore, it is important to have the correct skill-level, create a good cooperation for the decision-making process within the organizations to be able to fully lead decisions with data on the high-level decisions as well as managerial challenges seem to be emphasized when it comes to data analysis and that there is still room for decision makers to develop their ways of working.

The data-driven decision-making process is commonly presented in phases of actions in literature, including collecting data, analysing data, and using data (Kemppainen et al., 2022). The same goes for the data-driven decision-making process proposed by Mandinach and colleagues (2006), which is presented in figure 2. However, Mandinach and colleagues (2006) additionally list six crucial actions in the decision-making process, which makes it more thorough than the others presented within the literature research (Kemppainen et al., 2022). The process shown in figure 2 showcases how organizations use the collected data and change it to information, knowledge and ultimately wisdom. The six actions that are listed as crucial during the decision-making process to create insights from the collected data are collect, organize, analyse, summarise, synthesize, and prioritize. More specifically, to be able to become data driven in the decision-making on an organizational level, the data needs to be implemented on the websites to be able to collect it first. Secondly, the data needs to be converted into important information, such as key performance indicators and then being able to perform actions with these, such as business decisions. Since the process is divided in different steps, it can be used to point out in which areas within the organization the data driven process is lacking and where it is performing. However, the complexity in organizations operating in different industries might need some more background information to really become aware on why some steps might not be performing on the organization, which is why the process itself is not sufficient enough to point out the data-driven challenges.
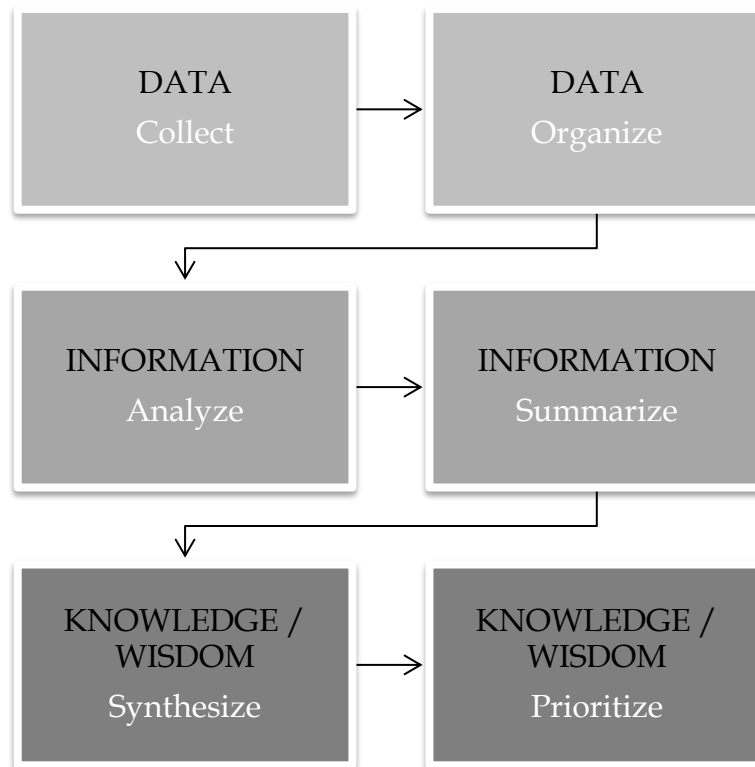
Figure 2: Data-driven decision-making process (Mandinach et al., 2006, p. 7)

2.3 Digital marketing activities and digital insights

The emergence of digital marketing has simultaneously created a way for digital tracking and web analytics (Lahtinen et al., 2023). This digital change made possible to analyse customers' digital behaviour on companies' websites, the keywords used for an online search or the interactions on a social media page (Lahtinen et al., 2023). Especially measuring online marketing has fuelled in growth, since it allows digital marketers to measure the effectiveness of their work with metrics such the number of users who reach the website and convert compared with the amount of money spent for acquiring these visitors (Beasley, 2013). It is emphasized that web analytics does not solely focus on measuring web traffic, but it will give insights on website optimization, and it is especially effective in unveiling new advertisement campaigns (Kumar & Ogunmola, 2020). Moreover, with the help of web analytics, the future campaigns can be directed to the correct target group, for instance, to an area that creates most traffic to the site (Önder & Berbekova, 2021). In addition, organizations nowadays have access to individual-level data that is used to different marketing purposes, such as customer relationship management, personalization and recommendations, one-

to-one messages and targeting ads as well as sophisticated customer segments (Salminen et al., 2018).

One of the most important cornerstones of effective digital marketing is the measurement of it (Lahtinen et al., 2023). It is stated that the measurement process should be integrated to the whole digital marketing activities and from the earliest stage as possible (Lahtinen et al., 2023). Järvinen & Karjaluoto (2015, p. 117), state that the web analytics data is used to gain insights on online customer behaviour, measure any responses for digital marketing actions and optimization for digital marketing actions towards enhancing the business performance. The process of web analytics metrics to be used in the digital marketing performance measurement should start by identifying the key performance indicators, also stated as KPIs, and differentiating them from other metrics used (Järvinen & Karjaluoto, 2015).

A research states that analysing the needs of individual customers is largely beneficial for the banking industry (Nguyen et al., 2022). Moreover, a 5% increase in customers retention rate within the banking industry can lead up to 85% of an increase in a bank's profit, which indicates that the industry benefits from analysis and understanding on the customer behaviour (Bakar et al., 2016). Banks have the ability to reduce the servicing costs, grow their revenue and increase knowledge about customer facing financial needs and issues when analysing their loyal customers (Nguyen et al., 2022). This process will, moreover, enhance their loyal customer base even more (Nguyen et al., 2022). Loyal customers tend to consume more products that the bank offers, present as a word-of-mouth spokesperson for the brand and stay loyal instead of looking for other banking providers on the market (Bakar et al., 2016). Therefore, there is a need to understand more of the customer data and customer behaviour in the banking sector.

2.4 The banking industry in Finland

The operating environments of banks are considered stable due to the capital levels and profitability levels within the organizations. They are controlled by central banks in each country as well as the European central bank for countries operating within the European Union markets (Kettunen, 2022). Banks usually operate to serve customers in large corporations, SME organizations, private banking and personal business. Banks offer their clients different online services in both online and mobile applications, accounts and payment facilities, credit and payment cards, loans and credits, savings and investment services, in some cases insurance and specific customer type of benefits.

At the end of 2022, Finnish market had 196 credit institutions (Kettunen, 2022). This entails deposit banks, mortgage credit banks, credit institutions, credit card companies and Municipality Finance Plc. From this, there were 12 banking groups operating within the Finnish banking industry (Kettunen, 2022). The

banking sector in Finland separates from other market due to its concentration. There are three main banks within the market holding the most shares: OP Financial Group, Nordea and Danske Bank had approximately 70% of the market share at the end of 2022 (Kettunen, 2022). OP Financial Group is the leading operator in the Finnish market with a market share of 34%, Nordea second with a market share of 24%, and Danske Bank with a share of 10%, also presented in figure 3 below (Suomen Pankki, 2023). In addition, another separation is the interconnection with the other Nordic countries through ownerships and investments, for instance Nordea is an example of this (Kettunen, 2022). Moreover, the industry in Finland differs from other countries in the big share of housing loans, small amount of non-performing assets and high-level profitability within the European market (Kettunen, 2022). The Finnish banking sector's profit in 2022 was approximately 6.6 billion euros and was mostly populated by household loans, corporate loans, and commission incomes such as asset management and the industry employ a bit under 20 000 employees at the end of 2022 (Kettunen, 2022).
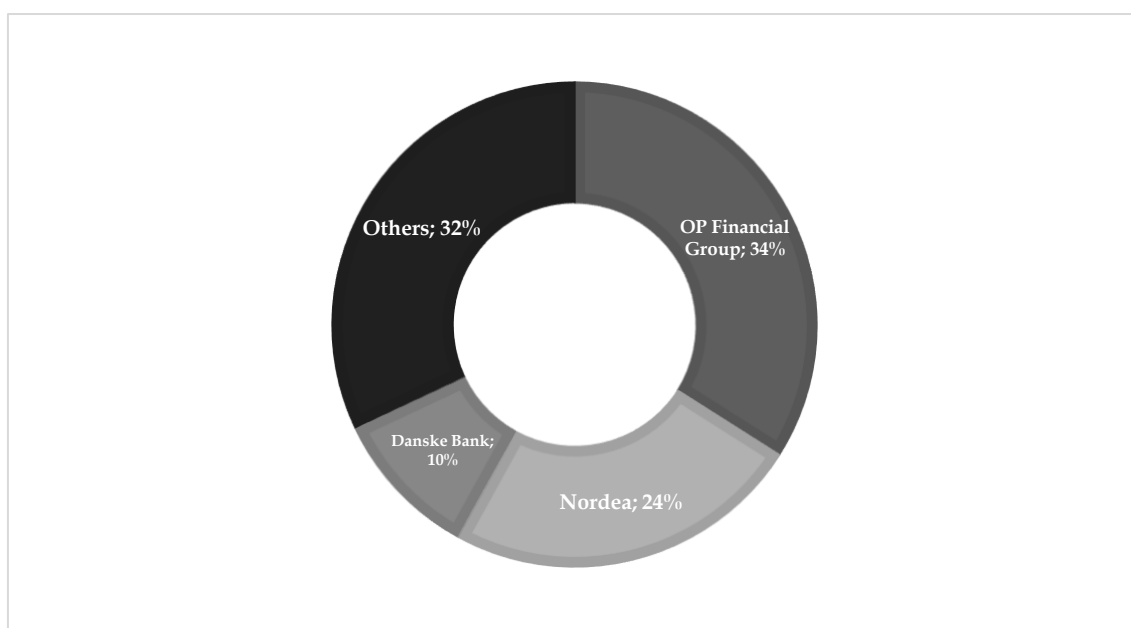
Figure 3: Market share within the Finnish Banking sector (Suomen Pankki, 2023, p.1)

