A case study of digital innovation in the online advertising industry in Norway

Jostein Bleken Hellerud and Karl Thomas Hauglid

Thesis submitted for the degree of Master in Software Engineering
60 credits

Department of Informatics
Faculty of Mathematics and Natural Sciences

UNIVERSITETET I OSLO
May 2018
A case study of digital innovation in the online advertising industry in Norway

Master’s thesis

Jostein Bleken Hellerud and Karl Thomas Hauglid

May 2018
A case study of digital innovation in the online advertising industry in Norway

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

IV
Abstract

The online advertising market in Norway is dominated by big players such as Facebook and Google.

In this thesis, we explore this situation and the challenges of the Norwegian advertising industry in the online advertising market and suggest how to strengthen the position of the Norwegian actors. This is done by (1) exploring how online advertising works and the challenges faced by the Norwegian industry. With the use of a conceptual framework of digital innovation, software-based platforms, and value networks, this thesis (2) looks at how the challenges can be met and (3) generalize our findings to propose a set of software design guidelines. The scope of this thesis is the online advertising industry in Norway.

A qualitative case study on the online advertising industry was conducted through studying a project led by a Norwegian startup established within the field of programmatic advertising. The data were collected through document analysis, interviews, observation of a workshop, and continuous discussions with the project lead from the startup.

The results show that a prominent challenge faced by the Norwegian online advertising industry is the dominant role of big international players. This has two primary implications; (1) that small Norwegian companies struggle to compete and (2) big actors become the preferred choice because they possess a superior amount of data.

Through a theoretical framework of value creation, this thesis explores how a software solution could increase the competitiveness for local software providers in the Norwegian advertising market. Furthermore, the study shows that success for a potential solution is tightly coupled to its ability to bootstrap itself.

This thesis contributes by, first creating a clear description and overview of the online advertising industry in Norway and how this relates to online information privacy. This target those looking for insight into the practices of the online advertising. Secondly, we describe how the challenges in the Norwegian advertising industry can be met through the lens of digital innovation, aimed at IT professionals in the industry. Lastly, we propose a set of design guidelines to maximize the attraction for a software service, aimed at designers of software services in online advertising looking to increase the attraction of a software service.
Acknowledgment

First of all, we would like to thank our supervisor Petter Nielsen for guiding us through this thesis and providing helpful feedback. In addition, we are also thankful to Kjetil Hustveit for allowing us to follow the project and providing us with good help and discussions. Additionally, we would like to mention the participants that have taken their time to contribute to this thesis.

We would also like to thank our fellow students and the employees of the research group Information Systems for valuable discussions, distractions, and coffee breaks. Especially, we thank Egil "Vandraren" Øvrelid for helpful discussions and inputs. Finally, we would like to thank our family and friends for motivating us.

Jostein Bleken Hellerud and Karl Thomas Hauglid
Oslo, May 2018
# Table of Contents

Abstract .......................................................................................................................... V  
Acknowledgment .......................................................................................................... VII 
Table of Contents .......................................................................................................... IX 
List of Figures ................................................................................................................ XIII 
List of Tables .................................................................................................................. XV 
1 Introduction .................................................................................................................. 1  
  1.1 Motivation .............................................................................................................. 1  
  1.2 Research Context ................................................................................................. 2  
  1.3 Research question ................................................................................................. 2  
  1.4 Thesis structure .................................................................................................... 3  
2 Background .................................................................................................................. 5  
  2.1 Online advertising ................................................................................................. 5  
    2.1.1 A brief history ................................................................................................. 7  
    2.1.2 The evolution of programmatic advertising ..................................................... 9  
    2.1.3 Online advertising compared with traditional advertising ................................ 11  
    2.1.4 Actors .............................................................................................................. 14  
    2.1.5 Programmatic advertising software ................................................................. 16  
    2.1.6 Tracking and identifying methods ................................................................... 19  
    2.1.7 Cookie matching ............................................................................................. 23  
    2.1.8 Summary ......................................................................................................... 25  
  2.2 Online information privacy ..................................................................................... 26  
    2.2.1 Defining user data .......................................................................................... 26  
    2.2.2 Economies of user data ................................................................................ 27  
    2.2.3 General Data Protection Regulation .............................................................. 28  
    2.2.4 Summary ......................................................................................................... 30  
3 Theory ........................................................................................................................ 33  
  3.1 Value network ........................................................................................................ 33  
  3.2 Software-based platforms ..................................................................................... 36  
    3.2.1 Two-sided markets ......................................................................................... 37  
    3.2.2 Summary ......................................................................................................... 38  
  3.3 Digital innovation ................................................................................................. 38  
  3.4 Internet Users’ Information Privacy Concerns ....................................................... 40  
  3.5 Summary ................................................................................................................ 43  
4 Research method ........................................................................................................ 45  
  4.1 Identifying the research question .......................................................................... 45
6.2 How to acquire a competitive amount of data .................................................. 92
6.3 Creating attraction for Digisphere Core ............................................................. 93
   6.3.1 Leveraging network effects to increase attraction ........................................ 94
   6.3.2 Digisphere as a platform ............................................................................. 96
   6.3.3 Digital innovation ...................................................................................... 97
   6.3.4 Summary .................................................................................................... 99
6.4 Factors for success ............................................................................................. 101
   6.4.1 Design initially for direct usefulness .......................................................... 102
   6.4.2 Bootstrapping the platform ......................................................................... 105
   6.4.3 Potential challenges and solutions .............................................................. 106
6.5 Design guidelines to increase attraction for a software service ......................... 107
6.6 Limitations ........................................................................................................ 111
6.7 Summary ........................................................................................................... 114
7 Conclusion and future research ............................................................................ 117
   7.1 Conclusion ...................................................................................................... 117
   7.2 Contribution .................................................................................................... 118
   7.3 Future research ............................................................................................... 118
Bibliography ............................................................................................................ 121
Appendix .................................................................................................................. 125
   Attachment A : Interview Guide Part A ............................................................. 125
   Attachment B : Interview Guide Part B ............................................................. 126
   Attachment C : Example of data reduction ....................................................... 127
   Attachment D : Examples of data display ......................................................... 130
   Attachment E : Examples of conclusion drawing ............................................ 131
# List of Figures

Figure 1: A simple overview of the advertising market ................................................. 14
Figure 2: Software components of programmatic advertising ........................................ 17
Figure 3: A model of value networks (Kothandaraman & Wilson, 2001) ............................ 34
Figure 4: Method location (Braa & Vidgen, 1999) .......................................................... 48
Figure 5: Participant’s position in the advertising industry ............................................. 51
Figure 6: Example mock-up of Digisphere Gate login on ToyCars.com .......................... 74
Figure 7: Mock-up confirmation page with Digisphere Bazaar ....................................... 78
Figure 8: Illustration of the public API of Digisphere Core ............................................. 82
Figure 9: Illustration of value stacking in Digisphere ...................................................... 84
Figure 10: The current architecture of Digisphere ............................................................ 87
Figure 11: Visualizing the combination of data to be competitive ................................... 93
Figure 12: Network effects in Digisphere ....................................................................... 100
Figure 13: Summary of the Digisphere platform and network ......................................... 101
Figure 14: Illustration of network effects ....................................................................... 109
Figure 15: Illustration of the relations between a platform and a network ....................... 110
Figure 16: Illustration of the relations between a network and digital innovation ............ 110
List of Tables

Table 1: Comparing the four categories of online advertising ........................................ 7
Table 2: Comparing the evolution of programmatic advertising ..................................... 11
Table 3: Comparing online methods for identifying users .................................................. 23
Table 4: Example of a Cookie Matching Table ............................................................... 24
Table 5: Summary of the theoretical framework used in this thesis ................................. 43
Table 6: Our research design based on Robson (2002, p. 81) ....................................... 47
Table 7: Overview of interviews done in part A and B ...................................................... 51
Table 8: Overview of the discussions .............................................................................. 55
Table 9: An overview of the collected documentation ...................................................... 56
Table 10: Our approach to threats to validity based on Robson (2002, p. 174) .................. 59
Table 11: Digisphere Gate example scenario ................................................................. 70
Table 12: Example of stored information in Digisphere Gate ......................................... 70
Table 13: Example of information stored as pairs in Digisphere Gate ............................. 71
Table 14: Example visualization of lookup in Digisphere Gate ........................................ 71
Table 15: Comparing cookie matching and Digisphere Gate ......................................... 75
Table 16: Example of Digisphere Mind for email ............................................................ 75
Table 17: Digisphere Bazaar example ............................................................................ 78
Table 18: Examples of cases for Digisphere Bazaar ......................................................... 80
Table 19: Comparing Digisphere Bazaar and Programmatic Advertising ....................... 81
Table 20: Summary of the purposes of the components of Digisphere ............................. 85
Table 21: The three core building blocks of the Digisphere value network ....................... 95
Table 22: Digisphere Core’s relation to the core components of platforms ...................... 97
Table 23: Summary of theoretical contributions to attraction ........................................ 100
Table 24: Digisphere in relation to the IUIPC framework ............................................... 103
Table 25: Design guidelines to increase attraction for a software service ....................... 108
1 Introduction

1.1 Motivation

Online advertising is a term used for all advertising delivered over the internet (Evans, 2009). The main actors in online advertising are buyers (advertisers) and sellers of advertising space (publishers), and the actors providing advertising software services. These services are highly dependent on user data, which are data collected about individual internet users.

User data has affected the evolution of advertising (Busch, 2016). Advertising can be roughly split into two different models. In the first model, ads are targeted at groups of similar people who get the same ad. One example is the ads shown on the television during commercial breaks targeting teenagers. In the second model, individuals, not groups, are sold online by publishers to advertisers in real-time without any manual input. This model is called programmatic advertising (Busch, 2016). Programmatic advertising means that two different individuals, in the same age group and with many of the same interests, can get different ads when connecting to the same webpage.

Collecting data about an individual user increases the potential value of showing this individual an advertisement through programmatic advertising (Datatilsynet, 2015). Advertisers are concerned about the precision of their advertising, to avoid showing an advertisement to individuals not likely to buy their product. Increasing the amount of information the advertisers have about an individual will increase the chance of showing an advertisement to their target audience. This is because if one actor knows that you are looking for a new bicycle they would want to pay to be able to show you an ad for a bicycle sale. As a result, almost every actor in the online advertising industry collects and utilize as much data as possible to improve the value of their advertising (Zuboff, 2015).

The introduction of the General Data Protection Regulation (GDPR) gives the user a deciding role in who can collect data about them. Mitigating users’ concerns related to data collection can be argued to be a success factor for actors in the online advertising industry.

The position that data has in this industry means that having less data than other actors is a weakness. This favors the big players because they can gather and analyze more data by having more contact points and a much larger user base than the smaller local actors. The
result of this is an industry dominated by companies such as Facebook and Google. Several hundred other companies still have their place in the industry, but struggle to compete. This includes companies of all sizes, holding one or more roles in the online advertising value chain.

A case study will be performed on a project driven forward by a startup company trying to increase their competitiveness. The project is an attempt to innovate with the purpose of establishing a software service called Digisphere within the domain of programmatic online advertising in Norway. We call this project for DITIC, which stands for Digital Innovation To Increase Competitiveness. We will examine DITIC and use the Internet Users’ Information Privacy Concerns (IUIPC) framework to assess how DITIC mitigates users’ concerns. Theories of value network, software-based platforms, and digital innovation will be used to explore how DITIC can help smaller actors survive the market pressure provided by Facebook and Google.

### 1.2 Research Context

This thesis is placed in the information systems (IS) research area. The empirical study was conducted over a period of one year, from mars 2017 to mars 2018, and was done in the online advertising industry. The participant in the study were all located in Oslo, Norway. To better understand the case and its context, existing literature about online advertising will be examined and applied.

### 1.3 Research question

The purpose of this thesis is to investigate how we can strengthen the position of the small local actors in the Norwegian advertising market. The research question that will be addressed throughout this thesis is the following:

- How can digital innovation enable increased competitiveness for local software providers in the Norwegian online advertising market?
1.4 Thesis structure

Following is a brief summary of the content of each chapter in this thesis.

**Chapter 2: Background** contains background information regarding online advertising and online information privacy.

**Chapter 3: Theory** contains our theoretical framework needed to answer the research question. The framework is based on network, platform, and digital innovation theories.

**Chapter 4: Research method** describes the applied research method and strategy for data collection.

**Chapter 5: Case description** consists of an introduction to the case under study and detailed descriptions of its individual parts, actors, and evolution. The chapter is based on insight gained from data collection.

**Chapter 6: Discussion** contains a discussion based on the chapters in background, theory and case description in order to answer the thesis’ research question. A set of design guidelines will be presented based on the generalization of the findings. Furthermore, the limitations of the study will be outlined.

**Chapter 7: Conclusion and future research** presents the conclusion to the research question, maps out our contribution, and propose what topics for future research.
2 Background

An overview of online advertising will be given as a basis for understanding the research context. This overview consists of two parts. The first part is about online advertising; this consists of an overview of the history and evolution of programmatic advertising. A description of the four main categories in programmatic advertising will be given, and the methods for identifying users and collecting data will be described. The second part contains a definition of personal data, the economies of user data, and a brief description of the General Data Protection Regulation (GDPR).

2.1 Online advertising

Online advertising is a term used for all advertising delivered over the internet, including advertising delivered to mobile phones (Evans, 2009). Like traditional advertising, publishers (content creators) attract viewers and offer advertisers access to those viewers. This is done by the publishers selling access to their space, where advertisers can display content designed to attract customers (ads). The first online ad was sold to AT&T by the web magazine HotWire in 1994 (Evans, 2009) and a lot has happened since then. Online advertising has transformed from being an online equivalent to newspapers to becoming a well-oiled machine, selling more than 1.3 million user impressions every second (Datatilsynet, 2015). 2016 was the first year in history where media spending’s on online advertising surpassed traditional television (Hamilton, 2017).

The online advertising industry is often divided into four main categories (Evans, 2008):

- **Display Advertising.** When advertising is displayed on webpages it is usually referred to as Display Advertising. This type of advertising consists of graphical content in the form of pictures, text and/or video, and may contain audio. The advertisement is presented to users looking to consume content provided on a webpage. Display advertising can be shown to the user alongside the desired content, sharing the space available to the webpage, or block the desired content, thereby forcing the user to view the advertisement before access to the content is provided. Examples of webpages using display advertising are many and can be found in all categories. Examples of
categories where they exist are online newspapers, blogs, social media or media platforms.