2.5   General regulation within the banking sector

A study shows that banking customers are concerned about their privacy when it comes to their bank transactions (Bakar et al., 2016). As can be imagined, the regulation is generally quite strict when it comes to the banking sector. Banks contain a great amount of sensitive customer data since the sector operates closely with customer secure banking references. There are multiple processes and actions that are done differently comparing banking industries to any other

industry that will have actions for both employees within the bank as well as the customers. For instance, banks have their own strict processes for knowing their customers (KYC) and risk management as they are regulated and monitored by the Financial Supervisory Authority (Finanssivalvonta, 2023). Moreover, any new programs or processes needs to go through a quality risk assurance (QRA) process where the risk for the organization is assessed. In big organizations, it is common that all software and systems need to be approved by the privacy and IT departments within the organization (Scheuing & Niininen, 2022). For instance, within the banking industry, the approved systems are monitored and regulated and for instance a popular web analytics system, such as Google Analytics, is forbidden to be used. It is also common that all software and equipment used by employees will have regular mandatory updates regarding security. In addition, it is common to have strict password protections in place for laptops as well as educate the employees constantly regarding safety (Scheuing & Niininen, 2022). Data sensitivity is highlighted, and it is important to secure the accesses for data users with credentials and passwords.

Research states that security is the second most important dimension for banking customers and that customers want to feel safe with a bank they select to be a customer of (Bakar et al., 2016). This entails not only users' personal safety but also the possessions of the user. Furthermore, all the banking customers' resources should be kept safe and confidential, and the rules should apply to all employees within the banking institutions (Bakar et al., 2016).

In addition to privacy, customers highlight the service quality within the banking industry to be an important factor that can be used to increase customer trust, loyalty, and brand image of a bank (Christanto & Santoso, 2022). The most important legal framework governing web analytics regulation in Europe and therefore also in Finland affecting web data collection is the General Data Protection Regulation (GDPR) that is the focus in the next sub chapter.

2.5.1 GDPR

Since the digitalization has evolved and moved most of the interaction between users and organizations to digital channels, it has become even more relevant to operate securely and store data about internet users in a safe way. Organizations have started to utilize multiple different digital marketing tools that allow them to advertise, customize content as well as target different stakeholders online to enhance their business objectives. GDPR, a term that stands for General Data Protection Regulation, is a regulation that was launched on May 25th in 2018 and it affects business activities within countries in the European Union (Siegert et al., 2020). The aim is to create a stronger data protection that would give EU residents more control over their personal data since some organizations may misuse their rights for web data collection (Scheuing & Niininen, 2022). Siegert and colleagues (2020, p. 16) state that: "The GDPR is designed to unify all data privacy laws across Europe while protecting the information of all EU residents against

information leakage and privacy breaches." Therefore, it is a regulation used to protect customers' personal rights with online tracking and data usage.

This regulation of GDPR aims to follow and ensure that personal data for website users is handled with privacy, security, and confidentiality, especially when processing personal data. Moreover, it focuses on a more secure way of collecting, storing, and processing data (Siegert et al., 2020). The regulation entails that the data used on the Internet needs to be used legally and for an actual legal purpose. In addition, the information needs to be valid and up to date and the information need to be kept and stored in a safe place. Moreover, the website user can ask the data to be altered if it is not correct or missing something as well as ask for all personal data to be deleted. (Siegert et al., 2020).

Every EU country has its own public authority taking care of monitoring that all website providers are obeying the regulation and for Finland this authority is Traficom (Peukert et al., 2022). As the process is monitored, there have also been unfortunate mis usages that have led up to either organizations dealing with a brand image hit or even bankruptcy. An example is Cambridge Analytica, that used personal information without consent from Facebook for targeting and personalization that led up to a 500,000 dollar fine as well as a bankruptcy for the organization (Van Oojien & Vrabec, 2018).

However, even though GDPR was launched six years ago, consumers are still having a hard time fully understanding the privacy statements that explain how data is used and what is collected, which is creating a gap between the data collector and the data subject (Van Oojien & Vrabec, 2018). When GDPR was launched, organizations were obligatory to provide a privacy statement where a website user can see what technologies and providers are used to collect personal data about users, what is the purpose of the usage and the duration of the usage (Scheuing & Niininen, 2022). Research shows, that the more users realize about what occurs to their personal data, the less info they can process and understand (van Ooijen & Vrabec, 2018). In addition, organizations cannot rely that the large part of users on their website will understand the privacy policies they are presenting, which is why it is crucial to make sure these texts are written in a clear language and that those are easy to understand (van Ooijen & Vrabec, 2018, Scheuing & Niininen, 2022).

As GDPR enhances individual control, Van Oojien and Vrabec (2018) present three stages in the consent-based data processing regarding the individual control within the web data era: The information receiving stage, the approval and primary use stage and the data reusage stage, presented in Figure 4. In the first stage, the data collector will provide information to the user about how the data is used as well as the privacy policy. In the second step, approval and primary usage, the user makes a decision about whether he or she wants data to be collected or not. In the step a user either accepts tracking and cookies in the consent banner or declines them. The third stage, reusage stage, contains the secondary data usage by the data collector. Individuals can access the privacy

policy page and see how the data is used as well as have the option to revoke consent. (Van Oojien & Vrabec, 2018).
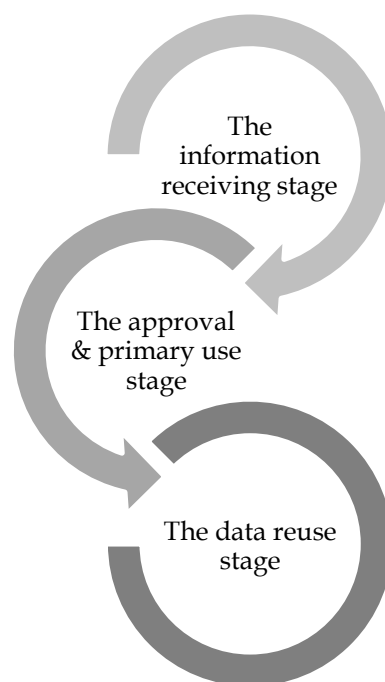
Figure 4: Individual control stages for data usage (Van Oojien & Vrabec, 2018, p. 93)

Big organizations that handle a lot of customer data need to take a risk-based approach when dealing with customer data collected (Scheuing & Niininen, 2022). GDPR requires a more protective implementation method if there is a higher risk for data subjects, for instance in cases where a big organization would have more data than others or if the data would include any sensitive data about customers, such as within the banking industry. Therefore, data pseudonymisation and encryption processes need to be in place as well as safe tools and software need to be utilized at minimum (Scheuing & Niininen, 2022). Moreover, it is important to be able to recover quickly for any security incidents and have processes in place for reporting these (Scheuing & Niininen, 2022).

2.5.2 Personal data and consent

For any type of data collection, GDPR raises awareness of personal rights and therefore, everyone who handle data needs to be aware of GDPR rules (Siegert et al., 2018). It is stated in the General Data Protection Regulation that user consent is needed if the website is using any web technologies that could identify a user, such as using fingerprinting technologies, collecting cookies from the website visitor, tracking scripts or collecting images. (Peukert et al., 2022). Therefore,

without consent given in a website cookie consent banner, no data is enabled to be collected for this website visitor. As users have the possibility to opt-out from tracking, GDPR also protects users with how their user data is implemented. Therefore, every user has additionally rights about knowing what personal data information the organization has collected about the specific person as well as knowing how the information is used (Siegert et al., 2020).

Personal data means all data and information that relate to directly or indirectly identified or identifiable natural person, that can also be called as the data subject, for instance through name, social security number or other identification number, location of that data subject (Scheuing & Niininen, 2022). Moreover, Siegert and colleagues (2020) identify personal data as any type of information, it needs to relate to a person by either by the content, purpose or result, the person needs to be identifiable and the person itself must be a living individual.

As the regulation brings customers with a lot of safety, it creates a need for organizations to pay attention to how they are acting and obeying to the regulations. In addition to changes done within the cookie consent banners and privacy policy page, organizations need to be aware of the customer data they are collecting in the future. Web analytics implementation teams need to be educated about data collection and be aware not to collect any personally identifiable information (PII data) about their customers if it will not be stored in a secure form. Therefore, this kind of PII data needs to be careful not to be implemented in companies' web analytics systems for instance (Tietosuoja.fi, 2022).

2.6   Summary of the literature review

This chapter introduced multiple key concepts for this study, which are web analytics, data-driven decision-making, digital marketing activities and insights, and general regulation within the banking industry. The main objective with the literature review was to present the key concepts, which are relevant for this research and ultimately support finding the answers to the research questions. The most important elements from the literature review, which are intended to be explored with the interview are presented in the table 1 below.

Key concept	Researchers	Key findings
Web analytics	Kumar & Ogunmola, 2020 McAfee & Brynjolfsson, 2012 Phippen et al, 2004 Bakar et al. 2016 Järvinen & Karjaluoto, 2015	• To be able to gain positive effects on web analytics, strong skill level and good tools within an organization are needed • Being able to measure data effectively will give decision-makers and managers radically more business insights and that knowledge can be transformed into

		• improved decision making and performance • Efficient usage of web analytics is a crucial management tool for analysing online performance today but also in the future • Monitoring performance on the banking industry needs to be consistent since it provides knowledge for managers about improvement ideas, ideas to maintain services and highlights the focus areas • Many practitioners are remaining sceptical towards the benefits of performance measurement and rely on intuition in decision making rather than data
Data-driven decision-making	Provost & Fawcett, 2013 Kemppainen et al. 2022 McAfee & Brynjolfsson, 2012 Zaitsava et al. 2022 Järvinen & Karjaluoto, 2015	• On an ideal situation, digital marketers would select advertisements based on data, rather than based on the experience in the industry • Data-driven decision-making is used to be able to make better and faster decisions on the organizational level • Organizations pretend to be a lot more data-driven than what they essentially are • Managers are still not fully trusting data for insights and as a tool for decision-making • The biggest issues relating to data-driven decision-making are how to be able to utilize data effectively on the strategic level, lack of resources, budgets and skill levels, low quality in data collected and poor web analytics tools. • The data needs to be trustworthy • It is not enough to only use web analytics to measure digital marketing performance, but to be able to turn the analysis into action within the organizational management level • Organizations do not succeed simply since they have better data collected, but since they have management level teams that set clear goals and ask the right questions for the analysis • Cross-functional cooperation is highlighted, meaning that it is important for the analyst to be able to discuss business language together with the leaders to enable them to reform the challenges towards a data aspect as well as to be able to help visualize the business data in a preferred way
Digital marketing	Kumar & Ogunmola, 2020	• Web analytics does not solely focus on measuring web traffic, but it will give

activities and insights	Lahtinen et al. 2023	• insights on website optimization, and it is especially effective in unveiling new advertisement campaigns • One of the most important cornerstones of effective digital marketing is the measurement of it and it should start from the earliest stage possible • The process of web analytics metrics to be used in the digital marketing performance measurement should start by identifying the key performance indicators
General regulation within the banking industry	Bakar et al. 2016 Scheuing & Niininen, 2022 Van Oojien & Vrabec, 2018 Siegert et al., 2018	• Banking customers are concerned about their privacy when it comes to their bank transactions • In big organizations, it is common that all software and systems need to be approved by the privacy and IT departments within the organization • It is common that all software and equipment used by employees will have regular mandatory updates regarding security, have strict password protections in place for laptops as well as educate the employees constantly • Research states that security is the second most important dimension for banking customers and that customers want to feel safe with a bank they select to be a customer of • Consumers are still having a hard time fully understanding the privacy statements • Big organizations that handle a lot of customer data need to take a risk-based approach when dealing with customer data collected • Data pseudonymisation and encryption processes need to be in place as well as safe tools and software need to be utilized at minimum • everyone who handle data needs to be aware of GDPR rules

Table 1: Key findings of the literature review

3 METHODOLOGY

This chapter presents the methodology used for this study, including the chosen research approach, data collection, sampling of the research group and the pre-analysis of the data collected. Methods can be divided into data collecting methods and analysing methods (Eriksson & Kovalainen, 2008). A Methods map can offer a clear view of the structured approach and can help identify the options for the research design (O'Gorman et al., 2015). Figure 5 showcases the methods map and design selected for this study research.

The scope of the conducted research is explorative, which is why qualitative approach was selected as the method. The purpose of the research is to explore how the professionals within the banking industry are navigating the challenges regarding legislation and other issues within the field related to collecting and using web data.
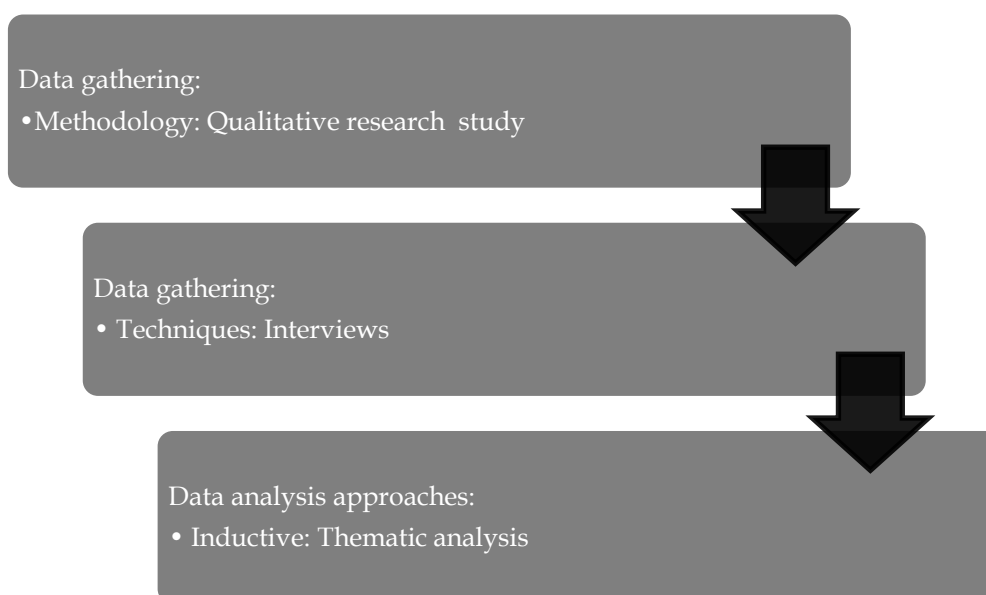

Figure 5: Methods map (O'Gorman et al., 2015, p. 51).

3.1 Qualitative research approach

A qualitative method aims to explain and gain a holistic understanding of the issues studied by producing and interpreting through cultural meanings (Eriksson & Kovalainen, 2008). Qualitative approaches can be a case study, focus group research, grounded theory research, ethnography, narrative, and

discursive research as well as critical, feminist and action research (Eriksson & Kovalainen, 2008).

A qualitative research study was selected to gain knowledge through qualitative methods, to understand how data is collected within the banking industry. Since there is not a lot of previous knowledge with the selected theme, the qualitative method is justified as the methodology in this case and semi structured interview was selected as the data gathering method to gain a holistic understanding on the subject in the research.

3.2   Data gathering

The data collection in this study is done by semi-structured interviews to gain a holistic understanding on web data collection and usage within the banking industry and detecting the main challenges.

In semi-structured interviews, the order of questions changes based on the conversion taking place as well as the wording of the questions will shift to personalise them to each interviewee (Arsel, 2017). The questions in semi-structured interviews are open ended and do not contain any response alternatives (Hair et al., 2015). Therefore, an interview structure was created guiding the process but giving enough flexibility to change the order of the questions throughout the process but also to keep the dialogue ongoing if needed to gain more knowledge as the process went on.

3.2.1   Interview questions

The interview questions were drafted to support these following research questions to gain knowledge and receive data on these subjects:

1. What is the purpose of using web analytics in the banking sector?
2. What are the main issues when collecting web data in the banking sector?
3. What are the main challenges when analysing web data in the banking sector?

The 16 interview questions are presented in appendix 1. With the help of these questions, this thesis seeks to gain understanding of the banking industry and the challenges that apply within the web data collection and usage.

The research questions as well as the interview questions were formulated based on the research problem presented in chapter 1.1. as well as the literature review in chapter 2 that discussed the main challenges amongst the key concepts.

The interview process started by sampling out the interviewees that are working closely with either web data collection, usage or privacy and would therefore have the best knowledge to answer the questions of web data usage in the banking industry. Since semi structured interviews were chosen, it allowed

to ask for more questions from the interviewees later during the process to gain better understanding of the whole subject. The interview questions were tested with two interviewees to see if the structure of the questions presents a good coherent implementation and later those were modified slightly with better wording for the final interviews.

3.2.2   Interviewees and sampling

The process or technique used for selecting the most suitable sample for the research conducted with the correct and most suitable characteristics or parameters is called sampling (Adams et al., 2014). Purposive sampling was used, where the sample is selected based on the most fitting characteristics for the conducted research (Adams et al., 2014).

Table 1 presents the interviewees collected for the study. The sampling technique to get the correct interviewees contained gaining understanding of the banking organization structures and reaching out to the people who work directly either with web data analysis, web data implementation or privacy and either in business insights roles or very technical roles in different areas within a banking organization operating within the Finnish market. The interviewees were approached and asked to participate in the interviews via email or within Microsoft Teams chat functionality. Altogether 14 invitation attempts were made from which 10 interviewees accepted the invitation. All of the interviewees that took part in the interviews have many years of experience on web data and are banking professionals. Thereby, the interviewees were able to provide insights on the research questions set for this study from different angles within the industry. The demographics of the interviewees consisted of both male and female and different age groups, which represents nicely a wider sampling of interviewees on the demographic level.

As soon as the interviewees confirmed their participation, an interview was booked that contained the structure and main points for the upcoming interview for the interviewees to be able to prepare for the sessions. When the interviews began, it was clearly presented that the data collected within the interviews as well as any identifiable information about the interviewees would not be mentioned in the study conducted, which was an important tactic to create a safe environment for the interview setting. Altogether ten interviews were conducted, and both held and recorded in Microsoft Teams. The duration for each interview was approximately 30 minutes. This was sufficient to go through the main subjects in the interviews but also to circle back to any questions that needed anything to be added. The interviews were held both in Finnish and in English and ultimately all materials were translated into English.