- **Search Advertising.** Advertising displayed in the search result of online search engines is categorized as Search Advertising. It refers to advertising displayed together with the search result of online search engines. The advertisement is presented in a similar fashion as the search results, as pure text with a headline, description and a link to the relevant webpage. Search advertising is usually presented above regular search results. These ads are usually visually differentiated in some way from regular search results to inform the user that they are looking at advertising. Search advertising has one key attribute not found in other types of advertising: The user uses the search engine to find information about something, making the user more perceptible to relevant advertising providing information about what the user looks for. Search engines can use display advertising on their sites as well, but this will not be presented as a part of the search results, but rather share the available space and be presented alongside the search-section of the search engine. Examples of online search engines are google.com, bing.com, and yahoo.com.

- **Classified Advertising.** Classified Advertising is usually associated with traditional printed media and is also known as Classified Listings or Classifieds. This form of advertising is also used online. The advertisement often consists of pure text together with images, displayed on dedicated sections of webpages without other content than Classified Advertising. These sections can be found on several types of webpages, including online newspapers. There are webpages specializing in Classified Advertising and offers this as their value. Users viewing Classified Advertising are often visiting a website with the purpose of finding these advertisements.

- **Email Advertising.** Advertisement distributed through email is often referred to as Email Advertising. The email can consist of display advertisement and include images, text, and video. The content can be tailored to the recipient, but this is not always the case. Email advertising can happen in several ways; (1) by distribution to potential customers at regular time intervals; (2) distribution could be triggered by user interaction with the advertiser; (3) as well as relevant activities by the advertiser, such as sales or promotion. Advertisers choose recipients by collecting email
addresses of users actively agreeing to receive the advertisement and/or through other means of acquiring email addresses without consent (e.g. buying lists of email addresses). When Email Advertising becomes too intrusive in the form of the time interval of the emails, their relevance, and the recipient’s willingness to receive, it is often referred to as Spam.

We have summarized these categories on location on webpage and examples of typical webpages in the table below.

<table>
<thead>
<tr>
<th>Type</th>
<th>Location on webpage</th>
<th>Examples of typical webpages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display Advertising</td>
<td>Alongside desired content</td>
<td>Newspapers (Vg, Aftenposten), blogs, social media (Facebook, Instagram)</td>
</tr>
<tr>
<td>Search Advertising</td>
<td>Search results</td>
<td>Search engines only (Google, Bing, Yahoo)</td>
</tr>
<tr>
<td>Classified Advertising</td>
<td>The main content of dedicated webpage or subsection</td>
<td>Newspapers (Vg), landing pages (Startsiden) or dedicated webpage (Rubrikk)</td>
</tr>
<tr>
<td>Email Advertising</td>
<td>Email</td>
<td>Not displayed on webpages</td>
</tr>
</tbody>
</table>

Table 1: Comparing the four categories of online advertising

2.1.1 A brief history

The early days of display advertising were largely concerned with selling banner ads. These were advertisement placed alongside the desired content on a webpage, like an advertisement in newspapers and publishing media in general. The adverts were equal for all users and had a static placement on the webpage. The economic model was based on the number of views the ad would get, making it similar to the model followed by the periods traditional publishing media. As a result, many ads at the time were sold as “cost per mille” (CPM) which means cost per 1000 viewers.

The CPM approach to charging for advertisement changed in 1996 when Yahoo! agreed to a deal based on “cost per click” (CPC), meaning that the publisher would only get paid when a user clicked on the ad. This approach was similar to the one used by traditional direct advertisers at the time, such as telephone solicitations (i.e. telemarketing). CPC dominated the industry for a while and with its adaptation by the online search engines (mainly GoTo.com)
innovated the business model further. This led to the introduction of auction-based advertising, enabling advertisers to bid in real-time for advertising spots on the page showing search results. Up until this point, online advertising (and services) had mimicked offline solutions, but the auction based advertising innovation marked a break from the traditional advertising industry (Evans, 2009).

Auction-based advertising developed into what is known as “Programmatic advertising” or “real-time bidding” (RTB), which is a key development in the evolution of online advertising. Oliver Busch describes it in his book “Programmatic advertising” as:

“Programmatic advertising describes the automated serving of digital ads in real time based on individual ad impression opportunities” (Busch, 2016)

The evolvement leading to programmatic advertising relied on identifying the individual user, with high precision, through advances in identifying and web tracking technology. This advancement was then used to enable advertisers to display advertisement based on the individual user visiting the webpage. This allowed re-targeting, one-to-one communication and persistent advertising towards the same user based on harvested data about the individual. The harvested data contains information collected about the user. Examples of the collected data could be information about a webpage the user has visited or internet search history. A more detailed description of personal data is given in chapter 2.2.1 Defining user data.

Programmatic advertising is the combination of identifying individual opportunities to display an advertisement to individual users (ad impressions) and auction-based advertising. Programmatic advertising is a broad term for auction-based buying and selling of advertising displayed for one specific user. The term includes the act of using data about the users to make marketing decisions in real-time, which is done to enable intelligent buying and selling of ad impressions to individual users. Programmatic advertising increased the efficiency of marketing compared to earlier approaches to online advertising (Busch, 2016).

The term ad impression is used to describe the moment a user views an advertisement. After displaying an ad to an individual user, this user has then been served an ad impression. It is used a lot in programmatic advertising because ad impressions are the products that are made available for sale, through the programmatic technology. It is an important term for the
history of online advertising, and it encompasses some of the contrast to the earlier days of online advertising where advertising space was for sale, not the individual impression.

Social media advertising is also an important part of the evolution of online advertising. With the introduction of social media in the early 2000s, advertisers could buy user impressions on a social platform in the same way as they could with other content webpages, allowing them to display ads next to a user’s social interactions. Social media platforms became a valuable resource for harvesting user data, as the users engaged with the content created opportunities for making assumptions about their interests (Russell, 2013).

Social media created an opportunity for free advertising for advertisers by engaging with the users of social media platforms through joining them on the platform. By creating a brand profile and actively engaging with users, advertisers could participate in a new way by creating content users enjoyed and spread by sharing with their friends and followers. This required different advertising than earlier and word of mouth promotion on the social platform was often the goal (Dehghani & Tumer, 2015). Social media platforms would also provide the opportunity for advertisers to pay to increase the visibility of their posts, and in that way, reach more users.

2.1.2 The evolution of programmatic advertising

Programmatic advertising has been through a transformation since its introduction. As described above, programmatic advertising has its roots in auction-based advertising and re-targeting individuals through identifying and tracking users. The first version of programmatic advertising (programmatic advertising 1.0) was merely concerned with re-targeting and had the advertising goal of reaching users that had shown interest in a product but not yet bought it. Data about the user were collected by the advertiser (first-party), and the data collected were concentrated on the user’s purchase history and the path to purchase. Advertisers were at this stage mostly focused on reaching the same user multiple times and the quality of the webpage where their ads were displayed were mostly irrelevant. This resulted in an open real-time bidding platform for individual ad impressions being the preferred choice for most advertisers (Busch, 2016).

The next focus of programmatic advertising (programmatic advertising 2.0) became the reach and targeting of potential new customers. This meant trying to create brand awareness with
users that had not engaged with the advertiser yet. To enable advertisers to reach users that were more likely to buy their products, programmatic products emerged that collected a massive amount of data about consumers. The data was then used to deduce possible interests of the user and in that way, build a targeting profile for individual users. This created a market for advertisers to buy data from third-party companies that based their income on collecting data and analyzing it (known as data brokers and/or market analyzers). The advertisers became more focused on reaching the users that had a targeting profile matching the one most likely to buy their products, instead of re-targeting individual users. As a result, advertisers negotiated deals with webpages that had a user base which matched the advertiser’s target audience, instead of focusing on re-targeting (Busch, 2016).

The third focus of programmatic advertising (programmatic advertising 3.0) became a Consumer-Centric view on advertising. This is a very data-driven way of doing online advertising and involves harvesting and combining as much data as possible about the user. Programmatic advertising 3.0 learned from the two previous versions and combines first-party data about purchase history with third-party data collected by data brokers. This data is used to create a much more selective buying process where buying techniques from both programmatic advertising 1.0 and 2.0 are used. This is often made possible by Private Marketplace Deals, where advertisers gain access to a bidding platform with increased quality of ad-placement and publishers. As a result, these private marketplaces can offer a more premium ad-format for advertisers (Busch, 2016).

The Consumer-Centric approach to programmatic advertising (and online advertising in general) is an essential and important change to online advertising. It differs from earlier approaches by starting the decision making related to bidding on ad impressions with the user. Traditionally, these decisions started with allocating budgets, choosing placement and then the users. With a Consumer-Centric approach, however, these decisions are different and are all based on the consumer. When a user is made available for sale on the bidding platform (or ad exchange) several decisions need to be made before a bid is placed.

- **Right context?** What is the content of the webpage the user is visiting? Is it appropriate content with regards to the potential ad?

- **Right quality?** What is the quality of the ad impression for sale? Where on the webpage will the ad be displayed?
• **Right time?** Is it the right time of the day and day of the week? How many ads from the advertiser has this user been exposed to lately? When was the user last in contact with the advertiser?

• **Right location?** Is the user using his/her mobile, pc or tablet? How will the device in use affect the advertisement?

• **Right price?** At what price can the advertiser buy the impression?

All these questions need to be answered and evaluated before a bid is placed. This approach makes the user the deciding factor and involves a key differentiation compared to earlier approaches to programmatic advertising (Busch, 2016).

Even though the later versions of programmatic advertising are evolutions of the first version, all versions are still used and serve different purposes for advertisers. Approaches to online advertising often include all versions of programmatic advertising, as well as a social media presence. The table below compares the different version of programmatic advertising.

<table>
<thead>
<tr>
<th>Version</th>
<th>Focus</th>
<th>Data source</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programmatic 1.0</td>
<td>Re-targeting</td>
<td>The advertiser (First-party).</td>
<td>Low</td>
</tr>
<tr>
<td>Programmatic 2.0</td>
<td>Brand awareness</td>
<td>Data brokers (Third-party)</td>
<td>Medium</td>
</tr>
<tr>
<td>Programmatic 3.0</td>
<td>Consumer Centric</td>
<td>The advertiser and data brokers (First- and third-party)</td>
<td>High</td>
</tr>
</tbody>
</table>

Table 2: Comparing the evolution of programmatic advertising

### 2.1.3 Online advertising compared with traditional advertising

Online advertising draws many advantages from the nature of the world wide web. In contrast to traditional advertising, such as newspapers, radios, and television, online advertising offers, as mentioned above, the ability to learn about the individual that is consuming the content. This can be done quickly with low precision by looking up the IP (Internet protocol) address of the user (Evans, 2009), or more complex with increased precision through the use of cookies (Datatilsynet, 2015). A basic principle in online advertising says that “Real value comes with real identity”. It has, therefore, become vital to be able to track the user with
precision across various devices (Busch, 2016). By identifying the same user in various places as he/she navigates through the internet, advertising networks have the ability to create a profile on the specific user, containing interests and other forms of data that could be of value for advertisers. This information is then used to target the user with ads directly aimed and personalized for the individual user, in contrast to traditional advertising’s ability to only target groups of individuals (Datatilsynet, 2015).

Another thing that differentiates online advertising and traditional advertising is the use of technology to create a highly efficient marketplace for selling and buying advertising space. By the development of real-time bidding into programmatic advertising, the process has been gradually more automated to the point of a fully automated market with real-time bidding, information, and creation. This makes it possible for advertisers to create consumer-centric advertising and creates opportunities for the advertisers to optimize their budgets and select, evaluate, price, and create an advertisement with high precision, aimed at individuals. Prices are based on real-time supply and demand for each individual ad impression opportunity at the moment of its creation. Advertisers can, therefore, specify what one ad impression towards that individual is worth for them. The content of the advertisement is automatically created for the individual based on their profile (often previous purchases or interests). All this automation means that the online advertising market is made efficient to a point far beyond traditional advertising could ever be (Busch, 2016).

These above-mentioned attributes imply that online advertising has requirements not seen in traditional advertising. The most eminent of which are (Busch, 2016):

- A high bandwidth and low latency data connection:

  The automatization of these processes creates a demand for transmitting data fast. Before an ad impression is bought, created and delivered to the user, its content and data could have traveled several times around the world.

- Computing power:

  To make continuous real-time marketing decisions data needs to be computed in milliseconds for several independent individuals every second. This requires immense computing power in order to not miss out on a potential impression.
• Fast data storage:

Storing every possible data point of interest for a vast number of individuals requires enormous amounts of storage. It should also be possible to access the data fast to make decisions in real-time.

These technological factors contribute to the fact that online advertising is a complex endeavor. For this reason, among others, online advertising is dominated by specialized platforms that handle the complexity. To avoid having to invest in expensive hardware and software most online publishers hire platforms to handle their advertising space (Evans, 2008).

Online advertising also differs from traditional advertising in non-technical aspects. Both marketing teams and agencies have their work changed dramatically with the introduction of online advertising. The professionals in the industry must know a lot more to create successful advertisements online than what is required for traditional advertising (Hamilton, 2017). Understanding and mastering the programmatic advertising landscape requires knowledge about how the software functions for demand and supply side, as well as the ad exchange.

Social media has also changed the requirements to the human side of the advertising industry. Traditional advertising, and even programmatic advertising, has been focused on creating marketing campaigns, with many time-consuming steps and discussions to create the best result. Social media requires instant replies and reactions to current events. One example of this is when brands responded to the blackout during the 2013 Super Bowl to post relevant ads on social media, such as an image with the line “you can still dunk in the dark” from the cookie manufacturer Oreo (Hamilton, 2017). This means that advertising teams and agencies need to work differently to be able to respond and maintain a brand’s image outside of planned campaigns.

To summarize, four main technology factors has emerged, distinguishing online advertising from traditional advertising: (1) The internet provides effective means for collecting information about individual users and targeting these with personalized ads; (2) Online advertising delivers a more efficient market for buying and selling advertising space and providing advertising content, programmatic advertising is an example of this; (3) Online
publishers hire specialized advertising platforms, such as Google, to handle the selling of their advertising space (Evans, 2008); (4) Online advertising requires advertising agencies and teams to work differently and possess knowledge about software and programmatic advertising (Hamilton, 2017).

2.1.4 Actors

Online advertising does have many similarities with traditional advertising, and the most eminent of which is the actors involved. The most important actors in advertising are the sellers and buyers of advertising space (or ad impressions). These actors are also the ones visible to the consumer.