Interviewee	Gender	Age group	Web analytics experience	Occupation	Main operation area	Interview duration time
Interviewee 1	Male	40-50	9 y	Senior implementation specialist	Nordics	25min 03sec
Interviewee 2	Female	30-40	6 y	Senior web analyst	Nordics	35min 30sec
Interviewee 3	Female	20-30	4 y	Web analyst	Finland	32min 10sec
Interviewee 4	Male	40-50	10 y	Senior data analyst	Nordics	28min 22sec
Interviewee 5	Female	30-40	2 y	Web analyst	Finland	38min 03sec
Interviewee 6	Male	30-40	2 y	Privacy specialist	Nordics	26min 33sec
Interviewee 7	Female	30-40	5,5 y	Implementation specialist	Nordics	36min 05sec
Interviewee 8	Male	40-50	12 y	Senior data analyst	Finland	34min 58sec
Interviewee 9	Male	20-30	1,5 y	Implementation specialist	Nordics	26min 20sec
Interviewee 10	Female	20-30	8 y	Senior implementation specialist	Nordics	25min 15sec

Table 2: Table of interviewees

3.3 Data management

The data management chapter discusses how the research data has been managed, stored, and used within the whole research period (January – March 2024). The data managed and stored should be handled differently, based on if it includes sensitive data or not (JYU, 2023). This research data does not include any identifiable sensitive data that could be pointed to a specific person and all the names of the interviewees have been pseudonymised into Interviewee 1, 2, 3 and so on when transcribing from the interviews into Microsoft Word. However, as the interviews were recorded with Microsoft Teams, the participants first names and voices can be seen and heard from the recordings and thus they could be identified. Therefore, safety measures were used for the interview recordings by using encryption and storing them in the personal university's Drive folder.

When the interviews were conducted, it was clearly presented that no identifiable information will be shared in the actual research.

When the research process was ongoing, the whole master's thesis was shared with the opponents and thesis supervisor within the university's shared Drive folder. This version did not include any sensitive data and could be stored without encryption.

3.4   Data analysis

The idea for the analysis process is to try to locate patterns within the collected data. Some of these patterns can be located for instance in the answers of the interview questions (Belk et al., 2013).

The interviews conducted were analysed with thematic analysis. O'Gorman and colleagues (2015) recommend six steps to perform thematic analysis:

1. Familiarizing yourself with the data
2. Generating initial codes
3. Searching for themes
4. Reviewing themes
5. Defining and naming themes
6. Producing the report

Therefore, the process with data analysing with thematic analysis should start with getting to know the data that is the subject for analysis. O'Gorman and colleagues (2015) mention that the data content should be read through at least once, depending on the complete size of the content. The step proceeds with getting all the notes organized based on codes they consist of. Bryman and colleagues (2007) mention that the coding is a crucial step for the analysis process and must be done consistently. The third step continues with combining these codes into broader patterns. The fourth step oversees if these patterns are similar and can be collapsed into each other and it follows with the final theming phase. Even though the semi-structure interviews were somewhat structured to focus on specific themes to guide the interviews forward, the idea for the thematic analysis was to try to locate new themes among the interview answers, which were first recorded in Microsoft Teams and furthermore analysed and coded in writing.

As the idea is to start the analysis process as soon as any qualitative data has been collected, the analysis process was kicked-off straight after the first interview and continued as soon as new interviews were held. As the interviews were translated and written down, the coding process could be started and ultimately when the data was gathered to themes, the most important topics were highlighted. Microsoft Word was actively used to write down the interview answers and for the entire coding process where topics

were searched among the interview answers. As soon as the process was finalized and all the themes were mapped, the final results could be drawn.

4   RESULTS AND ANALYSIS

This chapter presents the research findings of this study by looking into the data collected from the semi structured interviews. First, the main research findings are presented and then continued with the results focusing on the four macro themes that emerged with the thematic analysis. This is presented in more details in the main research findings below.

4.1   Main research findings

The main codes received with thematic analysis are presented in figure 6 below. These codes were highlighted in the interview responses and ultimately grouped into more specific patterns of 10, as illustrated in the middle section of figure 6. Eventually, the presented topics were further grouped onto four macro themes that are used as the main headlines below presenting the results. These headings and main themes grouping the research findings are privacy, data-driven decision-making and web analytics and skills.
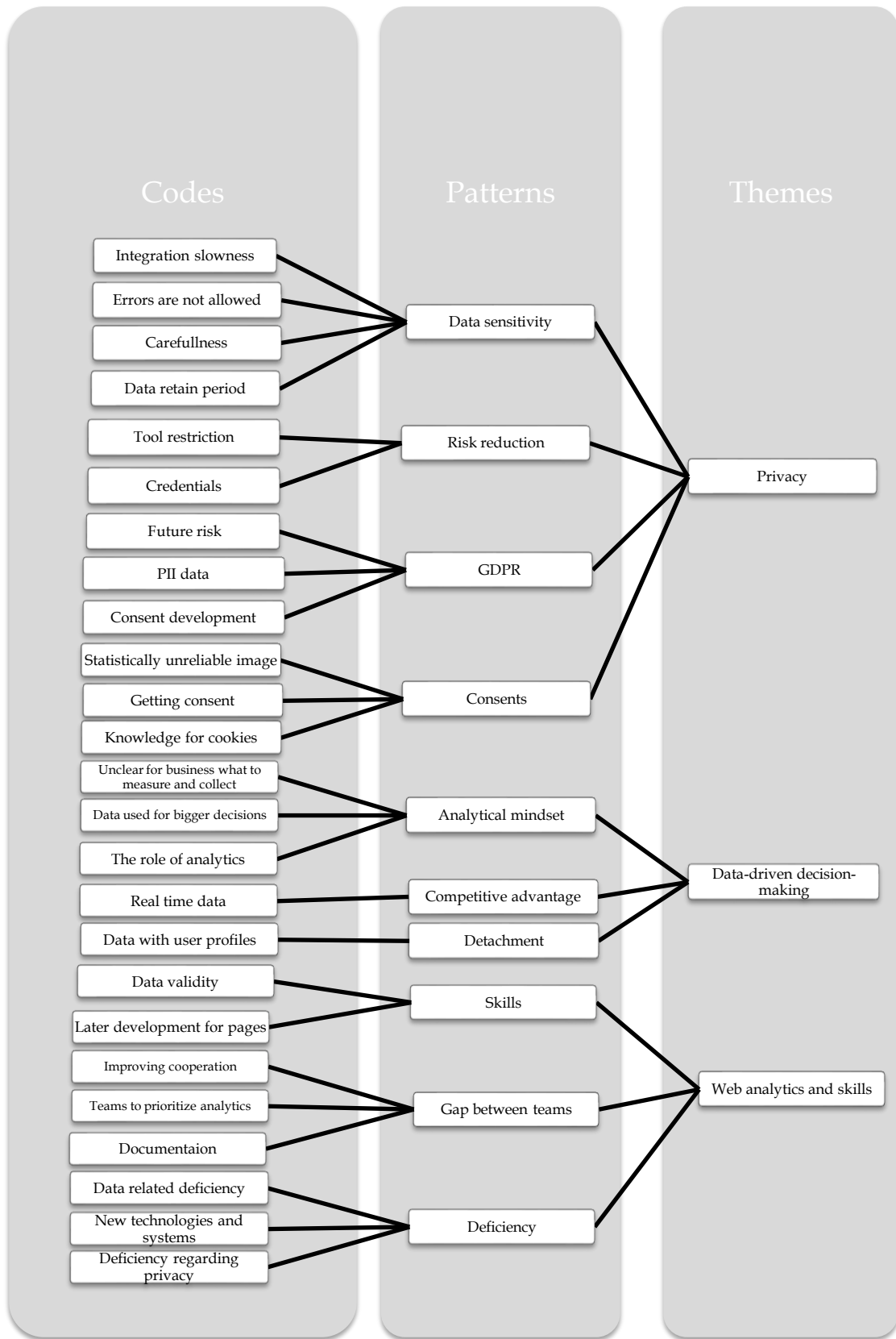
Figure 6: Defining themes

Most of the research findings are presented with the help of the data-driven decision-making process map with six steps presented in the literature review by Mandinach and colleagues (2006). The steps include different phases towards becoming data-driven in the decision-making process, which relate to collecting data, organizing data sets, analysing the information, summarising the findings, synthesizing the knowledge into wisdom, and prioritizing the findings for further development or decision-making. This will allow to visualize on which phase in the process map each of the research findings are currently at. Something that arises from the interview answers is that the banking industry in the Finnish market still has some development to do within the process map by Mandinach and colleagues (2006) towards becoming fully data driven. When analysing the different themes, it can be detected more specifically on which of these process steps the path is inoperative.

The main results of this study support the findings as literature review already presents for the biggest challenges in data collection and usage, which indicates that the same challenges persist also in the banking industry. Within the banking industry, customers' privacy is highlighted as it was mentioned in the research as one of the most important factors for customers and employees for banks, which is why it will be presented first within this results chapter before moving on to the results regarding data-driven decision-making and web analytics knowledge and skills. In addition, the findings show some issues regarding operations in big organizations. The result findings will be presented in more depth on their own subheadings below.

4.2   Privacy within the banking industry

As privacy is highlighted as an important theme already in literature review for both customers and organizations within the banking industry, this will be presented first in the results. The privacy results will be divided into privacy within the banking industry, GDPR and consents based on the emerging sub themes within the interview answers. As this research seeks to gain knowledge on the biggest challenges in web data collection and usage, it is important to know if privacy affects web data collection and usage within the industry. Therefore, privacy was also present within the interview questions to gain information if it is creating some challenges within the banking industry. Interviewees were asked if privacy and regulations affect the data collection or usage within their organization and whether it can be seen as a benefit or a challenge.

Privacy was mentioned as one of the most important challenges in data collection and usage within the banking industry by the interviewees. When challenges were asked solely focusing on web analytics collection and usage, privacy was not mentioned as a challenge. Therefore, we can assume that privacy

is an important and emerging theme when research focuses on the banking industry and in this case the collection and usage of web data within the industry.