The online advertising industry does not consist exclusively of buyers and sellers, and there are several other roles to fill in the value chain of online advertising. Datatilsynet (2015) points out that differentiating between the actors in this industry is a challenge because many actors play several roles in the value chain. We have modified the categories found in Datatilsynet (2015) and divided the value chain into four main categories. A description of the categories will follow, with a visual overview of the industry in Figure 1. The colors in Figure 1 represent the various categories and the arrows represent the relations between the actors.

Figure 1: A simple overview of the advertising market

**Vendors of advertising space (Supply Side)**
The companies selling advertising space are usually called Publishers. The term is a generalization for companies making their living from selling advertising space. This includes online newspapers, news portals, social media, landing pages and search engines.

Traditionally, media companies had the larger part of the revenue share in this sector, but in the latest years, this is no longer the case. Advertisers have moved to other publishers such as
Facebook and Google, and media companies are now struggling. This is largely because these companies are able to offer greater precision in targeting individuals, through utilizing the data they possess about their users. However, there is still a market for ads through traditional publishers. One advantage of this is that publishers can target content to increase the value of their advertising spots. In this way, a publisher can offer advertisers for bikes to place their ads beside content that attract potential bicycle buyers (Datatilsynet, 2015).

There are also companies in this space that do not make their living from vending advertising space but still wishes to sell potential ad impressions. Examples of this are newspapers with subscription options, mobile games or other online services that have a lot of user traffic but does not rely on an advertisement for income. These are still vendors of advertising space but may not fit the traditional model of publishers.

**Buyers of advertising space (Demand Side)**
The buyers of advertising space are known as *advertisers*. The advertisers rely on the publisher to show their ads to the desired potential buyers/users of their product. Every company or organization wishing to market their product to consumers could be a potential advertiser. To be able to place advertising without extensive knowledge about online advertisement it is common for advertisers to hire *media agencies*. The media agencies assist advertisers in placing advertisement wisely but are not strictly necessary for online advertising. There exists independent software that provides solutions for advertisers to handle this on their own.

The advertisers usually collect data about their customers through loyalty programs or previous purchases. It is important for advertisers to be able to use this data in a way that improves their advertising and media agencies. Solutions that possess data who could aid in this process are often preferred.

**Vendors of programmatic advertising software**
The automatization of online advertising (programmatic advertising), requires software to function. This opens the value chain to software providers, producing platforms and solutions to enable programmatic advertising. Both advertisers and publishers need software platforms to enable real-time buying and selling. In addition to this, there is a need for a marketplace to link the two sides, which is called the ad exchange. The companies producing this software
have several places in the value chain and there are several vendors in this space (Datatilsynet, 2015).

The individual parts of programmatic advertising software are described in Chapter 2.1.5.

Data analysis companies
This group of companies sells data and market analysis to both supply and demand side. They are a part of programmatic advertising and remain invisible to the consumer. The companies collect, store and analyze data about the individual users for the purpose of selling this information to others. The companies collect data through publicly available sources (e.g. social media), commercial partners and the government. The data is then analyzed to find interests and habits and then places the individual in one or more categories. This is a large part of the revenue in online advertising and is important to the ecosystem because it provides data that makes targeting of ads more precise and effectful.

Companies in this category may also provide market insight. This could be measuring the effect of a campaign in real-time to adjust the campaign if it does not produce optimal results. Another contribution could be providing insight to help companies target the right demographic.

Whatever the insight provided is, this category depends on tracking and collecting data about user behavior in real-time. Not all the companies in this category do everything described in this section but may specialize in one or more parts of the chain.

2.1.5 Programmatic advertising software
As mentioned, programmatic advertising refers to the system for automating buying and selling online ads in real-time and is where the different actors meet. There are, however, several details in how this is achieved not yet described. The system consists of three distinct software parts. The relation between the parts is visualized as arrows in Figure 2.
Supply side platform (SSP)
The SSP is the publisher’s software platform for offering ad spots (potential ad impression opportunities) to advertisers. It is used to manage the selling of space on the website where the publisher wants to display ads. The SSP notifies potential advertisers of a potential ad impression through the ad exchange when a user loads their webpage. When the advertiser willing to pay the most for this opportunity is selected, their ad is loaded and displayed to the user navigating the publisher’s webpage.

There are several companies offering such software, some of the most well-known are “Google Doubleclick”, “Rubicon Project” and “Pumatic” (Datatilsynet, 2015).

Demand side platform (DSP)
The DSP is the advertiser’s software platform. It is software that serves ads on behalf of the advertisers. The advertisers specify rules that determine where, how and to what price their ads will be displayed. Its functionality is as an automatic bid manager and purchases ad impressions based on targeting criteria. The targeting criteria is often an algorithm based on several sources with the goal of predicting the ad impressions value for the advertiser. Ad impressions are valued based on the likelihood of it leading to a purchase with the advertiser. There are several factors affecting this, but the most common algorithms combine data from the advertiser (f. ex. previous purchase history), behavioral data collected from the DSP (f. ex. known interests) and data from third parties (f. ex. Social media) to make a bid.

Together with the opportunity to place the ad on the publisher’s webpage the DSP winning the bid will place a cookie in the user’s browser. This is used to track how efficient their algorithms are and measure its effect based on the user’s response to the ad. This is also used to identify and target the same user again on different webpages.
Companies offering DSP software are several, some well-known examples are “Google DoubleClick Bid Manager”, “Flurry” and “Xaxis” (Datatilsynet, 2015).

**Ad exchange**
The ad exchange acts as an intermediator between the SSP and DSP. When the SSP has a potential ad impression the ad exchange is notified and relays information about the potential ad impression to the registered DSPs. The ad exchange is like a stock exchange and is a marketplace for selling and buying advertising space. It acts as the natural middle ground between publishers and advertisers.

The ad exchange “conducts” an auction to fill the potential ad space on a webpage with ads. In order to do this, the ad exchange has to receive what is known as an “ad call”. The ad call starts the process and is sent to the ad exchange from the user’s browser when the user opens a webpage that is a part of an SSP. After the ad call is received by the ad exchange, it has access to read the cookies installed on the user’s browser. There will be a cookie identifying the user if the ad exchange has previously shown ads to the user. This enables the ad exchange to accumulate information about the user. This information usually contains which ads the user has seen previously, technical information about the computer (OS, browser and other software), IP, address, and location. If this is a user, the ad exchange has not seen previously, a cookie is placed in the user’s browser and this information is gathered for the first time. This information is then sent to all the DSPs linked to the ad exchange. This is done by another “ad call” and its purpose is to notify advertisers about the potential ad opportunity. As with the previous “ad call”, this enables the DSPs to read the cookies they have installed in the user’s browser.

After each individual advertiser has figured out what they are bidding for the ad impression the ad exchange will receive the bids. The auction takes place in real-time and is conducted in such a way that each participant only can bid once. This is called a second-price auction.

The winning advertiser is notified by the ad exchange, through their DSP of choice, and are provided with the ability to send the code to the user’s browser that will set up the ad, together with a cookie to identify the user at the next crossroads.

There are many actors offering ad exchanges. Almost all big internet actors offer their own ad exchange (Datatilsynet, 2015).
2.1.6 Tracking and identifying methods

There are several methods for identifying users through different webpages. We have mentioned the most common methods for identifying and tracking users, IP, and cookies, but have not explained how it is done or what it is. There are other ways of tracking users and user behavior, not mentioned so far. Tracking and identifying users are vital to the online ad industry and advances in this area adds value to both publishers and advertisers. This builds on the principles of “Real value comes with real identity” (Busch, 2016).

It is important to understand what identifying really implies. By confirming that two or more set of data belongs to the same user, for example through the use of cookies, we have identified and can keep track of the user.

The next section will describe the most common methods. A comparison is presented in Table 3 at the end of the chapter.

IP tracking

Everything using the internet has an IP address, which makes it a valuable tracking tool. The IP address is the first piece of information the publisher receives about a user visiting their website. The user often operate from the same IP address over a period, meaning it is possible to identify and track users based on repeated visits from the same IP. In addition, looking up the IP address of a user will give information about where the user is located (Evans, 2009).

Despite its advantages, IP tracking has some dominating shortcomings, resulting in the fact that it is not suitable for tracking users over a longer period of time. The main reasons for this are:

- People using the same router will, in general, have the same IP address. This means that people in the same home and workplace will have the same IP address.

- One user using the same device will have different IP addresses when using the internet at home, over cellular networks, and at work. If tracking was done solely with IP this would appear as different users.
• The IP address of a stationary network will change over time. Internet providers usually do not give private homes static IP addresses. They change unpredictably, which makes IP tracking hard as a permanent solution.

• It is possible for users to change IP address, and appear as if they are someone accessing the internet from the other side of the world, through the use of Virtual Private Networks (VPN).

Because of the above-mentioned arguments, IP tracking provides limited opportunities for webpages to learn about users. Using information from this method to make inference about the user, such as age, gender and interests, is hardly possible.

However, using IP tracking together with other tracking methods enables IP tracking to be a valuable tool. When relying on other methods for identifying users, IP tracking is often used to provide insight into where users are located, and which users are located in the same place (connecting from the same IP address) (Evans, 2009). It can also be used to quickly identify if a user is traveling, through looking up the IP address. IP tracking is also used as a fallback method of tracking if other methods fail.

**Tracking cookies**

A cookie is a term used for a small file stored in the user’s web browser while the user is browsing a specific website. When a webpage’s server sends a file to the user’s web browser and this file is stored, this is known as a “cookie insertion” (Evans, 2009). A cookie usually stores information used to identify the user. When the user navigates to the same webpage that inserted the cookie, the user’s web browser transmits the cookie back to the webpage’s server, and the webpage can identify the user and modify its behavior accordingly. Only the webpage that inserted the cookie has access to read it. Some cookies are deleted after the user’s “session” is over, other cookies can exist in a user’s browser for several years (Datatilsynet, 2015).

Cookies can only be inserted by the webpage’s server but can be controlled by other parties. This is what creates the distinction between, what is known as, first-party and third-party cookies. First-party cookies are cookies inserted and controlled by the webpage owner. Third-party cookies are placed by the webpage owner but controlled by a company other than the website owner (a third-party). Third party cookies are usually inserted on behalf of data
analysis companies looking to gain insight into the browsing habits of their users. By having third-party cookies on a huge number of webpages, these companies can identify individual users as they navigate different webpages and collect data about users’ navigation and behavior on the internet.

The use of cookies for identifying and tracking does have several disadvantages (Datatilsynet, 2015):

- Cookies cannot track users across devices, because cookies are inserted into web browsers. This means that different browsers will appear as different individuals.

- Cookie tracking requires the webpage to inform the user that cookies are used, and what they are used for. The user is also required to give their consent.

- Users can reserve themselves against the use of cookies. Web browsers enables the user to stop webpages from inserting cookies. Both the use of first-party and third-party cookies can be restricted altogether. Users also have the option to delete some, or all of their cookies. Both these acts mean that webpages lose their ability to track the individual user. Deleting cookies enables webpages to insert a new cookie and start tracking the user again but preventing the insertion of cookies means that the webpage must find other solutions to identify and track the user.

Identifying users based on cookies is an improvement over IP tracking but does not yield factual information. The information gathered about a user is inferred knowledge and its value will therefore not be as high as it could be. Data analysis companies apply the information they learn from the browsing habits of the individual and deduct age, gender and interests to build an information profile for all users.

**Tracking pixel**

Pixel tracking, also known as a web beacon, is an invisible pixel-sized graphic image embedded in a webpage by the website owner. Tracking pixels are often used together with other identifying methods, such as IP tracking and cookies. It works by making the user’s web browser request the hidden “image” from the advertiser’s server. This is done by telling the web browser that an image should be shown to the user, and it can be found at a specified web address. The web browser then sends an HTTP request to the server to fetch the image. When
this done, the browser passes data to the server, along with the request for the image. The image is never shown to the user, it is used as a way of “tricking” the web browser to send a request to the advertisers’ server. The data passed to the advertiser is often previously inserted cookies, as well as IP information and other data gathered about the user, such as type of web browser, device and language settings.

Tracking pixels can be used by both first-party and third-party companies. By third-party companies, it is used as a method to insert and retrieve cookies, as well as other information. The snippet of code that tells the web browser to request an image must be placed by the webpage owner (first-party).

Tracking pixels can be used to extract more data than other tracking methods mentioned above. It can be used to provide information on the actual behavior of the individual user and their activities on the webpage (Evans, 2009). This method relies on cookies and IP tracking for identifying and categorizing the user and does have the same weaknesses as its utilized methods for identifying and tracking users, the biggest of which is the lacking ability to track users across devices.

In contrast to cookies, it is not possible for users to reserve themselves against pixel tracking. However, since the method relies on cookies and IP addresses to identify the user, pixel tracking does not yield much of value if the user has blocked cookies (and use a VPN).

**Unique ID**
Collecting data by identifying individual user’s visits and activities when browsing the internet does increase the individual’s value on the ad exchange, but there is still more to be gained by aggregating the data and tracking across devices. Without the ability to follow the user between devices, the user for sale on the ad exchange is not actually the user but a presence on one device. This is a problem because advertisers cannot target the individual on more than one device, and lack information about the user’s behavior on other devices. Solutions to solve this and, in turn, build better user profiles is to identify the users by making them log in are created by some of the large internet companies. Facebook was the first large company to create a login solution with unique identifying of the user; and other actors, such as Google, Microsoft, and Amazon, have later done the same. These solutions connect data from various devices through identifying the user’s different devices when the user logs in from a new device (Datatilsynet, 2015). The user is given a unique ID which is applied to the
cookies stored on the various devices of the user, to better be able to track him/her in the future. When the user is logged in he/she is often asked to provide additional information, such as age, full name, and address. This is also used as a data collection method.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>IP tracking</th>
<th>Cookies</th>
<th>Tracking pixels</th>
<th>Unique ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requires consent</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes (consent is given when using the service)</td>
</tr>
<tr>
<td>Tracking between devices</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Possible to hinder</td>
<td>Yes (through VPN)</td>
<td>Yes (through the web browser)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Type of data collected</td>
<td>Individual location data and users in the same location</td>
<td>Browsing habits through tracking yields deducted information such as age and interests.</td>
<td>Actual behavior and activities on the webpage</td>
<td>The unique ID and additional information that was given by the user.</td>
</tr>
</tbody>
</table>

Table 3: Comparing online methods for identifying users

### 2.1.7 Cookie matching

An integral part of programmatic advertising is the ability to know which ad impression is for sale and cookie matching (also known as Cookie Syncing) is used to solve this problem. For buyers to decide if they want to bid on an ad impression or how much they want to bid, they have to know how much the user is worth to them. This is done by identifying the user in a way that lets the buyer create a match with one of the users in their database. This is not straightforward because the buyer and advertiser do not have the same identifiers and information on the same user. This means that even though both actors have stored information about one user it is no guarantee that they have stored the same data, or that this data is sufficient to know that this is the same user. This is solved by linking the profile of a single user in the databases of two independent companies, called Cookie matching. One example of the use of cookie matching in programmatic advertising can be (Datatilsynet, 2015):
1. User X is put up for sale on the ad exchange.

2. The demand side platform (DSP) is given access to data about user X by the ad exchange.

3. The DSP uses cookie matching and determines that user X is the same as user Z, which the DSP has stored a profile on at an earlier point in time.

4. The DSP is now able to accurately determine how much an ad impression to the user is worth for them.

To accurately determine that user X is the same as user Z (step number 3 above), a cookie matching table is used. The cookie matching table maps the different IDs a user is known as into one table. One example of a cookie matching table could look like the one presented below (Zawadziński, 2015):

<table>
<thead>
<tr>
<th>Table ID</th>
<th>Actor A’s ID</th>
<th>Actor B’s ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X1</td>
<td>Y1</td>
</tr>
<tr>
<td>2</td>
<td>X2</td>
<td>Y2</td>
</tr>
</tbody>
</table>

Table 4: Example of a Cookie Matching Table

In this example, there are three columns representing different IDs one user is known as by different actors. The first column is an ID used by the specific cookie matching table to identify the user. The next column is Actor A’s internal ID for this user. The last column is Actor B’s internal ID for this user. Each row in the table represents a new user. This table would then allow for Actor A and B to determine that user X1 and Y1 are, in fact, the same user. The same goes for user X2 and Y2.

The challenge of cookie matching, however, is populating the cookie matching table. Determining what IDs belong to the same users is not a trivial task in programmatic advertising. As the name implies, the process relies on cookies and it starts when a user visits a webpage that contains programmatic advertising. The process is as follows (Zawadziński, 2015):
1. A user visits a webpage containing programmatic advertising

2. The browser sends a request to the ad exchange. The ad exchange examines if this user has been identified by the platform before. If the ad exchange has seen this user before it fetches the user’s ID. If this is not the case the ad exchange creates a unique user ID. This ID is then stored in a cookie in the user’s browser.

3. After having created a cookie, the ad exchange redirects the request to a predetermined actor. This actor is usually a part of the ecosystem for programmatic advertising connected to the specific webpage. Along with the redirect, the ad exchange will pass on the ID used to identify the user.

4. The predetermined actor reads the ID created by the ad exchange and fetches its own cookies stored in the user’s browser containing an ID. The actor does now have both the ad exchange’s user ID and its own user ID for the same user. These are now stored in the cookie matching table belonging to the actor.

5. The actor redirects back to the ad exchange and passes its user ID along.

6. The ad exchange reads the user ID passed along together with the redirect and fetches the cookie stored in the user’s web browser. The ad exchange does now have both user IDs and stores these in its own cookie matching table.

However, there are some problems with cookie matching because of the possibility for data leakage in the process described above. During the process of cookie matching, there has been claimed that as much as 27% of the user’s web history can be leaked to the DSP (Datatilsynet, 2015). This is a problem and represents a challenge with cookie matching.

### 2.1.8 Summary

Online advertising is a term used for all advertising delivered over the internet. It has evolved from being an online equivalent to newspapers to programmatic advertising. Programmatic advertising relies on user data to improve the value of ad impression.

The actors involved in programmatic advertising are vendors and buyers of advertising space, as well as actors involved in the ad exchange and software providers. The software providers create the software services that make programmatic advertising possible and the most central
components are the SSP, DMP and ad exchange. Other software parts that are vital for programmatic advertising are tracking and identifying technology and cookie matching.

### 2.2 Online information privacy

Information privacy refers to an individual’s claim to control over when, how, and to what extent information about them is shared with others (Malhotra, Kim, & Agarwal, 2004). As established, almost every actor involved in online advertising collects and stores a great amount of information about the users. The data collected about each user is often extensive and include personal data. The companies collecting this data often store it over an extended time period and processes the data for various purposes (Evans, 2009). This data is valuable to the companies, and some actors earn a living collecting and trading data to others. Almost all parts of this widespread practice violate a user’s information privacy because the user has no control over when, how or the extent of which their information is used.

Online information privacy is relevant for this thesis because it is tightly related to online advertising. We will in this section define “user data” for future use in this thesis. Furthermore, we will present the economics of user data as a way of providing understanding for why user data is collected and why violation of users’ information privacy is important. Lastly, we will present the General Data Protection Regulation (GDPR) because it is created to protect the online information privacy of EU citizens and have the potential to impact the online advertising industry greatly.

#### 2.2.1 Defining user data

User data is all the data collected about a user, which is also the short definition of personal data. This definition is not very precise and this thesis will, therefore, use the definition of personal data that originates from Article 4 in the General Data Protection Regulation (GDPR). A description of the GDPR will be given in chapter 2.2.3.

Personal data is defined as *any information relating to an identified or identifiable person*. A person in this context refers to a human being, and not a “legal person” which could be a firm. The important part of this definition are these two elements “any information” and “relating to”. “Any information” must, in this case, be taken literally. The information can be a name, an identification number, location data, an online identifier, one or more factors specific to the
physical, physiological, genetic, mental, economic, cultural or social identity of that person. Other examples of identifiers are information about the address, phone number, email, IP address, pictures, fingerprint, iris, or head shape. The second part, “relating to”, means that the information must be in relation to an identified or identifiable person. Having a list of street names would not be considered to be personal data. Adding a first and last name to the list of street names and the data becomes personal. A list of first names would, by itself, not be personal data, it could just be a list of the most common names. Adding information such as eye color, hair color, or height, and the data is now considered personal.

With a definition of personal data, it is also important to have a clear definition of an amount of personal data. The amount of data can be measured in two dimensions. The first is the number of persons, and the other is the amount of data that is collected about each person. Increasing the amount of data would then mean that either the number of persons is increased, or that the amount of data about one or more person is increased. The total amount is increased if either of the dimensions are increased.

2.2.2 Economies of user data

The reason for companies to collect user data is the potential money to be made by analyzing it. Zuboff (2015) describes this phenomenon and addresses it as “Surveillance Capitalism”. The foundation for surveillance capitalism was the emergence of Big Data and its possibilities. Big Data created an opportunity to extract value from large data sets. This motivated companies to increase the amount of data collected about their consumers, with an expectation that it would benefit them in the context of Big Data. This has spiraled into the common belief among companies that the actor that possesses the most user data has the power to produce the most revenue and control the market.

Companies offering their services online has additional motivation for the collection of user data because of the personalization of their services (Awad & Krishnan, 2006). This is done to increase the consumption of their services by creating a better product for the individual consumer. To achieve this data about the consumer is collected and stored.
2.2.3 General Data Protection Regulation

Understanding the General Data Protection Regulation (GDPR) is important for online advertising and online information privacy. The GDPR is a regulation intended to replace the old Data Protection Directive from 1995. It was designed to unite and standardize data privacy laws across Europe, to protect and increase all EU citizens’ control of their data privacy and change the way organizations approach data privacy. The GDPR is, in simpler words, a regulation where organizations will get stricter and more rules to follow with regards to data privacy, and EU citizens will get more power over their own data. The reason for this new regulation is because the world is becoming increasingly data-driven compared to the time in which the 1995 directive was established (European Commission, 2016). A brief explanation of the key changes will be given below. Keep in mind that these are just short summaries.

For the purpose of these summaries, a quick explanation of the most used terms will first be given:

- **A natural person** is a real human being.

- **A legal person** is, in contrast to a natural person, not a real human being, but instead is a non-human legal entity such as a public or private company.

- **The data subject** is an identifiable natural person who can be identified, directly or indirectly, by personal data.

- **The controller** is the entity, being it a natural or legal person, agency, or another body which, alone or in collaboration with others, determine the purpose and means of the processing of personal data.

- **The processor** means a natural or legal person, agency, or another body which processes personal data on behalf of the controller.

**Extended jurisdiction**

GDPR comes with extended jurisdiction, as it applies to all companies processing the personal data of EU citizens, regardless of the company’s location. This means that it will apply to the processing of personal data by controllers and processors in the EU, regardless of
whether the processing takes place in the EU or not. It will also apply to controllers or processors not established in the EU, so long as the processing of personal data is of data subjects residing in the EU.

(Article 3)

Penalties

The consequences of not being able to comply may be severe “Under GDPR organizations in breach of GDPR can be fined up to 4% of annual global turnover or €20 Million (whichever is greater)”. It’s important to note that this is the maximum fine that can be imposed for the most serious breaches. These rules apply to both controllers and processors.

(Article 83)

Consent

The requirements for collecting consent has been strengthened. Companies will not be able to use illegible terms and conditions full of nonsense. The request for consent must be given in an intelligible and easily accessible form, and the consent should cover all processing activities carried out for the same purpose or purposes. The consent must also be as easy to withdraw as it is given.

(Article 7)

Breach notification

A data breach is an incident where private information is taken from a system without the knowledge or authorization of the system’s owner. If a data breach is likely to result in a risk for the rights and freedom of individuals, notification will be mandatory in all member states of the EU and should be given not later than 72 hours of first having become aware of it. Data processors will also be required to notify their customers, the controllers, without unnecessary delay after first becoming aware of a data breach.

(Article 33)

Data transparency

The data subject has the right to obtain confirmation from the data controller as to whether or not personal data concerning them are being processed, also where and for what purpose it is being processed. The data controller shall also be able to provide a copy of the personal data,
free of charge, in an electronic format.

(Article 12)

**Data erasure**
The data subject has the right to be forgotten. This entitles the data subject to have the controller erase his/her personal data without undue delay, inform potential third parties to halt data processing concerning the data subject. The conditions for erasure include that the data no longer are necessary in relation to the original purpose, or a data subject has withdrawn consent, or the personal data have been unlawfully processed. The data subject has no right to erasure if the data includes information concerning the public interest, such as criminal records.

(Article 17)

**Data portability**
The data subject has the right to obtain personal data in a structured, commonly used and machine-readable format. They have the rights to transmit that data to another controller without hindrance from the controller to which the personal data have been provided.

(Article 20)

**Privacy by design**
Privacy by design is not an entirely new concept, but it has now just become part of a legal requirement because of the GDPR. Privacy by design has the intended effect to make sure the controller “implement appropriate technical and organizational measures… in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects”. Only personal data which are necessary for each specific purpose of the processing are processed (data minimization).

(Article 25)

### 2.2.4 Summary

To summarize, we establish that user data is defined as any information relating to an identified or identifiable person. One example of user data could be a list of addresses together with the first and last name of the residents. User data is collected by companies
because there is a potential to earn money by analyzing it. This has spiraled into the common belief among companies that the actor that possesses the most user data has the power to produce the most revenue and control the market. The introduction of GDPR aims to do something with this trend and is a regulation where organizations will get stricter and more rules to follow with regards to data privacy. A company failing to comply with the GDPR could be fined up to 4% of annual global turnover or €20 Million (whichever is greater).
### 3 Theory

The theoretical framework in this thesis is focused on supporting our understanding of how the case under study can create value for local actors in the online advertising industry. The purpose of this framework is to examine how value can be created for the actors joining a software service. This is used to understand the potential for actors in the Norwegian online advertising industry to build a competitive advantage through participating in the case under study. First, theories of **value networks and software-based platforms** are used to describe the case under study and how value creation is achieved. Second, **digital innovation** is introduced to discuss the potential for local actors to increase their competitiveness and its effect on value creation. Third, the **internet users’ privacy concerns** are presented to discuss how the case under study can mitigate user concerns because this is viewed as a potential factor for success in the online advertising industry.

#### 3.1 Value network

Theory of value network will provide insight into how networks create value. Value network theory originates from a business perspective as a generic value configuration model used to understand and analyze firm-level value creation logic across a broad range of industries and firms (Stabell & Fjeldstad, 1998). This will be used to understand how the case under study could increase the competitiveness of Norwegian software providers in online advertising.

**Defining a value network**

Stabell and Fjeldstad (1998) outline that a value network is the use of communication technology to link actors that wish to be inter-dependent. A network creates value in its facilitation of information exchange between actors. Many different value creating relationships can exist in a network. How valuable the network is depends on the relationships it can facilitate. The value of the network will vary from actor to actor depending on their perceived value of the other actors in the network.

An important distinction is that one firm cannot be the network but can facilitate and provide a networking service. For the firm facilitating the network all participants in the network are viewed as customers and the firm does usually not participate in the network. The facilitating firm has a managing role in the network. The firm defines the potential relations between the
actors in the network. This is done by creating the services and protocols used in interactions and in this way manage what the network can be used for and, to some degree, the relations within it. The managing firm does also have a role in enrolling new actors in the network. To become a part of the network actors would have to enter into an agreement with the managing firm. This agreement has the potential to exclude actors not wanted by the firm or actors not agreeing to the terms and has, therefore, the potential to greatly affect the potential relations in the network. One example of this is the transportation network company Uber, where they exclude drivers that do not have a driver’s license as potential members of their network. This alters the relationships in their network by avoiding potentially unsafe relations between drivers and passengers.

A value network will typically have low value to the initial customer, and at the same time have a high cost associated with the creation and introduction of the network. This makes the rollout phase of a value network inherently challenging.

Kothandaraman and Wilson (2001) add to our understanding gained from Stabell and Fjeldstad (1998) by establishing that value networks have three core building blocks. The three building blocks are important because they help us to better determine the potential of a value network. The model in Figure 3 outlines the relations between the three building blocks:

![Figure 3: A model of value networks (Kothandaraman & Wilson, 2001)](image-url)
The figure above and its relations can be understood as:

- **Superior customer value.** This is the objective of a value network. Customers are in this case the actors joining the network. It should, therefore, have a superior value for the involved actors compared to alternatives outside the network. Superior customer value determines the core capabilities and reinforces the relationships in the network.

- **Core capabilities.** The core capabilities of the actors in the network are what creates superior customer value. Actors in the network combine their capabilities to create additional value. Core capabilities constrain the relationships in the network and create superior customer value.

- **Relationships.** The relationships between the actors in the network define how they combine their capabilities to create value. Relationships facilitate superior customer value and maintain the core capabilities.