"The regulation like GDRP is of course the same for everyone but it still might be that within the financial & banking industry the regulation is taken more seriously and thus there could be a bit more focus on privacy, regulation, and security. But I wouldn't say that these are special challenges that are only applicable for banking industry." *(Interviewee 4).*

One of the first interview questions asked about the tools used for web data analysis. The organizations within the banking field are using Adobe Analytics due to privacy reasons. This indicates that the tools used for web data analytics within the industry are regulated. When asked further, the tool itself seems to serve its purpose and does not cause any challenges even though it is regulated. What was mentioned, however, is that other supporting tools could be beneficial, which are not allowed to be used due to privacy regulations within the industry.

"Google Analytics is usually used within organizations, but due to privacy and them not telling where the data is stored, the only option was to use the Adobe toolkit." *(Interviewee 7).*

"I do not see any specific issues with web data within the industry we operate in. However, the biggest challenge could be strict law interpretation that will not allow the same products to be used with web analytics as normally would be common to use such as Google Analytics, big query and so on." *(Interviewee 8).*

There seems to also be a strict, maybe even cautious way of working relating to digital marketing activities, which affects not only the tools chosen to be used within the banking organization but also some digital marketing actions performed. For instance, heatmaps, session recording tools, optimization and being able to combine data with other tools are not as straight-forward operations within the banking industry due to privacy as in other businesses.

"Moreover, some heatmap or session recording tools might not be dared to use." *(Interviewee 8).*

"Strongly regulated industry creates a challenge for targeting and segmenting; there would be great potential to be able to sell specific products or target specific customer groups, but it is not allowed." *(Interviewee 5).*

"We might be a bit too careful when it comes to tracking and being able to combine web data into other programs and tools has not been made simple." *(Interviewee 3).*

Moreover, due to privacy and careful approach within operations, data sensitivity is highlighted in the interview answers. Time-consuming technology

adoption and having credentials to all tools used were mentioned in the interview answers related to privacy within the industry. As mentioned in the literature review, big organizations that collect a lot of customer data need to take a risk-based approach when dealing with customer data (Scheuing & Niininen, 2022). Therefore, the time-consuming technology adoption relates partially to difficult processes in big organizations, however, partially it relates to the privacy and data sensitivity within banking where trustworthiness is important for the brand. Due to the fact that consumers rely on banks being safe, it becomes a big image hit if anything were to happen and the bank would lose its credibility for safety. Unfortunately, this occurred for S-Pankki in 2022 and they were forced to answer a lot of questions regarding the safety for the customers (Yle, 2022). In addition, the interview responses highlight that credentials are needed for all tools as a safety measure.

"Data sensitivity is emphasized within the banking industry. One needs to be extra careful with data accesses and tracking. For instance, pixels are not allowed to be used in the industry." (Interviewee 2).

"There is no room for errors in this industry, which makes new technology adoption very time consuming and hard." (Interviewee 1).

"As an analyst it is sometimes difficult not to be able to get credentials on some tools that could help in understanding the setup. This was not an issue in my previous job within another field of business." (Interviewee 3).

"Credentials are needed on all tools, which is also a privacy aspect within the industry." (Interviewee 9).

Furthermore, something that is brought up in the responses is that data storage period, which is mentioned to be shorter than in other industries due to privacy. Having the data available for shorter period of time, could cause issues in longer time period analysis and hence privacy affecting the quality of web data analysis.

"Web data is stored for less amount of time than in other organizations outside the banking industry, which make longer period web analysis harder to accomplish." (Interviewee 9).

In addition, some further ideas for improvement were mentioned focusing on enhancing the privacy processes. For instance, privacy being designed and built to all systems to enable better privacy processes as well as creating better processes for data collection regarding privacy. These points would enhance privacy aspects operating smoother and not causing them to become blockers of any operations.

"One improvement point would be ensuring that privacy is built into all the processes ("privacy first design") and thus ensuring that privacy related questions are taken into account early on and that the privacy related questions don't become blockers." *(Interviewee 4).*

"Setting the data collection rules should be happening faster. Unfortunately, many times the work for analysing is blocked since teams are unsure not about if they can do it but if it is allowed. If this would be faster, we could forward our work with the same pace as other industries." *(Interviewee 2).*

4.2.1   GDPR

One of the biggest regulations for data collection for the companies operating within the European Union is the GDPR. It stands out within the interview results when discussing about challenges overall for web data collection as well as asking directly about how the interviewees think privacy is affecting the collection process. The topic did not, however, solely stand out as the main challenge probably due to professionals in the industry already being used to it by now, as the regulation was launched in 2018. Instead, the privacy within the industry was highlighted more. Relating to privacy and collecting data as securely and maybe even a bit too carefully, the following was suggested as an improvement and a comment about the future becoming even stricter:

"My development idea for the banking industry would be enhancing the cookie collection or collecting anonymous data from users who accept statistics or necessary cookies." *(Interviewee 5).*

"In the future, the main issue will be the removal of cookies and the restriction to use third party cookies." *(Interviewee 2).*

One of the interview answers mentions a risk of collecting personally identifiable data (PII data) from customers, which is not allowed to be done with the web analytics tools used within banking. Therefore, it creates a challenge for everyone working with data collection, such as the analysts and the development teams to fully understand the privacy aspect and what they can track and what should be left out from the tracking due to security reasons.

"On top of my mind there are no bigger issues within the banking industry in using or collecting web data, but if we are not being careful, it could be something relating to any PII data, that might be passed in the URL/Link and interactions on pages have confidential and PII data in them." *(Interviewee 6).*

Moreover, something that was mentioned is that the regulatory aspect on interpreting GDPR seems different based on who the authority in each country

is. This is an interesting point to be aware of that even though GDPR states something, it might be interpreted differently in different European countries by the authority.

"Traficom which is the authority in Finland seems to interpret GDPR rules and regulations a bit more strict than other Nordic countries, for instance Denmark and Norway." (Interviewee 7).

4.2.2   Consents

Many interview answers directly focused on GDPR consents. Overall, the interview answers state that unless the organization gets a consent to collect data, the data is not coming in, creating the number one challenge on the web data collection.

"In an ideal world, we would be able to collect all the web data." (Interviewee 8).

"The main challenge is getting consent - if we don't have the correct consent, we won't collect data." (Interviewee 6).

Moreover, to be able to work efficiently and be able to rely on the data collected, the web analyst needs to be aware of how the consent rules and banners have been implemented and have an overall understanding on GDPR.

"Getting consent is diverse on different pages and the web analyst needs to be aware on how the consent tool operates as well as stay up to date on the European regulations and any updates relating to the consent tool." (Interviewee 10).

"Depending on the analytics tool used, the data can be collected differently. Some statistical data is allowed to be collected anonymously and obeying GDPR, but this might be an ethical issue if we are allowed to do so if a user asks not to be tracked – even though the user data would be anonymous and without any PII data." (Interviewee 5).

Quite many of the interviewees, especially the ones working with analysing the collected data, mentioned consents as the most significant challenge from the angle that it becomes hard to know what sample size gives a reliable image of data that can be used for data-driven decision-making.

"Collecting data only from users that accept in the cookie consent banner might give an unreliable and statistically wrong image for instance if the data collected is not large. It is also hard to know what amount of data is sufficient to use it for decision-making." (Interviewee 9).

"Depending on the site or application volume the data analysis done based on solely accepted cookies can give a totally unreliable and statistically wrong image. Especially if

the sampling is small, the possibility for wrongly analysing data is large." (Interviewee 2).

"It can be hard to know what size of a sample gives a reliable image for data-driven decision-making of the users accepting cookies since the sample is never every user on the page." (Interviewee, 3).

However, it appears that the Finnish banking industry is perceived trustworthy and therefore the user acceptance rate for the consent banner is higher than for other industries. All of the professionals interviewed for this research work in big organizations within the banking industry, meaning that overall, the websites they are analysing already receive more visitor traffic compared to smaller organizations. Therefore, the amount of data collected from users should be sufficient and can be utilized for trustworthy analyses.

"Some smaller organizations might have issues in getting enough data since the cookie consent has emerged. However, luckily the banking industry is regarded as a loyal and trustworthy industry, which means that the consent acceptance rate is quite high." (Interviewee 7).

"Luckily, we have so many hits that even though users can decline all data within the consent banner, we still have quite a lot of data to begin with. Banks are luckily considered as trustworthy operators, which means that on average, we get more people to accept data collection than within the other industries." (Interviewee 9).

Moreover, an improvement idea was mentioned for consents relating to communication. Better communication and storytelling towards customers were suggested in the copytext in the consent banner to improve the consent acceptance rate.

"What could be done better relating to consent banners and tackling the users to accept the consent with better percentages is better communication towards the customers as well as making it transparent for customers to see what benefit it brings to grant access to their data." (Interviewee 1).

4.3   Data-driven decision-making

As defined earlier, data-driven decision-making is the process of basing business decisions on data rather than on pure intuition (Provost & Fawcett, 2013) and within organizations is preferred since it can offer great amounts of benefits, such as competitive advantage on the market. Different key performance indicators (KPIs) and goals for the operations should be set and measured. One of the main realisations in this research is that even banking industries, where data was

introduced to decision-making as one of the first industries, seem to struggle with basing their decisions on data every time. It seems that banks are introducing new products and services to the market without fully reflecting about the KPIs and metrics for tracking, which could indicate that also skillset should be improved to know what can be done and how.

"The main challenge in banking in working with data, in my opinion, is the same as in many other industries. It has been learned that web analytics needs to be there but usually it is there only seemingly and will not guide the development. Usually, web analytics is taken into planning in a too late stage." (Interviewee 8).

It appears that the data collection is not comprehensively planned and that the business teams do not know what metrics to measure and follow but instead are tracking everything. This suggests that as the comment above mentions that they know that it is important to have tracking, but solely that is insufficient for data-driven decisions.

"It is important to know what to collect business wise and what not to and have an analytical mindset." (Interviewee, 2).

"Overall, a big challenge within organizations is to realise what (web data) to collect and what not to collect. If the products or services that are developed have clear goals, the web analytics can be built to support those goals. However, if the product goals would be on a level of: "nice to have", it will be insignificant on a data-driven decision-making level." (Interviewee 8).

"Data is definitely collected but not always used for bigger decisions. Moreover, often the wish is just to have everything tracked, which is not ideal businesswise. Therefore, there is still a lot of room to develop within the data driven aspect." (Interviewee 7).

"One of the biggest challenges with web data collection is that business does not have a realistic picture on what they want to measure and collect." (Interviewee 3).

As mentioned previously in the literature review, some managers are still sceptical towards web analytics and leading with data and would still prefer to trust their gut feeling instead (Järvinen & Karjaluoto, 2015). The interview answers relating to data-driven decision-making also indicate that some decisions are based on a feeling instead of numbers and something that is used to describe analysis work within a few answers is the word rushed. Perhaps the data-driven aspect is not fully utilized and turned into smart KPIs to guide the business forward.

"Sometimes business leaders forget about data existence and decisions are made based on a gut feeling." (Interviewee 6).

"There is also some kind of a rush and things are not analysed enough with time before the decision has been made. Unfortunately, this is an issue that persists still. (Interviewee 7).

"Sometimes it feels that the teams who are developing a new product and also adding tracking do not really think about the business aspect and KPIs when the tracking should be added, and it feels a bit rushed." (Interviewee 9).

It seems that data is being collected nicely by implementation specialists and tracking codes implemented with tag manager systems and there are web analysts doing the process of mapping and analysing the data and ultimately turning it into reports and summaries. When analysing the process with the data-driven decision-making map presented in the literature review, it appears that the final parts within the mix are missing, that would turn the summarized data into insights that will lead the development as well as guide the business. This is presented below in the figure 7 to present the area affected by this challenge.

Collect › Organize › Analyze › Summarize › Synthesize › Prioritize

Figure 7: Data-driven decision-making process challenges

"Numbers and quantities are easy to look at but solely based on those it is hard to develop anything so that web analytics would have a significant role. How to tackle this is if the actual website / app has clear goals that can be driven with analytics." (Interviewee 10).

"The biggest challenge with web data usage is using (all) the data that has been collected. Way too often the focus is at collecting the data instead of actually using it for something that provides business value." (Interviewee 4).

Furthermore, also this section discussing data-driven decision-making received some future development ideas that can be turned into advantages mentioning real-time data collection that was also presented within the literature review.

"Something that could create competitive advantage, is if the data used for decision-making would be real-time. Then customers would be phased faster with for instance targeting and overall, it would create better digital marketing." (Interviewee, 7).

4.4 Web analytics and skills

As mentioned in the literature review, it is important to monitor performance within the banking industry consistently to be able to gain improvement ideas and highlights the focus areas (Bakar et al., 2016). As the research seeks to gain understanding on the main challenges within the industry on web data usage and collection, it was important to ask about web analytics usage in the interview.

It appears that web analytics is broadly used within the organization, and micro conversions are built to follow the performance on some specific products and services.

"Web data initially is used to gain knowledge on the website and app usage and on different actions the users perform, that we call micro conversions." (Interviewee 7).

"We are using web analytics to analyse and measure how certain functions are operating within the websites and possibly develop a page or service further if we see anything related to that in the data." (Interviewee 1).

One specific theme that occurs in the interview answers is the basic knowledge in web analytics, especially how tracking works and what needs to be taken into account within product development by developers and product owners inside the organization. The interview answers under web analytics knowledge are divided into three emerging themes: web analytics skills, gap between teams and deficiency with the data.

4.4.1 Skills

As mentioned previously, strong skillset and good tools are the ground for successful web analytics (Kumar & Ogunmola, 2020). Therefore, the skillset should be in place for both the business to have an analytical mindset but also for the developers that add the tracking to pages. Based on the interview answers it seems that web analytical skills and understanding how tracking has been setup within the organization creates gaps and challenges for web tracking and analysis.

One of the biggest themes that repeat in the interview responses is the risk towards development teams and their knowledge on tracking. If the development teams change any page elements or structure without understanding how tracking works, it could cause the existing tracking to break, which will lead to web analysts not being able to work with the data.

"Even though not all data is received to the systems due to GDPR, it is acceptable since it is more important to have the skills to follow trends instead of total amounts." (Interviewee 8).

"Any changes in the page structure or technical changes within the application is the number one challenge. It is common that the developers will touch the source code of the page or the elements on the page in a way that the current tracking breaks." (Interviewee 5).

"The pages are continuously developed, which causes a risk for existing tracking to get broken." (Interviewee, 1).

"Especially if the organization is large and changes as well as development is done constantly for the pages, it is very possible for the existing tracking to break. It is crucial that all units operating within the websites recognize the consequences to their changes." (Interviewee, 3).

"Technically data collection itself is no longer an issue on my opinion as long the development teams have enough resources and knowledge on their own product/service." (Interviewee, 8).

"One of the biggest challenges in web data usage is making sure that the data is valid and that we can trust it 100% - sometimes we maybe use metrics that have been set up incorrectly and then we can make decisions that are based on misleading data." (Interviewee 4).

When analysing the skillset results on the data-driven decision-making process scale, it seems that data is collected but due to skill levels, it might not be trustworthy for the organizing phase. There could be cases when the page elements are changed, which would break the tracking, meaning that data would not be correctly collected, or it would break completely. Therefore, the challenge regarding skill level on the data-driven decision-making scale handles about the organize step within the process, shown in figure 8 below.

Collect → Organize → Analyze → Summarize → Synthesize → Prioritize

Figure 8: Data-driven level challenges on skills

4.4.2   Gaps between teams

Something that is highlighted in the interview results are challenges of big organizations. It seems that information is scattered and not always reaching the right people who should be involved in getting the information. This could also

be labelled as knowledge and fully understanding the big picture of what data is needed for decision-making and tracking.

"In big organizations, data can be collected in silos that can make combining or analysing it challenging or even impossible. It is also common that data collection has not been implemented transparently and through the whole organization, which is why it can be on different levels in different teams." (Interviewee 5).

"It becomes a challenge when business does not know what data development teams have implemented and business is not always included in the thinking process regarding web analytics data collection. Overall, there is a fast pace in getting things done but also the main challenge within big organizations that the information does not float." (Interviewee 7).

"Banking industry is pushing hard to provide all the services online, which can lead to gaps in data collection. Instructions can be done a bit differently towards implementation since it is impossible to know all the products and services provided within the organization. Luckily, we do have frameworks that ultimately works the same way across the organization, that provides the structure in data collection." (Interviewee 10).

"When the development teams actually prioritize tracking implementations, there will be fewer issues with the web analytics overall." (Interviewee, 3).

The challenges in the scattering information and knowledge bring issues for web analysts who do not know if the data can be reliable throughout the process and within the whole organizational level. Therefore, the challenge presented in figure 9 is on the analyse level within the process.

Collect Organize Analyze Summarize Synthesize Prioritize

Figure 9: Data-driven level challenges on information gaps

Another aspect that four different interviewees mentioned was decent documentation related to web data collection. It was mentioned that in big organizations the pace for implementation is fast and that the first thing that usually is left out in a hurry is the documentation. This is listed as a challenge since it can cause issues in the data quality in the future if information lies only behind few individuals.

"For web analysts, partial documentation challenges and slows down the reporting process, since we might not know if some data is being collected if there is no documentation of it or the document is not up to date." (Interviewee 3).

"It is important for a web analyst to know that data collection has been documented well and in a way that is easy to understand." (Interviewee 2).

"Documentation starting from the very first stakeholder has room for improvement. Good documentation strengthens and lessens incorrect interpretations." (Interviewee 5).

"The documentation is unfortunately not up to date for both data implementations, analysing and guiding the developers. This is a challenge that would need to be fixed and a better process of ownership should be in place." (Interviewee 7).

Within the data-driven map in figure 10, documentation has the biggest effect on summarizing step, since most of the interview answers for documentation were reported by web analysts having issues in knowing how to build reports and summarizing the data collected.

Figure 10: Data-driven level challenges on documentation

Furthermore, collaboration is emphasized in the interview answers as an improvement area. Partially, collaboration issues are highlighted within large organizations where information is not easily passed to all relevant parties as discussed above but it also affects the quality of processes regarding data collection and usage. This section is hard to place down in the map of becoming data-driven, created by Mandinach and colleagues (2006) with the steps of collecting, organizing, analysing, summarising, and synthesizing since it could affect all phases of the process, all the way from data collection to analysis and to decision-making.

"I would say that the primary challenge is collaboration with other units/teams. This is because typically the team who is responsible for the online data collection doesn't own the whole value chain. Thus, typically the actual implementation must be done with the development teams who own the web sites or apps. Quite often this leads to practical

challenges like agreeing the priorities across different teams; ensuring that the development team knows what they need to do and so on." (Interviewee 4).

"What I see as the biggest improvement areas for banks is improving the collaboration with other teams." (Interviewee, 10).

"The cooperation needs to work better with analysts and development teams. Analysts can create tracking themselves with different kinds of Tag Manager implementations but there is then a risk of tracking not being accurate and giving false results if any changes are done to the page later by the development team." (Interviewee 5).