**Network effects**

It is important to understand how networks add value and the effect additions to the network have on existing members. This is used to analyze the implications of network effects might have on the case under study.

Katz and Shapiro (1985) outline that the foundation of a network is that the value of a product increases with the number of other actors consuming the same products. Network effects are the effects that appear in a network. Liebowitz and Margolis (1994, p. 135) define it more precisely by stating that “The circumstance in which the net value of an action (consuming a good, subscribing to telephone service) is affected by the number of agents taking equivalent actions will be called a network effect”. One of these effects is that actors joining a network increase the value of the network for both the actors already enrolled and the actors considering joining. This thesis will use two of the three network effects outlined by Katz and Shapiro (1985):

- **A direct effect of the increased number of consumers.** This effect is the direct result of an increased number of actors in a network. An example of this is telecommunications companies where an additional user on the network increases the
value of the network for the consumers already enrolled on the network and potential new customers.

- **An indirect effect of the increased number of consumers.** This is the indirect effect an increased number of actors in a network. An example could be that the number of apps available to the iPhone increases when more people buy the iPhone.

**Summary**

This thesis will use the theory of value networks and direct and indirect network effects to analyze what effects could be present in the case under study and how these might affect the outcome.

### 3.2 Software-based platforms

Theory of *software-based platforms* will provide insight into how software platforms create value and how software platform ecosystems can become a competitive advantage. The concept of platforms comes from engineering design (Baldwin & Woodard, 2008), and we define therefore a software-based platform to be software engineering design. Tiwana (2013) refers to software-based platforms as a technology solution. This will be used to understand how the case under study could increase the competitiveness of Norwegian software providers within online advertising. We will also examine if the case under study could be considered a software platform.

Tiwana (2014) presents that a software platform is software designed as a foundation on to which complementary software products can be built. Platforms are designed to attract a diverse developer community motivated to produce the software complementing the platform. Developers are attracted to the platform by providing them with access to potential software consumers and the use of the platform’s components. By doing this a platform can attract developers outside the firm and this way develop new software faster and with capabilities unforeseen by the firm creating the platform. The platform and the actors surrounding it is called a *platform’s ecosystem*.

A software platform consists of a distinct set of elements (Baldwin & Woodard, 2008; Tiwana, 2014). According to Tiwana (2014) the most fundamental of which is the platform and complementary software that connects to the platform to extend its functionality (called
apps). Baldwin and Woodard (2008) argue that the infrastructure connecting the platform and the complementary components should also be considered as part of the core components of a platform. The platform contains functionality which the apps share and the infrastructure the apps interoperate through. A platform’s architecture describes the ecosystem. In addition to these components, a platform has end-users, rival platforms ecosystems, and a competitive environment (Tiwana, 2014). The end-users are all current and potential adopters of the platform. Rival platforms ecosystems are the competition from other platforms and their ecosystem, which makes up the competitive environment the platform exists in.

Many definitions of platforms have focused on the act of sharing assets or components across multiple products (Bresnahan & Greenstein, 1999, p. 4; Robertson & Ulrich, 1998, p. 20; Simpson, Maier, & Mistree, 2001, p. 3). Eisenmann, Parker, and Van Alstyne (2006) and Tiwana (2014) add to this understanding and argue that a platform has to facilitate interactions between two distinct groups (two sides of a market). To examine if the case under study could be defined as a platform we will base our definition on the sharing of assets across multiple products and the facilitation of two-sided markets.

### 3.2.1 Two-sided markets

Establishing that platforms need to facilitate two-sided markets has implications for the theory on how competitive advantages can be achieved through platforms. The theory is used to identify the two market sides of the case under study and gain insight into how this might affect the competitiveness.

Two-sided markets consist of two distinct groups that benefit from each other. The platform connects the two sides of the market. One side of the software-based platforms is the developer community and the other side consists of consumers. One example of a platform in two-sided market can be a video game console platform. On this platform, the two sides are the players and the creators of video games. The sides depend on each other for the platform to succeed. The platform will not attract persons interested in playing video games if no games are created. On the other side, no video game developers will invest time and money to create games for the console if there are no potential players.

There is an overall agreement that a platform ecosystem in a two-sided market will always face the “chicken-or-egg problem” (Armstrong, 2006; Jullien, 2001; J.-C. Rochet & Tirole,
This means that for the platform to succeed and the ecosystem to thrive both sides of the market must be present. Hanseth and Lyytinen (2010) address this problem in Information Infrastructures but calls it the “Bootstrap problem”. They propose three design principles to tackle the bootstrap problem. (1) design initially for usefulness, (2) draw upon existing installed base, and (3) expand installed base by persuasive tactics. These design principles can also be applied to bootstrap a platform.

Network effects are present in two-sided markets, as outlined by Parker and Van Alstyne (2005). They discuss that it is possible to have networks on both sides of the market and that they might have both direct and indirect effects on each other.

3.2.2 Summary

This thesis will use the theory of software-based platforms to explore how the case under study could increase the competitiveness of Norwegian software providers in online advertising. This is done by leveraging the competitive advantage of a platform’s ecosystem. A platform’s architecture can affect how successful the platform’s ecosystem will be.

Theory of two-sided markets is presented to complement platform theory because a platform connects two sides of a market. It is therefore used to investigate the challenges this brings for the case under study. We also establish that network effects can present on both sides of a market and might affect each other.

3.3 Digital innovation

The theory of digital innovation can be used to gain understanding of innovation processes where digital technology is central. The theory could provide insight into why networks and platform theory can facilitate digital innovation and the implications that this might have.

In the research commentary put forward by Yoo, Henfridsson, and Lyytinen (2010) digital innovation is defined as: “The carrying out of new combinations of digital and physical components to produce novel products”. A similar definition is used by Lyytinen, Yoo, and Boland Jr (2016, p. 49) who defines digital innovation as “significantly new (from the perspective of a particular community or market) products or services that are either embodied in information and communication technologies or enabled by them”. An important
distinction by Yoo et al. (2010) and Lyytinen et al. (2016) is that their definition focuses on product innovation and that digital innovation is concerned with digital technology. It is therefore important to understand how digital technology differs from analog technology. Yoo et al. (2010) present three main characteristics separating digital and analog technology:

1. Reprogrammability enables one device to perform a variety of different tasks.

2. Homogenization of data enables the same device to store, transmit, display and process all digital content.

3. The self-referent nature of digital technology accelerates the creation and availability of digital devices.

Digital innovation differs from earlier innovation by utilizing the characteristics above. Yoo et al. (2010) argue that this is important because these characteristics make a layered and modular architecture a possibility and change the nature of innovation. One example of this is the success of the Amazon Kindle, which demonstrates how the digitization of the book changed a competitive landscape.

Modular and layered architecture creates a base for digital innovation where stable core components and rapidly changing complementary components interact through standardized interfaces between them (Nielsen, 2017). This offers opportunities for combining components in new ways and create additional value.

A layered and modular architecture also creates an opportunity for innovation to be distributed. This is done by combining different actors’ components in new ways. It allows for innovation to happen based on new relationships between multiple actors. Digital innovation is about breaking up an industry’s existing silo structures in a way that create networks and bring actors together to combine digital technology in new ways. These networks typically consist of heterogeneous actors that come together and interlink complementary digital components (Nielsen, 2017).

Building on Yoo et al. (2010), Lyytinen et al. (2016) and Nielsen (2017) we can think of digital innovation as a process where existing silos structures are broken down and networks are created, enabling the creation of significantly new combinations of complementary digital components to create additional value.
Digital innovation theory is relevant for our research because digital innovation can provide an understanding with regards to observations made during data collection and could enhance insight into the case. The potential of digital innovation, if the theory could provide insight into the case under study, is that local actors could use this to change their current situation and increase their competitiveness in the online advertising industry in Norway.

3.4 Internet Users’ Information Privacy Concerns

The Internet Users’ Information Privacy Concerns (IUIPC) framework presented by Malhotra et al. (2004) is important for this research to support our understanding of potential hinders for value creation in the online advertising industry. The theory is chosen because it allows us to understand the concerns individuals have when it comes to online advertising and their information privacy. This is important because mitigating user concerns can be argued to be a success factor for increasing the competitiveness of local software providers in the online advertising industry.

Information privacy concerns is a reference to the subjective view of an individual on fairness in the context of information privacy (Malhotra et al., 2004). Malhotra et al. (2004) offer a theoretical framework on the dimensions of the concerns internet users have for their information privacy. Their theory is based on Social Convention (SC) theory and state that when this theory is applied to information privacy it would suggest that:

“(…) a firm’s collection of personally identifiable data is perceived to be fair only when the consumer is granted control over the information and the consumer is informed about the firm’s intended use of the information” (Malhotra et al., 2004)

The framework is based on the concept of a fair exchange. A user gives up something with an expectation to get something in return. If the value of what the user received after giving up information is less than perceived drawbacks or potential negative outcomes, then the exchange is not regarded as fair. This can be discussed in the context of distributive, procedural, and interactional/informational justice.
This can then be broken down into three main factors to characterize IUIPC:

- **Collection (distributive)**
  This is the initialization of concerns regarding information privacy. Without an actor collecting the data, legal or illegal, there would be no concerns. In relation to the IUIPC framework it refers to:

  “The degree to which a person is concerned about the amount of individual-specific data possessed by others relative to the value of benefits received” (Malhotra et al., 2004)

  This is rooted in the concept of fair exchange and collection is when a user gives up something and therefore expects something of value in return. If the user expects negative outcomes, the collection of data will be increasingly associated with concerns for the individual user. It is also argued that there exists a “privacy threshold” level for the collection of data users are willing to endure.

- **Control (procedural)**
  Individuals perceive control of the collected personal data to be of importance. Control is usually exercised through approval, modification, and withdrawal of personal data. When the social contract of fair exchange is broken and there exists an opportunity for commercial actors to gain a profit from the collected data, the requirement for control becomes greater for the individual user. The framework of Malhotra et al. (2004) proposes that:

  “An individual’s concerns for information privacy center on whether the individual has control over personal information as manifested by the existence of voice (i.e., approval, modification) or exit (i.e., opt-out)”

  It is also demonstrated that when users are given the explicit choice to opt-out or must give a company permission before data collection, they become less concerned about the collection of data. Users want to be able to control the information collected about them, and in that way influence policies in companies that they find intolerable by opting-out of such agreements.
• Awareness of Privacy Practices (interactional/informational)
  The awareness of Privacy Practices refers to the degree of concern a user has towards
  the lack of information provided about an organization’s privacy practices. When
  users are made aware of how the organization uses individual-specific data their
degree of concern towards the trade tends to lessen. Failure to make users aware that
  their data is shared with other actors, stored for a long time, or used in marketing
could all be sources of concern for internet users.

  These elements can be summarized as; “whether the exchange of personal information is
equitable” (collection), “whether I have control over the data” (control), and “whether I am
adequately informed about the use of the data” (awareness).

  Malhotra et al. (2004) created a measurement model based on control, collection, and
awareness, as well as previous scales related to internet users’ concerns. The model consists
of ten items:

  1. Collection of data
  2. Errors in the data collected data
  3. Unauthorized secondary use of the data
  4. Improper use of the data
  5. Awareness of privacy practices
  6. Control of the data
  7. Global information privacy concern
  8. Trusting beliefs
  9. Risk beliefs
  10. Intention to give personal information

  The score related to IUIPC was calculated from a series of questions within the ten topics of
the model; a high score would mean a high concern. Even though the IUIPC framework was
developed to measure the change in the individual user’s perceived fairness of the widespread
use of the internet, it could be used as a general framework because of its solid roots in SC theory. The framework can be argued to be applicable in most privacy-sensitive settings and could, therefore, be used to gain insight into how a solution or innovation might affect the concerns users possess towards online marketing. By examining the framework and the ten topics in relation to our thesis, we can use the topics as a foundation for discussing how the case under study potentially could mitigate users’ concerns.

### 3.5 Summary

In this chapter, we have built a theoretical framework based on concepts of value networks, software-based platforms, and digital innovation – summarized in the table below. The aim of this framework is to support our understanding of value creation in the case under study and the potential for actors in the Norwegian online industry to build their competitive advantage. Complementing this framework, we present the IUIPC theory to explore a potential success factor for value creation specific for the case under study.

<table>
<thead>
<tr>
<th>Theoretical concept</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value networks</td>
<td>The value of a network increases with the number of other actors enrolled.</td>
</tr>
<tr>
<td></td>
<td>A value network creates value in its facilitation of information exchange</td>
</tr>
<tr>
<td></td>
<td>between actors. Network effects increase a network’s value for actors in</td>
</tr>
<tr>
<td></td>
<td>the network and are divided into direct and indirect effects.</td>
</tr>
<tr>
<td>Software-based</td>
<td>A software platform connects two sides of a market by establishing a</td>
</tr>
<tr>
<td>platforms</td>
<td>foundation on to which complementary software can be built.</td>
</tr>
<tr>
<td></td>
<td>Value is created by providing developers of complementary software with</td>
</tr>
<tr>
<td></td>
<td>access to consumers and consumers are provided with additional</td>
</tr>
<tr>
<td></td>
<td>capabilities from the complementary software.</td>
</tr>
<tr>
<td>Digital innovation</td>
<td>Digital innovation enables both the breaking down of existing silo structures</td>
</tr>
<tr>
<td></td>
<td>to create networks and the distribution of innovation between multiple</td>
</tr>
<tr>
<td></td>
<td>actors.</td>
</tr>
<tr>
<td></td>
<td>Digital innovation generates value in its creation of significantly new</td>
</tr>
<tr>
<td></td>
<td>combinations of complementary software. This is enabled by layered and</td>
</tr>
<tr>
<td></td>
<td>modular architecture. Digital innovation establishes a potential for local</td>
</tr>
<tr>
<td></td>
<td>actors to change their situation and increase their competitiveness.</td>
</tr>
</tbody>
</table>

Table 5: Summary of the theoretical framework used in this thesis
Value network originates from a business perspective as a model used to understand value creation logic across a broad range of industries and firms (Stabell & Fjeldstad, 1998). The concept of software-based platform is viewed as software engineering design and is referred to as a technology solution (Tiwana, 2013). Digital innovation is viewed as an innovation process where digital technology is used.

The theories of digital innovation (a process), value network (a model), and software-based platform (a technological solution) are used in relation in this thesis to form the value creation framework. The theories were chosen because they can enhance our understanding of different aspects of value creation. The theories complement each other and by linking them together we can explore value creation in the case under study. The relations between these theories are described below:

- Value networks enable software platforms by being one side of the software platform, making it attractive to creators of complementary components. Value networks enable digital innovation by allowing actors to combine complementary digital components in new ways.