4.4.3   Deficiency

When mapping out the answers regarding deficiency, it was recognisable that the answers in this category fall under deficiency regarding data, deficiency regarding privacy and deficiency in devices and systems. This topic is listed under web analytics and knowledge, since overall, it has to do with usage of web analytics.

The first interview answers in this theme focus on data deficiency. For web analysts, it seems to be hard to be able to do data-driven decisions in cases where the same device is used by multiple persons, since then the unique user is hard to detect. In addition, all the data is never collected due to GDPR and consent, the tool does not provide real time data analysis, and insufficiencies might be involved in the data collection process, which is why the data might not be completely trustworthy for usage. However, in big organizations, the amount of hits on websites are large to be able to use the data still accurately for decision-making.

"In my opinion, the number one challenge in collecting data has to do with deficiency in the data, either regarding the user using the data or with the deficiency caused by the system used." (Interviewee 1).

"The biggest challenge in data collection is being able to parse the data down to just one unique user. Especially when the same person uses many different browsers and devices. Often, also the same browser, for instance in libraries, is used by many different persons that transfer into one unique user in the data, that causes issues in detecting user paths." (Interviewee 2).

"Moreover, the biggest challenge also in data usage handles deficiency and being able to build customer paths reliably. When you know that data is not fully deficient it is hard to also be able to utilize data on a trustworthy way. Instead, you have to use a "best guess" principle." (Interviewee 2).

Partially the same subject was mentioned regarding the skill levels, but from a different aspect. When data collection implementations are not fully consistent,

it might break parts of the tracking and it cannot be used sufficiently, which also causes the data in the system to be deficient. As presented previously, this has mainly to do with the skill level on development teams if they will do any changes to the pages or elements on the pages that already have tracking in place.

"Within a large organization, data collection is unfortunately not always consistent, and some parts of the tracking do not fully work as intended. When this is the case, the usage of the data is challenging, and the comparability and trustworthiness of the data suffers. (Interviewee 3).

"I would say the number 1 challenge in using web data is the inadequacy and detachment. Rare number of organizations have web data as a solid part in a user profile." (Interviewee 1).

Furthermore, deficiencies within web analytics relate also to consents and not being able to track the user all the way to the correct entry channel of visit. Therefore, it is challenging to analyse the performance on a channel level to be able to guide the marketing budgets to the correct digital marketing platforms, for instance regarding digital marketing campaigns.

"One big challenge is analysing digital marketing performance. When a big chunk of traffic cannot be tracked due to declining the consent or any type of privacy settings within browsers, it is hard to optimize advertisements and channels through data. This will lead to web analysts being confused about which platform actually performs. (Interviewee 9).

Additionally, an interview comment mentioned new technologies within the markets that will create better user experiences but are more challenging tracking-wise. These new technologies are for instance single page applications (SPA). Moreover, it is challenging to be able to fully follow the user journey if users are changing from a device to another, especially within the banking industry where users are logging in and out of the netbanks. Therefore, the session break might create a deficiency in the analytics.

"New CMS systems and different technologies such as iframes, headless and SPA can cause challenges in data collection, especially if we want to be able to parse customer behaviour from one platform into another (cross-device and cross-domain). For instance, it is always a bit unsure whether sessions are transferred within logins." (Interviewee 5).

You could argue that the deficiency effects on the same level in the decision-making process as the skill level presented above, taking place in the collect phase. As the interview answers say, the deficiency is a challenge both in web data collection as well as analysing. However, if the collection phase would work, it would be a bit easier to tackle the challenges within the analysis part. This is presented below in the figure 11.

Figure 11: Data-driven level challenges on deficiency

5   DISCUSSION & CONCLUSIONS

This chapter discusses the study's main findings as well as draws some main conclusions on the conducted research. The purpose of this study was to understand the main challenges in data collection and usage within the Finnish banking sector. The idea for this research was to gain more insights to how web data is collected and used on a highly regulated industry and what the main issues are. Due to lack of research done for the banking industry and a high need for more research in web analytics, three research questions were created:

1. What is the purpose of using web analytics in the banking sector?
2. What are the main issues when collecting web data in the banking sector?
3. What are the main challenges when analysing web data in the banking sector?

Finally, the limitations for this study are presented and future research suggestions are proposed.

5.1   Discussion

To answer the three research questions, semi-structured interviews were conducted for 10 industry professionals working with web data within the Finnish banking sector. The received interview data were analysed with thematic analysis and presented in chapter 4.

"What is the purpose of using web analytics in banking sector?"

As stated by Nguyen and colleagues (2022), banks benefit from analysing their customers and Bakar and colleagues (2016) stated that just a 5% increase has a big impact on the banks overall profit. As anticipated, privacy is highlighted when looking into the first research question about web analytics usage in the banking sector. What can be seen from the interview answers is that data is analysed and even though the industry is highly regulated, web analytics data is collected from all users who accept cookies, which is a mandatory regulation within GDPR. However, an important recognition by Järvinen and Karjaluoto, (2015), state that before any analysis is done, KPIs should be identified. This indicates that the usage of web analytics is planned from the start and organizations are able to also lead with the data collected. It appears that the purpose of web analytics in the banking sector is mainly to identify general trends on their customers instead of directly linking it to goals and key performance indicators.

"What are the main issues when collecting web data in a banking sector?"
"What are the main challenges when analysing web data in a banking sector?"

When analysing the second and third research questions relating to web data collection and usage, there should still be development happening on the managerial level and the data deficiency. Moreover, as mentioned in the literature review, to be able to gain positive effects in web analytics usage, the site performance should be monitored and optimized with good tools and good skill level (Kumar & Ogunmola, 2020). This seem to be present within the organizations interviewed, at least based on answers related to how is web analytics used. Consistent with the previous research and as mentioned by Scheuing and Niininen (2022), organizations that handle a lot of customer data need to operate with a risk-based approach. Therefore, of course, the high level in regulations and privacy affects the possible tools used within a banking organization, but it did not seem to affect the web analytics work itself.

The results of the study indicate that there are no bigger issues in data collection, however, the skillset of teams working with editing webpage structures should be improved, since it may cause existing tracking to break. Many answers also highlighted data deficiency and the importance of being able to rely on the data collected. As listed in the literature review, the main known issues in effective data utilization have to do with lack of resources, low budgets, bad skill level, low quality data and poor analytics tools (Kemppainen et al., 2022). In addition, Zaitsava and colleagues (2022) mentioned that data needs to be trustworthy but there is always a risk when a human implements it. Therefore, improving knowledge within the analytics collection reduces the risk of data becoming insufficient.

The results of the study confirm the initial findings by Järvinen and Karjaluoto (2015), McAfee and Brynjolfsson (2012) as well as Zaitsava (2022) that organizations still have some issues in being fully data-driven, and within the managerial level the managers are not fully utilizing the data collected and turning it into business actions. Data collection and utilization seems to be an important functionality within banking organizations, but the utilization all the way into decision-making seem to be still lacking. What can be summarized from the interview answers is that data is not used for bigger decisions and planned before collection but instead a mindset of collecting everything is present, which indicates that no KPIs are set and followed later on.

Another challenge that can be drawn from the interview answers on both data collection and data usage are the issues relating directly to bigger organizations. In many cases it seems that relevant information does not always find its way to all relevant parties when the organization has many teams and thousands of employees. Besides, it appeared that the documentation relating to tracking and data usage was not always up to date for web analysts to know how the data could be used for reporting purposes. Therefore, collaboration between different teams and units within the bigger organizations could be improved.

5.2   Limitations of the study

Even though the study was carefully planned and executed, there are still some limitations within this research that provide also possible openings for future research. Firstly, the study is conducted on the Finnish market solely by interviewing 10 field professionals. A bigger sample would be required to be able to fully benefit the key findings and generalize the study results to this area. Moreover, the interview answers solely represent the subjective views from this selected sample of web analytics professionals operating within the banking industry. Secondly, as this is a qualitative study, the research findings cannot be generalized to all banking organizations in Finland. Moreover, since the research conducted focused on bigger organizations within the market, some of the research answers might not be present within smaller organizations. Purposive sampling was used, that could cause another limitation of selecting the sample solely based on the interviewees' knowledge on web analytics.

The research seeks to report the findings as openly as possible, but since the pseudonymization on the organizations and interviewees, the transparency is limited to the reader. Additionally, the analysis is done with thematic analysis with the best knowledge of the researcher but might include some limitations with the way the results are presented and understood. However, the interview questions were carefully planned and tested with the first few interviewees to ensure the reliability for this study. A few tiny modifications were made based on the first interview rounds before the final interview question setup and the interviewees were let openly speak regarding the answers. Moreover, the interviewees knew that any identifiable data of them would not be presented in the study, which ensured a good atmosphere to collect the data as openly and truthfully as possible.

5.3   Future research ideas

Since this study focuses on the Finnish market solely this already creates an opening to focus on other markets as well. Moreover, since the research was conducted with qualitative methods, more focus and in more deeper understanding could be gained with quantitative methods towards web analytics collection and usage, perhaps by broadening this to other countries and analysing more data with a survey analysis. In addition, this study focused on the bigger banking institutions on the market, which is why more research could be focused on the smaller organizations. Therefore, future studies could replicate this research in other regions and smaller organizations and compare the results with this one.

This study wanted to gain knowledge on the challenges within web analytics from the employees working closely with it. More in-depth knowledge

on the data-driven decision-making could be collected with interviewing the managerial level to gain more knowledge in why data is not fully utilized there yet. Furthermore, since machine learning and other artificial intelligence tools are constantly developing and new types of data is collected, such as voice data, this study would be interesting to redo in a couple of years, to see how the decision-making has evolved.

REFERENCES

Adams, J. Hafiz, Khan, T. A., Raeside, R. Research Methods for Business and Social Science Students. Vol Second edition. Sage Publications Pvt. Ltd; 2014. Accessed January 28, 2024. https://search-ebscohost-com.ezproxy.jyu.fi/login.aspx?direct=true&db=nlebk&AN=784882&site=ehost-live

Arsel, Z. 2017. Asking Questions with Reflexive Focus: A Tutorial on Designing and Conducting Interviews. https://academic.oup.com/jcr/article-abstract/44/4/939/4104525

Bakar, A., Clemes, D., Bicknell, K. 2016. A comprehensive hierarchical model of retail banking. International Journal of Bank Marketing. Vol. 35 No. 4, 2017, pp. 662-684

Beasley, M. 2013. Practical Web Analytics for User Experience : How Analytics Can Help You Understand Your Users, Elsevier Science & Technology, 2013. ProQuest Ebook Cen-tral, http://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=1249219.
Created from jyvaskyla-ebooks on 2024-01-30 08:34:28.

Beattie, A., Khartit, K., Jasperson, H. 2023. The Evolution of Banking Over Time – From the ancient world to the digital age. https://www.investopedia.com/articles/07/banking.asp

Belk, R., Fischer, E., & Kozinets, R. 2013. Approaches to data analysis, interpretation and theory building for scholarly research (Chapter 7) in Qualitative Consumer and Marketing Research, Thousand Oaks CA: Sage.

Boegershausen, J., Datta, H., Borah, A., Stephen, A. 2022. Fields of Gold: Scraping Web Data for Marketing Insights. Journal of Marketing 2022, Vol. 86(5) 1-20 © The Author(s) 2022. https://doi.org/10.1177/00222429221100750

Bryman, A. & Bell, E. 2007. Business research methods. 2nd ed. Oxford: Oxford University Press.

Bucklin, R. E., & Sismeiro, C. (2009). Click Here for Internet Insight: Advances in Clickstream Data Analysis in Marketing. Journal of Interactive Marketing, 23(1), 35-48. https://doi.org/10.1016/j.intmar.2008.10.004

Christanto, Y., Santoso, S. The influence of service quality, corporate image, and customer satisfaction on customer loyalty in banking sector in Yogyakarta. Research in Business & Social Science. IJRBS VOL 11 NO 7 (2022) ISSN: 2147-4478

Eriksson, P. & Kovalainen, A. 2008. Qualitative methods in business research. Los Angeles, [Calif.] ; London: SAGE

Germann, F., Lilien, G., Rangaswamy, A. 2012. Performance implications of deploying marketing analytics. Intern. J. of Research in Marketing 30 (2013) 114–128.

Hair, J. F., Wolfinbarger, M., Money, A. H., Samouel, P. & Page, M. J. 2015. The essentials of business research methods (Third edition). London: Routledge.

Jansen, B. J., Ciamacca, C. C. & Spink, A. 2008. An Analysis of Travel Information Searching on the Web. Information technology & tourism, 10(2), pp. 101-118. doi:10.3727/109830508784913121

Jyu, Jyväskylä University, 2023. Store and share. https://openscience.jyu.fi/en/tutkimusdata/store-and-share#storing-and-sharing-sensitive-data--special-categories-of-personal-information--otherwise-secret-data-

Järvinen, J., Karjaluoto, H. 2015. The use of Web analytics for digital marketing performance measurement. http://dx.doi.org/10.1016/j.indmarman.2015.04.0090019-8501/© 2015 Elsevier Inc.

Kettunen, J. 2022. Finnish Banking 2022. Overview of the Finnish banking sector and the key changes in its operating environment in 2022. Published 25 APR 2023. https://www.finanssiala.fi/en/publications/finnish-banking-2022/

Kemppainen, T., Frank, L., Makkonen, M., Kallio, A. 2022. Barriers to Data-Driven Decision-Making Among Online Retailers. In A. Pucihar, M. Kljajić Borštnar, R. Bons, A. Sheombar, G. Ongena, & D. Vidmar (Eds.), 35th Bled eConference : Digital Restructuring and Human (Re)action (pp. 327-342). University of Maribor. https://doi.org/10.18690/um.fov.4.2022.20

Kumar, V. & Ogunmola, G. 2020. Web Analytics for Knowledge Creation:
A Systematic Review of Tools, Techniques, and Practices. International Journal of Cyber Behavior, Psychology and Learning
Volume 10. Issue 1. January-March 2020

Lahtinen, N., Pulkka, M., Karjaluoto H. & Mero, J. 2023. Digital Marketing Strategy. Create strategy, put it into practice, sell more. Edward Elgar Publishing.

Mandinach, E., Honey, M., Light, D. 2006. A Theoretical Framework for Data-Driven Decision-Making. Presented at The annual meeting of the American Educational Research Association, San Francisco, CA.

Nguyen, Q, Dao, T, Tran, N. 2022. Factors Affecting the Customer Satisfaction and Bank Brand Loyalty of Depositing Savings Individual Customers in Vietnam. Journal of Positive School Psychology http://journalppw.com2022, Vol. 6, No. 5, 5524–5541© 2022 JPPW

O'Gorman, D., MacIntosh, K. & MacIntosh, R. 2015. Research Methods for Business and Management: A Guide to Writing Your Dissertation, Goodfellow Publishers, Limited, 2015. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=4531612
Created from jyvaskyla-ebooks on 2024-01-15 15:09:03.

Peukert, C., Bechtold, S., Batikas, M., & Kretschmer, T. 2022. Regulatory Spillovers and Data Governance: Evidence from the GDPR. https://doi.org/10.1287/mksc.2021.1339

Phippen, A., Sheppard, L., & Furnell, S. 2004. A Practical Evaluation of Web Analytics. Internet Research, 14, 284–293.

Provost, F. & Fawcett, T. 2013. Data Science and its relationship to big data and data-driven decision-making. BD51, vol. 1, no. 1, 2013.

Salminen, J., Jansen, B. J., An, J., Kwak, H. & Jung, S. 2018. Are personas done? Evaluating their usefulness in the age of digital analytics. Persona Studies, 4(2), pp. 47-65. doi:10.21153/psj2018vol4no2art737

Sandhu, S. & Arora, S. 2022. Customers' usage behaviour of e-banking services: Interplay of electronic banking and traditional banking. International journal of finance and economics, 27(2), pp. 2169-2181. doi:10.1002/ijfe.2266

Scheuing, S., & Niininen, O. 2022. GDPR guidelines for academic research in marketing. In O. Niininen (Ed.), Contemporary Issues in Digital Marketing (pp. 139-151). Routledge. https://doi.org/10.4324/9781003093909-17

Seddon, J., Currie, W. 2016. A model for unpacking big data analytics in high-frequency trading. Journal of Business Research 70 (2017) 300–307

Siegert, O., Varod, V., Carmi, N. & Kamocki, P. 2020. Personal data protection and academia: GDPR issues and multi-modal data-collections "in the wild". Online Journal of Applied Knowledge Management. Volume 8, Issue 1, 2020.

Sleep, S., Hulland, J. & Gooner, R.A. 2019. The Data Hierarchy: factors influencing the adoption and implementation of data-driven decision-making. Academy of Marketing Science Review, 9, 230-248. https://doi.org/10.1007/s13162-019-00146-8

Suomen Pankki. 2023. Suomessa toimivien luottolaitosten markkinaosuudet. https://www.suomenpankki.fi/fi/Tilastot/rahalaitosten-tase-lainat-ja-talletukset-ja-korot/taulukot/rati-taulukot-fi/markkinaosuudet_luottolaitokset_fi/

Tietosuoja (2022, November, 8). Processing of personal data https://tietosuoja.fi/en/processing-of-personal-data

van Ooijen, I. & Vrabec, H. 2018. Does the GDPR Enhance Consumers' Control over Personal Data? An Analysis from a Behavioural Perspective. Journal of Consumer Policy (2019) 42:91–107 https://doi.org/10.1007/s10603-018-9399-7

Yiu, C. S., Grant, K. & Edgar, D. 2007. Factors affecting the adoption of Internet Banking in Hong Kong—implications for the banking sector. International journal of information management, 27(5), pp. 336-351. doi:10.1016/j.ijinfomgt.2007.03.002

Yle, 2022. Sadat S-pankin asiakkaat pääsivät kirjautumaan toisten verkkopankkeihin – Finanssivalvonta: poikkeuksellinen ja vakava tapahtuma. https://yle.fi/a/3-12621116

Zaitsava, M., Marku, E., Guardo, M. 2022. Is data -driven decision-making driven only by data? When cognition meets data. European Management Journal 40 (2022) 656–670. https://doi.org/10.1016/j.emj.2022.01.003

Önder, I., Berbekova, A. 2021. Web analytics: more than website performance evaluation? VOL. 8 NO. 3 2022, pp. 603-615, Emerald Publishing Limited, ISSN 2056-5607. DOI 10.1108/IJTC-03-2021-0039

APPENDIX 1   The interview protocol

0. Explanation on how the collected data will be used and that it will be pseudonymised.

1. What is your age group: 20-30, 30-40, 40-50, 50-60?
2. What do you work with within the case organization?
3. How long have you been working with web data?

4. What is the tool used for data collection?
5. Why is that specifically used?
6. How is data collected within your organization?
7. How is privacy and regulations affecting it?
8. How is data used within the banking industry?

9. What would you say are the biggest issues with web data collection?
10. What of these is the main issue and focus point?
11. How is web data used within the organization?
12. What are the main pitfalls with utilizing web data?
13. How do you see the organization and management level utilizing the data?

14. What would you say are the challenges in data usage and collection within the banking industry?
15. What could be done better?

16. Anything you would like to add regarding the interview questions?