- A software-based platform increases the number of actors in the network through access to complementary components. The platform enables further digital innovation by providing third-party developers with access to consumers and re-use of the platform’s components.

- Digital innovation creates networks by offering the potential to break down silo structures.

Because our theoretical framework is focused on value creation, we have chosen to include the IUIPC theory to explore potential blockers of value creation specific to the online advertising industry in Norway and is, therefore, a factor for success for the case under study. Because our case is situated in the online advertising market we use the Internet Users’ Information Privacy Concerns framework to support our understanding of this factor for success for actors in the Norwegian online industry to build their competitive advantage.
4 Research method

This chapter will present the chosen research method and our motivation for choosing it, types of data collection and technique for analysis.

This chapter will relate our research method to various parts of the case under study. We therefore wish to establish a distinction before continuing; the project “Digital Innovation To Increase Competitiveness” (DITIC) aims to create a software service. This software service is called Digisphere.

4.1 Identifying the research question

When choosing a research question the most obvious question is often not the best choice, as stated by Easterbrook, Singer, Storey, and Damian (2008). This holds true for our research question as well as it was not defined clearly from the start but was instead a continuous process. As this thesis focuses on the project, DITIC, which is in the initial stages, we chose an exploratory type of question as this felt best suited. The reason for this is that we wished to gain knowledge by attempting “to understand the phenomena, and identify useful distinctions that clarify our understanding” (Easterbrook et al., 2008). Our goal for the research was to explore and understand the challenges that the small actors face in the Norwegian online advertising market and viable solutions to these problems. This goal has therefore been the foundation for all iterations of the research question. Our research goal stayed the same throughout the research, changes to the research questions was thus a result of increased understanding of the research area. One example of an earlier research question is:

“How can digital innovation be used to level the competition in the skewed online advertising market in Norway?”

This question had a wide scope and lacked understanding of the position of digital innovation theory. It used the phrase “level the competition” which we later realized were used wrong. The question addressed the “skewed online advertising market in Norway” which were discarded at a later point because we understood that this was not actually what we were investigating. After several evaluations we eventually landed at the research question for this thesis:
“How can digital innovation enable increased competitiveness for local software providers in the Norwegian online advertising market?”

4.2 Qualitative research

The approach of qualitative research was decided early because the project under study and type of research question was chosen in the initial phase of this thesis. Furthermore, the situation would be described from the perspective of those involved with a focus on context and meaning. These attributes are among the typical features of a flexible research design strategy, also known as a qualitative strategy (Robson, 2002). This approach was chosen as an appropriate method for our exploratory research.

Robson (2002) outlines three design strategies within the flexible approach. These are case study, ethnographic study, and grounded theory study. Case study was chosen for this thesis, for reasons provided in chapter 4.2.2 Case Study.

4.2.1 Research design

Research design is the act of turning a research question into a research project. It is therefore crucial for answering the research question that the right approach is chosen (Robson, 2002). One way of defining an approach is through the framework provided by Robson (2002). Table 6 contains our research design which emerged from discussions with our supervisors.

<table>
<thead>
<tr>
<th>Component</th>
<th>Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>The study aims to (1) explore how online advertising works and the challenges faced by the Norwegian industry; (2) look at how the challenges can be met and (3) propose a set of design guidelines to maximize the attraction for a software service.</td>
</tr>
<tr>
<td>Theory</td>
<td>The study will use a theoretical framework based on concepts of value networks, software-based platforms, and digital innovation to support our understanding of value creation and attraction in the case under study.</td>
</tr>
<tr>
<td>Research question</td>
<td>This study will answer “How can digital innovation enable increased competitiveness for local software providers in the Norwegian online advertising market?”</td>
</tr>
</tbody>
</table>
Methods

The study will use various methods for collecting data. Semi-structured interviews, discussions and literary review will be used. In addition, we will participate in a workshop to gain further insight.

Sampling procedures

Data will be collected from the DITIC team continuously through the project period via discussions and documents. The interviews and workshop will be conducted at the beginning of the project. Information about the online advertising industry in Norway will be collected through interviews with actors holding various positions in the market. These interviews will be held towards the end of the research period.

Table 6: Our research design based on Robson (2002, p. 81)

4.2.2 Case study

A case study has a distinct advantage when the research question is on the form *how* or *why* (Yin, 2013). Case studies are also an effective way to gain deep insight and knowledge about an object of study (Flyvbjerg, 2006), case study was, therefore, a good fit to answer our research question. Braa and Vidgen (1999) argue that a goal of case studies is to understand and generalize beyond the individual case. For these reasons, the approach of a case study was chosen.

As stated by Easterbrook et al. (2008) a case study research may contain a single case or multiple cases. Even though a multiple case design can give greater validity, a single case can sometimes be sufficient. This is because a single case can be critical, extreme, or unique. Extreme and unique cases are often expected to yield interesting insights. Because of the time constraint in a master’s thesis, and the difficulties involved in finding a project with a similar goal, a single case design was chosen.

A case study can be conducted in two ways, exploratory and confirmatory (Easterbrook et al., 2008). The first one is used as initial investigations of some phenomena to derive new hypotheses and build theories, while the second one is for confirming existing theories and especially important for refuting theories. Because of the investigative nature of the research, this case study was conducted as exploratory and aims to build theories.

Action research was considered as a possible research method but was not chosen because our participation was considerably lower than first thought. This is described in more detail in chapter 4.2.3. Our participation in DITIC should not be neglected even though it was lower than first planned. This is important because it means that our research is not purely
observatory. We have discussed DITIC with the project lead and presented our thoughts. Our discussions have been with the intended purpose of gaining understanding and testing hypotheses. It can also be argued that we have established a relationship with the participants. These factors might have affected the outcome of the research. However, an inherent attribute of case study is that the researchers contribute to change (Braa & Vidgen, 1999). Examining our contribution to DITIC in the context of the trade-off between understanding and change, see Figure 4, concludes that the amount of change inflicted by our contribution is not enough to cause a reconsideration of our intended research method. The main argument for this is that the intention with the participation was to gain insight and understanding, not to change the outcome and that the participation was low.

![Method location](image)

**Figure 4: Method location (Braa & Vidgen, 1999)**

Without a well-defined unit of analysis, the data collected becomes less valuable and it is used to ensure that the study focuses on the intended phenomena. According to Easterbrook et al. (2008) and Flick (2014), the unit of analysis in software engineering might be a company, a project, a team, and individual developer, a particular episode or event. In our thesis the unit of analysis is Digisphere.

**4.2.3 Action research as a possible method**

The research method for this thesis was first considered to be Action Research. The reason for this was a planned participation in DITIC by being involved as developers. This participation would have given us a deeper technical insight to the project and could have given us an
opportunity for intervention with the purpose of improvements, while simultaneously studying the experience (Davison, Martinsons, & Kock, 2004; Easterbrook et al., 2008). However, this approach lost some of its relevance when our participation in DITIC was decided to be significantly reduced. This decision was made because of the time frame of this thesis and the project’s progress was slower than first expected.

4.3 Data Collection

Qualitative data is the preferred choice for a case study research. This is because a variety of different data sources offers rich insights into the case. The data for this thesis was collected in two parts. Part A was conducted first and focused on DITIC and the resulting software service, Digisphere. Here the main objective was to follow DITIC from the start, as well as understanding the problem that this project was trying to solve. Part B was conducted after part A was completed and is about understanding how the online advertising industry works today and the current challenges. This was done to get a better understanding of the context of Digisphere and if they were focusing on the correct challenges. The interviews were focused on the interviewees’ side of the market, what works and what doesn’t work in their respective field.

For the first part, understanding Digisphere, we had interviews with the people involved in DITIC, and several documents were provided to us for clarification and information about the project.

A considerable effort was put down to understand how the advertising industry is built up. This was done by literary review and gave insight to what questions we wanted to focus on in our interviews for part B.

4.3.1 Interviews

The interviews were conducted as semi-structured interviews with both external and internal sources. The interviews were structured around an interview guide which was created based on our theoretical knowledge of the industry, our research goal, and discussions with the project lead. Different guides were used for part A and B. An examples of interview guides can be found in Attachment A and Attachment B.
The interview subjects for part A were chosen based on their role in DITIC. One interview subject from both companies taking part in DITIC was selected. We started with interviewing the project lead, which referred us to a who he believed to be the right person to interview in the other company taking part in DITIC.

To locate interview subjects for part B a combination of contacting potential subjects after online research and snowball strategy (Vogt & Johnson, 2011, p. 368) was used. We started with researching actors in the online advertising industry in Norway online and contacted interview candidates via email. After the first interviews were established the interview object referred us to other potential subjects.

The interviews were recorded. This was to keep an audit trail and because it allowed us to focus on the interviewee. All the interviews lasted for about an hour and were held in Norwegian. Our focus during the interviews was not to follow the questions very strict but instead have an open dialog with the interviewee. We used the questions to keep track of what we had talked about and to make sure that all the questions were covered once we were done.

The recorded interviews were partially transcribed after we had finished interviewing all the participants. Since the interviews were in Norwegian the selected quotes that are used in the thesis have been translated into English and approved by the interviewees.

The positions of the participants in the online advertising industry are shown in Figure 5. The solid lines in the figure illustrate the relationship between components of the industry. Dotted lines place actors in relation to the components. Table 7 gives an overview of the interviews conducted.
Figure 5: Participant’s position in the advertising industry

<table>
<thead>
<tr>
<th>Date</th>
<th>Part</th>
<th>Who</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>08.09.2017</td>
<td>A</td>
<td>Startup Alpha</td>
<td>Karl Thomas, Jostein, Kjetil</td>
</tr>
<tr>
<td>18.10.2017</td>
<td>A</td>
<td>Startup Beta</td>
<td>Karl Thomas, Jostein, Per Elling (Startup Beta)</td>
</tr>
<tr>
<td>19.01.2018</td>
<td>B</td>
<td>Participant C</td>
<td>Karl Thomas, Jostein, Torgeir (Participant C)</td>
</tr>
<tr>
<td>26.01.2018</td>
<td>B</td>
<td>Participant B</td>
<td>Karl Thomas, Jostein, Stian (Participant B)</td>
</tr>
<tr>
<td>30.01.2018</td>
<td>B</td>
<td>Participant E</td>
<td>Karl Thomas, Jostein, Jan Morten (Participant E)</td>
</tr>
<tr>
<td>30.01.2018</td>
<td>B</td>
<td>Participant D</td>
<td>Karl Thomas, Jostein, Torgeir (Participant D)</td>
</tr>
<tr>
<td>31.01.2018</td>
<td>B</td>
<td>Participant A</td>
<td>Karl Thomas, Jostein, Peter (Participant A)</td>
</tr>
</tbody>
</table>

Table 7: Overview of interviews done in part A and B.

A total of seven interviews were conducted. Two of these were conducted in part A and five in part B.
Part A
The background work, part A, for this thesis began by interviewing two representatives from DITIC, one from Startup Beta and one from Startup Alpha. We asked the same questions to both participants to get an understanding of how Digisphere would help their position in the advertising market. By doing it this way we could get two perspectives of DITIC. The interview with Startup Alpha was held at Forskningsparken, while the interview with Startup Beta was done over Skype.

Part B
Five interviews were conducted in part B. The participants had different positions in the advertising industry, and in this way, we could get a different perspective and compare their answers. The reason for this approach was to get a broader understanding of the industry and the challenges the participants face.

A set of base questions were made as a blueprint for the interviews, this allowed us to modify the questions for each side of the advertising industry while keeping a base theme. The interviews were held at the interviewees’ workplace.

A detailed description of the participants is given below. Where relevant, age, size, location, and products they provide are described. This is done to provide detailed information regarding their position in the industry.

Participant A is a technology company that focuses on delivering services and software on the demand side of advertising, which gives them a central position in this market. They started in Stockholm, Sweden in 2002. They have more than 60 employees spread out between their offices in Oslo, Norway; Denmark; the Netherlands; and Germany.

They offer targeted and intelligent advertising solutions. The products they offer are:

- **Product A** is their demand side platform, it has access to the leading supply side platforms and ad exchanges and works as a display, video, and mobile advertising service station for planning purchasing, and tracking advertising in real-time.

- **Product B** is a data management platform for client specific user-data. This is a solution where publishers can build their own data sets, from the data that they import
and store on the platform. The data is stored in a closed environment allowing only the data owner access to it.

- **Product C** focuses on Big Data and contains over 300 million profiles. These profiles are used for advertising targeting and are all targetable on an individual level. The data is processed and analyzed, enabling large and scalable targeting activities towards target groups.

- **Product D** is their solution for real-time reporting and provides real-time executable data and customizable reports. It can also analyze the target group together with historical campaign execution to gain even more insights regarding customers’ interests and behavior.

**Participant B** started in 1996 and have now about 60 employees. They are located in Oslo, Norway. They are primarily a publicist but have several different products that they offer:

- **Product A**: a portal with shortcuts to the “rest” of the internet.

- **Product B**: an online newspaper.

- **Some smaller products**: for example the TV-guide

The key element for these products is that they deliver news, utility services such as Web Search in partnership with Google, and useful links. The portal is meant to “sort” the internet.

By aggregating, editing, guiding and distributing, they work to provide good and useful content.

The services that Participant B gives them a natural place on the Supply Side of advertising. By analyzing users and users’ behavior they are able to group and offer a targeted audience. They do not sell information about users but instead offer a targeted audience for a more effective ad campaign.

**Participant C** is a software-based company that offers cloud services to help companies comply with GDPR. They do this by developing and operating legal and privacy-enhancing information and communications technology and provide related services and products. The company started in 2015 and consists of 2 employees. Currently, they have three products that they offer.
- **Product A** helps businesses create GDPR-compliant Privacy Policies. This is done by answering questions that are linked to specific parts of the GDPR.

- **Product B** is their solutions for managing user consent. This allows businesses to easily store and retrieve user consents.

- **Product C** helps companies track and detect third parties running directly or indirectly on a website. This solution provides a unique analysis of third parties, help keep up to date agreements, and helps to be transparent to the end user.

**Participant D** is a creatively driven media agency located in Oslo. They produce music, movie, art, ads, and design. They started in 2011 and consist of 53 employees. As a media agency, they are on the demand side of advertising and can be seen as a content provider for advertisers.

**Participant E** is an interest and service organization for Norwegian advertisers. They organize about 170 members, primary the major marketers, such as DNB, Orkla, and Telenor. They offer a lot of courses and conferences, they conduct consultation statements and have network groups. They also do surveys which they use as a basis for presentations and articles that they publish. The turnover is about 13 million NOK, and they have 5 employees.

### 4.3.2 Discussions

The discussions were held with the project lead. He is also an external supervisor for this thesis.

Discussions were conducted as a type of unstructured interview. The discussions were an additional method of data collection to the semi-structured interviews. Several discussions were held with the project lead during the entire duration of the study. The choice of discussions, in the form of unstructured interviews, in addition to the semi-structured interviews were done because we in some cases had a general area of interest and concern and would let the conversation develop within this area. The discussions were held with the intended purpose of gaining understanding and testing theories. Some of these discussions were held with our main supervisor present. An overview of the discussions can be seen in Table 8.
<table>
<thead>
<tr>
<th>Date</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.10.2016</td>
<td>Karl Thomas, Jostein, Kjetil</td>
</tr>
<tr>
<td>27.02.2017</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>16.03.2017</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>06.04.2017</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>24.04.2017</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>08.05.2017</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>04.09.2017</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>22.11.2017</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>13.12.2017</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>16.01.2018</td>
<td>Karl Thomas, Jostein, Kjetil</td>
</tr>
<tr>
<td>19.02.2018</td>
<td>Karl Thomas, Jostein, Kjetil</td>
</tr>
<tr>
<td>15.03.2018</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
<tr>
<td>22.03.2018</td>
<td>Karl Thomas, Jostein, Kjetil, Petter</td>
</tr>
</tbody>
</table>

Table 8: Overview of the discussions

The output of these discussions was an increased understanding of DITIC and Digisphere. We gained insight into how the project progressed, the challenges that were met and how they were solved. In addition, we learned a lot about the challenges faced by a software provider in the online advertising industry and why DITIC was started.

### 4.3.3 Workshop

In addition to the interviews, we participated in a workshop between Startup Alpha and Startup Beta in Mars 2017. Our role in the workshop was as observers and consisted of listening and taking notes. The workshop consisted of technical and business people from both companies and the workshop lasted for a full day. The intentions were to map out the project plan, figure out what needed to be done, the technical requirements and explore possible challenges that could arise. Data were collected in the form of documents and presentations used during the workshop and summaries created by us.
4.3.4 Literature review

Multiple sources of literature have been used in this thesis. The literature review was conducted with two main purposes; (1) examining documentation in the form of evidence from the case study, and (2) explore existing knowledge on the online advertising industry. The first purpose resulted in the collection of several types of documents, ranging from powerpoint presentations to sketches of the platform. The second purpose resulted in the collection of background literature in the form of reports, journal articles, books and webpages. The types of literature used in this thesis are presented in Table 9.

<table>
<thead>
<tr>
<th>Literature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articles, reports,</td>
<td>Collected through search queries in Google Scholar. Examples of query strings is: “online advertising”, “programmatic advertising”. One article or</td>
</tr>
<tr>
<td>and books</td>
<td>report could often lead us to other relevant literature. This literature was created by other researchers.</td>
</tr>
<tr>
<td>Documents</td>
<td>The documents were sent to us after discussions or collected through a referral from the project lead. The documents were mostly created by</td>
</tr>
<tr>
<td></td>
<td>professionals in the online advertising industry in Norway. Examples of documents are application document for funding, annual reports of the</td>
</tr>
<tr>
<td></td>
<td>advertising market in Norway, regulations.</td>
</tr>
<tr>
<td>Pictures, sketches,</td>
<td>Pictures, sketches, and illustrations were used during explanations made in the interviews and discussions. These were collected afterward and</td>
</tr>
<tr>
<td>and illustrations</td>
<td>are made by participants in the interviews. Examples of what was collected are sketches of the relations in the online advertising ecosystem, how</td>
</tr>
<tr>
<td></td>
<td>parts of Digisphere functions and how existing technology in the industry works.</td>
</tr>
<tr>
<td>Presentations</td>
<td>Provided by the project lead, and interviewees. Sent to us after interviews to elaborate on the discussed topics. These are created by the</td>
</tr>
<tr>
<td></td>
<td>interviewees. Examples are marketing material, business proposals, project progress reports.</td>
</tr>
</tbody>
</table>

Table 9: An overview of the collected documentation

4.4 Data analysis

During this study, the theoretical lens was applied after the data was collected. This is as suggested by both Easterbrook et al. (2008) and Robson (2002) to avoid researcher bias and focus on data collection.
Data analysis was based on three flows of activity outlined by Robson (2002, p. 476): data reduction, data display, and conclusion drawing. These three flows create a continuous iterative process of data analysis. Data collection comes in addition to these activities. Our method for data analysis used the three activities in the following way:

- **Data reduction:** Our method of keeping the data manageable consisted of reducing the data collected from interviews to information that had value for the thesis. Workshops and discussions were summarized. Documents and literature were reduced by highlighting valuable data. An example from data reduction from the workshop can be found in Attachment C.

- **Data display.** The data was displayed and visualized through the use of matrices, charts, drawings, and diagrams. Some of the visualizations are presented in this thesis, other were used purely for data analysis. Examples can be found in Attachment D: Examples of data display.

- **Conclusion drawing.** The process of conclusion drawing was done through noting possible structures, mechanisms, and explanations. The process was accompanied by a verification process consisting of questioning conclusions and discussing their reliability. Examples can be found Attachment E: Examples of conclusion drawing.

Theory has a significant role in relation to our data analysis. Theory has been applied in the conclusion drawing phase of data analysis. This was done after the data collection was completed. We did, however, have knowledge about the theories before they were applied, which might have affected the analysis. Theory has been used in this thesis to understand the case under study and provide insight into how a competitive advantage for local software providers in the online advertising industry in Norway could be achieved. A lot of time has been spent relating the theoretical concepts to each other and examining the empirical data in relation to the theories one by one and together.

### 4.4.1 Creating the design guidelines

One example of the how the data analysis was done and its relation to theory and empirical findings is how the design guidelines presented in chapter 6.5 were created. The guidelines emerged towards the end of our research and the process started with extracting relevant parts
of the data collected. This was then analyzed to extract relevant parts for the design guidelines. This was also done with what we had concluded and discussed earlier in our research with relation to theory. We were then inspired by the design guidelines presented by Hanseth and Lyytinen (2010, p. 9) and decided to visualize the findings in the same format. This helped us structure and understand the data better. The guidelines were then explored in relation to our theoretical framework and our earlier conclusions. Possible mechanisms present in the design guidelines were discussed and explained in relation to theory before we concluded with what is presented in this thesis.

### 4.5 Validity

The validity of qualitative research has many threats. Validity is the trustworthiness of the research conducted (Robson, 2002) and in our case, this refers to the trustworthiness of our proposed design guidelines and research contribution. Reducing the threats to validity for this research is attempted through reducing bias and the degree that the researchers’ presence interferes with the result (known as reactivity). Robson (2002, p. 174) present six strategies for handling threats to validity in qualitative research by dividing the topic into reactivity, respondent bias, and researcher bias. With this Robson establishes a strategy for dealing with the threats to validity and steps one can take to increase the trustworthiness of the research. Our approach for dealing with threats to validity based on Robson’s strategies is presented below. Some of this approach has been conscious choices and other has been results of other factors along the way. Although the strategies presented in Table 10 help towards reducing the threat to validity it is not a way of guaranteeing validity.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Theory</th>
<th>Our approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prolonged involvement</strong></td>
<td>Prolonged involvement can reduce the threat to reactivity and respondent bias but increase the threat of researcher bias.</td>
<td>Our involvement in DITIC has been approximately one year and two months. This is a relatively long involvement and reduce the threat to respondent bias, but can increase our bias.</td>
</tr>
<tr>
<td>Approach</td>
<td>Description</td>
<td>Example</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Triangulation</td>
<td>Triangulation can reduce researcher bias, respondent bias and the threat of reactivity.</td>
<td>We reduce the threat to validity by using multiple sources of data collection (interviews, literary review, and discussions). In addition, we have been two observers through the entire study and apply more than one theoretical concept (digital innovation, platforms and value networks). We have not combined quantitative and qualitative approaches, which is a weakness.</td>
</tr>
<tr>
<td>Peer support</td>
<td>Peer groups can reduce the threat of researcher bias but have no effect on reactivity or respondent bias.</td>
<td>We reduce the threat to validity to some degree by having been two students through this study.</td>
</tr>
<tr>
<td>Member checking</td>
<td>Member checking can reduce researcher and respondents bias and the threat of reactivity.</td>
<td>We reduce the threat to validity by having returned understandings and interpretations we have made to participants in some circumstances. This has, however, not been done every time, which is a weakness.</td>
</tr>
<tr>
<td>Negative case analysis</td>
<td>Negative case analysis can reduce the threat to research bias but have no effect on respondent bias and reactivity.</td>
<td>A negative case analysis is not performed and is a weakness of our research.</td>
</tr>
<tr>
<td>Audit trail</td>
<td>An audit trail can reduce researcher bias but have no effect on reactivity or respondent bias.</td>
<td>We reduce the threat to validity by having kept a full record of all activities while carrying out this study. All interviews have been recorded and transcribed. Documents received has been stored.</td>
</tr>
</tbody>
</table>

Table 10: Our approach to threats to validity based on Robson (2002, p. 174)

**Generalizability**

The generalizability of a theory refers to its applicability outside of the setting it was developed, tested or confirmed (Lee & Baskerville, 2003). Using the framework presented by Lee and Baskerville (2003, pp. 232-238) we can position our research and its potential generalizability. Positioning our generalization as “Type ET Generalizability: Generalizing from Description to Theory” is fitting for our research because it can be used to generalize from case studies to theory and generalize of the resulting theory beyond the domain of the case study. These cases have been criticized and it is claimed that case studies have no generalizability beyond the given case. According to Lee and Baskerville (2003) researchers...
conducting a case study may strive to develop a theory that is generalizable within the case setting. Generalization beyond the case setting where it has been tested and confirmed can never be achieved. Incorporating this into our research we will strive to present a theory based on the case setting, keeping in mind that generalizability beyond this setting can never be achieved without testing and confirming in the new setting.

4.6 Summary

Our goal for the research was to explore and understand the challenges that the small actors face in the Norwegian online advertising market and viable solutions to these problems. This was done by answering the research question “How can digital innovation enable increased competitiveness for local software providers in the Norwegian online advertising market?”. Case study was chosen because case studies are an effective way to gain deep insight and knowledge about an object of study. The methods chosen for data collection are interviews, discussions, and literature review. We had some knowledge of theoretical concepts that would fit our project, but this was not applied until after the data collection was finished. Our data analysis is based on three flows of activity: data reduction, data display, and conclusion drawing. Theory has a significant role in relation to our data analysis and have been applied during conclusion drawing. A result of our data analysis is the design guidelines presented in chapter 6.5. Our approach for dealing with threats to validity is based on the six strategies presented by Robson (2002).
5 Case description

The project under study is called DITIC and is conducted by a Norwegian startup, Startup Alpha. The project is an attempt to innovate with the purpose of establishing a software service, called Digisphere, within the domain of programmatic online advertising in Norway. This chapter will start with describing the online advertising market in Norway. It will then describe each component of Digisphere, with a description for the various iterations and a concluding remark on the status. The last section will contain a brief description of the actors behind Digisphere.

5.1 Online advertising in Norway

From the interviews conducted we learned that programmatic advertising is a relatively new concept in the Norwegian market, compared to other markets such as Sweden or the American market. This also backed up by document analysis. In 2015, only 18% of the turnover in the advertising industry in Norway came from online advertising (Statistics Norway, 2018). This is also the case when it comes to programmatic advertising.

Online advertising in Norway has largely been dominated by direct buys of advertising space based on deals between the advertiser and publisher. However, in 2015 the real-time bidding (RTB) in Norway increased by 175% compared to 2014, largely due to publishers opting to use Open ad exchanges. In 2015 the supply side platforms in Norway increased their available inventory to handle over 12 billion bid requests each month (Delta Projects, 2015). This indicates that 2015 was one the first years with rapid growth in programmatic advertising in Norway.

5.1.1 Market status

Our impression from the interviews is that there is a high level of activity in the online advertising industry in Norway. Norwegian companies are in all parts of the industry and are competing against international actors. The local actors situated in either supply side or demand side has a stronger position than international actors. Supply side actors experience high traffic from local consumers and indicate that their main competition is Norwegian, not from international actors. The demand side actors do not experience to be out-bid on ad
impressions by international actors in any concerning degree and compete mostly with Norwegian actors.

Norwegian software providers experience heavy competition from international actors. Advertising software is not restricted to one geographical market but is applicable for multiple markets. This implies that there is no obvious reason for Norwegian companies to choose Norwegian software providers. Facebook and Google are often the preferred choices even in Norway, which presents a challenge for the local actors.

5.1.2 Google and Facebooks presence in Norway

Google and Facebook have a substantial market share in the Norwegian advertising market. This chapter will briefly describe their position.

**Google**

Google is present on 87 percent of all Norwegian newspapers (Datatilsynet, 2015) and is present in several of the roles in the advertising market. Google DoubleClick is a service which alone has products in DSP, SSP, and Ad Exchange. By being a worldwide service provider with products such as Gmail, Maps, Google+, they can get a steady stream of user data that they can use in the advertising products. From our interview we learned that Google has approximately 2/3 of the programmatic market share in Norway, taking more than 1 billion NOK of the market shares.

**Facebook**

Facebook has a strong position in the role as a data collector by being the world’s biggest social network. The social network has approximately 3,45 million Norwegian users, which accounts for approximately 84% of the Norwegian population over 18 years (IPSOS, 2017). They have a unique opportunity to gain insight into peoples’ interests through peoples’ actions and behaviors. Facebook is also an advertising company, by having deep insight into user’s life, they can offer pinpoint accuracy for their ads. By supplying a social platform through Facebook, Instagram, and Messenger, they take on the role as a supply side platform, allowing ads to be shown on the webpages and apps. Through Facebook Audience Network they take on the role as a DSP to serve ads across the web on other webpages.
Summary
With numbers like the ones presented above, it’s clear that Facebook and Google have a considerable influence on the online advertising market in Norway. Their position in the industry is remarkably strong, because of Google’s many services and Facebook’s numerous users.

5.1.3 Challenges in the Norwegian industry

Although online advertising has many positive sides, there are also challenges. Online advertising has even been described as “not trusted, not wanted, and not needed” (Hamilton, 2017). Through interviews, document collection and literary review we have gained an understanding of what we believe are the dominating challenges the Norwegian online advertising industry is facing at the moment.

Ad fraud
Ad fraud has been described several times during our interviews as the biggest problem within online advertising right now. Google’s ad network has been described to us during data collection as “the world’s biggest ad fraud”, meaning that the network suffers from fake ad clicks and several of the other parts of ad fraud.

The history of ad fraud began as a result of the introduction of cost per click (CPC) and the tracking of each ad click. The performance of different vendors within online advertisement was measured by how many ad clicks was generated on a specific advertisement. This became a standard key performance indicator (KPI) for many actors on the supply side and publishers would earn more money if more clicks were registered for the advertisements on their page. After a while, actors emerged that could generate clicks by paying people to spend their day clicking on advertisement – usually referred to as click farms. These actors would charge for this service and ad fraud (or false traffic) has been a problem since. With the introduction of automated bidding platforms and programmatic advertising, the ad fraud space has become more complex. There exist several ways of doing ad fraud (Delta Projects, 2016):

- Laundered impressions. This is when advertisers are fooled to think that they buy ad space on a legitimate webpage, but instead buy ad space on an unfavorable webpage. A typical example can be when an advertiser buys ad impressions in a newspaper, but the advertisement is really shown on a webpage for illegal file sharing. The advertiser
is then fooled to pay for an advertisement they would not buy if they knew where it actually was shown.

- **Ad stabling.** The problem of ad stabling is when several advertisements are placed on top of each other, making only the ad on top visible to visitors of the webpage. This fraud is based on selling ad impressions that are never displayed to the user.

- **Pixel stuffing.** The technology utilized by pixel stuffing is the same as the one used in pixel tracking. A pixel is placed on the webpage with a link referral to where the browser can find that pixel, and the browser will then load whatever is located at the link. In pixel stuffing the link points to other webpages, making it look like the user is visiting them, but the webpage is never visible to the user. This fraud makes advertisers buy ad impressions on a webpage that is never shown to the user, and the user does not see the advertisement.

- **Fake webpages.** These are webpages without human traffic, generated to sell ad impression to users that do not exist or never visit the webpage (for example through pixel stuffing).

- **Fake clicks.** This refers to when a person, bot or an automated script clicks on an advertisement with no other intent than clicking. In this fraud, advertisers pay for clicks that never will result in a purchase of their product.

- **Non-human traffic.** There are computer programs that simulate human behavior on a webpage and fool advertisers to buy ad impressions that are never shown to a human. These are harder to stop than pure fake clicks because it is harder to know if the visitor is an actual human being or not.

Ad fraud can quickly become costly for advertisers and fighting it is a focus for the industry.

**Lack of transparency**
Actors earning money from data collection, such as data brokers, analysis companies, and actors within online advertising in general, does not reveal much about their methods and users are generally not aware of their existence. During our interviews, we spoke to people indicating that a lack of transparency might be a problem that the industry faces, but the majority did not address transparency as a challenge. We believe that this might be because
our interview object is a part of the industry and therefore have a better understanding of the
practices, than people outside the industry.

Several sources back up this belief. The Commission (2014) published a report called “Data
Brokers: A Call for Transparency and Accountability”, which claim that users do not know
that these companies collect data and that their practices are very rarely questioned. Liu,
Sheth, Weinsberg, Chandrashekar, and Govindan (2013) have a similar view and present that
this is also the case for targeted advertising in general. Consumers notice that the ads change
based on their behavior but know very little of why this happens or what data are used and
stored about them. Ur, Leon, Cranor, Shay, and Wang (2012) mirror their view and claim that
people are concerned about this type of marketing. Addressing these concerns together with
the lack of awareness of data policies is a problem that can be argued to stem from the lack of
transparency in the online advertising industry.

There are some voices saying that this is now better than what has been seen before
(Hamilton, 2017), it seems that it still remains a problem. Busch (2016) even states that:

“Associations of National Advertisers around the world think that the world of online
advertising has reached a dead end where there is no transparency, a lack of measurability,
unviewable ad impressions and even fraud” (Busch, 2016).

Furthermore, Hamilton (2017) concludes his article “A New Take on Digital Advertising:
Theory, History, and Society” with:

“We need to continually subject the nature, processes, and operations of
digital advertising more fully to public scrutiny, assessment, and debate.”

This literature does, however, not relate directly to the situation in Norway. That being said,
most companies addressed in the literature have a presence in Norway and we believe that
there is not much data that indicates that this is not the case in our country. Datatilsynet
(2015) back this up by stating that there is a lack of transparency in the Norwegian industry as
well. Norwegian newspapers have also been reporting on several stories linked to personal
data and advertising, indicating that the practices of actors within online advertising is new
information to the Norwegian people. Examples of news articles are Johannessen (2018),
5.1.4 Data is a challenge for Norwegian software providers

Data is an enabler for competitiveness within programmatic advertising. From our data collection, we have learned that because user data adds value to ad impressions for both demand and supply side this is highly sought after for almost all actors in the online advertising industry. It has also become apparent that having less data than other actors in the industry is a weakness. This is especially true for actors in the industry that are not buying or selling ad impressions but have a role as software providers.

An actor holding a position as a software provider within programmatic advertising is particularly sensitive to having less data than competitors. This is because the products in this category often are used to take real-time decisions concerning ad impressions. An increased amount of data can enable one product to be superior to another. One example of this could be software products that decide what ad impressions to buy for demand side actors. If one product can offer more insight into the interests and purchase intents of individual users than a competitive product, it would be the preferred choice because it could increase the sales of a product without spending more on advertising. This advantage can be achieved by having more personal information about the user in question and data regarding the user’s previous online activities, such as previous purchases and browsing history, than competitors. Although, this is just one example of software it seems, from our investigation, that this holds true for most products offered by software providers within online advertising.

Gaining the amount of data required to compete is not possible for all actors. The actors dominating the industry have a huge advantage in their established online presence. Facebook has over 2 billion monthly users (Facebook.com, 2017) and harvest data through tracking pixels as well as users behavior on Facebook. Through their broad specter of products, Google is in an even stronger position to harvest data. By offering products such as a search engine, YouTube, Android, Maps, and Gmail they have an extraordinary position when it comes to collecting user data. The position the dominating actors have obtained as collectors of user data is hard to match for Norwegian actors. This is because it would be an extremely capital and time demanding effort, and even then, not feasible because of the head start Facebook and Google already has.

Another factor which makes data harvesting a challenge for Norwegian software providers is their lack of direct contact with the users. The supply and demand side are often the actors in
contact with the users and can collect data in that way. Software providers rarely interact with the users directly. They are able to collect data through other actors, but this cannot replace having direct user contact. The actors dominating the industry as software providers have many points of direct contact with the user through other products and can utilize this in their software.

Buying data to solve the challenges actors face in collecting data, is not possible in Norway. An alternative to collecting the data would be to acquire the data in some other way. Data can be bought through Data Brokers. However, the data offered for sale through Norwegian Data Brokers are not offering the same data that actors can collect on their own. The local Data Brokers aggregate public information and sell it but does not offer user data such as previous purchases, browser history or interests.

5.1.5 Summary

Programmatic advertising is a relatively new concept in the Norwegian market. The online advertising in Norway has largely been dominated by direct buys of advertising space based on deals between the advertiser and publisher. Lately, this trend has changed, and the industry experiences a rapid growth in programmatic advertising.

The local supply side and demand side actors in Norway do not struggle with competition from international actors. This is, however, not the case for local software providers which experience heavy competition from big international players, such as Facebook and Google. These actors have a remarkably strong position in Norway because of their widespread presence on Norwegian webpages and numerous users.

The online advertising industry in Norwegian face challenges related to ad fraud and transparency. For the local software providers, these challenges are trumped by their struggle to collect enough data to be able to compete against the international actors. Data is an enabler for competitiveness within programmatic advertising and having less data than competitors is a distinct disadvantage for software providers. Norwegian actors are not able to acquire the amount of data needed to compete with Facebook and Google.
5.2 Description of Digisphere

Digisphere consists of five distinct parts called Gate, Sprawl, Mind, Vault, and Bazaar. All parts have distinct roles in the solution. Digisphere Gate has a key role; it is a component the other parts rely on to be realized and is, therefore, the first step of Digisphere. In the following chapters each part of Digisphere is described.

5.2.1 Digisphere Gate

Digisphere Gate is a solution with two purposes; (1) an ID binding solution and (2) a solution for authenticating users. It is based on the idea that an ID is the smallest atomic unit in the context of user data on the internet and therefore needs protection. The solution solves problems related to the issues experienced when several actors need to communicate about the same user, without exchanging sensitive information.

The ID binding problem

Almost all parts of programmatic advertising rely on all actors in the chain being able to identify the user. The involved actors will usually have a unique identifier (UID) for each user in their internal storage solutions. This means that for the same user, all involved actors will have different internal UIDs. If one actor wishes to recognize that a user is someone they have stored data about at a previous point in time, with today’s solutions, the actor would deploy a technique called cookie matching (described in 2.1.7). This method, however, reveals every actor’s internal UID, which is not wanted by the actors in the industry. Cookie matching does also have problems with data leakage. The members of DITIC believe that there is a better way, which will be described in the next section.

The ID binding solution

The solution is a way of communicating about a user without using identifying information from any of the involved parties. This is done by creating ID pairs and storing them in a solution owned by a neutral third party without commercial interest in the data. ID pairs are two IDs that have a binding between them, making them linked together. The ID pairing is done at a point where both actors have identified the user and the user has approved the storage of data and the binding of the IDs. In most cases, this is when the user opens a webpage and cookies are loaded and the user is authenticated. At this point, every actor with
cookies stored in the user’s web browser and an agreement with the webpage owner to place a script (or tracking pixel) on the page to retrieve the cookies, can identify the user. This point is then chosen to bind two IDs, implying that these identify the same user.

The ID pair is then used for future communication between these two actors, regarding the specific user referenced by the IDs. Future communication will be based on random generated IDs, to avoid leaking sensitive data. The random generated IDs are created when one actor wishes to communicate with another actor about a user. Digisphere Gate generates a random identifier when it receives an ID and creates an ID pair of the two IDs (the received and the random generated). This random generated ID is then passed to other actors. The actors receiving the random generated ID will send it to Digisphere Gate and receive their internal ID in return. IDs can only be returned to the same actor that inserted the ID into Digisphere Gate. By doing it this way there is no risk of leaking sensitive data and by deleting the binding within an ID pair the solution allows for easy access management to personal data. A scenario explaining Digisphere Gate further is presented below:

<table>
<thead>
<tr>
<th>Step</th>
<th>Scenario: Cooperation between two actors – a new user is encountered.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>A web shop, ToyCars.com, uses Digisphere to authenticate users on their webpage. They also enter into an agreement with a company delivering online product recommendations, Recommendations24/7. These parties agree that Digisphere will be used to handle communication between them in the future.</td>
</tr>
<tr>
<td>1</td>
<td>The web shop, ToyCars.com, is visited by a new user (Kjell). Kjell has never visited ToyCars.com before and is prompted to authenticate himself by Digisphere. If Kjell had registered with Digisphere at an earlier point he would have been able to log in. In this case, Kjell has not used Digisphere before and registers himself with an email address and password, which is stored in Digisphere. Kjell is presented with information about which actors wish to identify him, the purpose they have by identifying him, and if data is going to be stored about him. Kjell is given an option to agree or disagree. Kjell chooses to agree and this is stored in Digisphere with the context of what Kjell agreed to.</td>
</tr>
<tr>
<td>2</td>
<td>Both ToyCars.com and Recommendations24/7 use cookies to identify the user. ToyCars.com identifies the user as AID1 and sends this ID to Digisphere Gate. Digisphere Gate creates a random ID, RID1, and binds the two IDs (AID1 and RID1). After the pairing is completed successfully Digisphere Gate returns the newly created random ID, RID1, back to ToyCars.com, indicating that the pairing was successful.</td>
</tr>
</tbody>
</table>
After ToyCars.com receives the random created ID, RID1, from Digisphere Gate, the ID is passed on as an argument when loading the script by Recommendations24/7. This allows Recommendations24/7 to access their cookie and the user is identified by Recommendations24/7 as BID1.

Recommendations24/7 received an ID from ToyCars.com when the script was loaded. This ID (RID1), together with the ID they identified the user as (BID1), is sent by Recommendations24/7 to Digisphere Gate with a request for pairing. Digisphere gate confirms that the pairing was successful.

Recommendations24/7 can now choose to generate a new random ID if they wish to communicate about the user (Kjell) to ToyCars.com or they can choose to store the ID received by ToyCars.com (RID1), since this just a random ID and not sensitive information for ToyCars.com. In this case, Recommendations24/7 do not wish to store anything else than their internal ID for the user (Kjell).

Recommendations24/7 does now wish to communicate to ToyCars.com about the user (Kjell). They send their internal ID (BID1) to Digisphere Gate. Digisphere Gate generates a random ID (RID2), binds it with BID1 and returns it to Recommendations24/7. This ID can now be used to communicate about the user (Kjell) with ToyCars.com

ToyCars.com receives information from Recommendations24/7 about the user (Kjell) with the ID RID2 but cannot identify the user based on this information. ToyCars.com passes the ID, RID2, to Digisphere Gate. Digisphere Gate knows that this ID (RID2) is paired with (BID1). BID1, however, does not belong to ToyCars.com and is therefore not returned. BID1 is paired with RID1, which in turn is paired with AID1. AID1 does belong to ToyCars.com and is therefore returned as a successful lookup.

Both ToyCars.com and Recommendations24/7 are now both able to identify the customer in question (Kjell), without exchanging any sensitive or personal data about the user. Both actors can only use data they have stored internally, avoiding data leaks or exchange of data without the user’s consent.

Table 11: Digisphere Gate example scenario

<table>
<thead>
<tr>
<th>ToyCars.com</th>
<th>Recommendations24/7</th>
<th>Random ID 1</th>
<th>Random ID 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>DGID1</td>
<td>AID1</td>
<td>BID1</td>
<td>RID1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RID2</td>
</tr>
</tbody>
</table>

Table 12: Example of stored information in Digisphere Gate
The same information can be visualized as pairs with the order they were paired in Digisphere Gate:

<table>
<thead>
<tr>
<th>Order</th>
<th>Pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AID1</td>
</tr>
<tr>
<td>2</td>
<td>RID1</td>
</tr>
<tr>
<td>3</td>
<td>BID1</td>
</tr>
</tbody>
</table>

Table 13: Example of information stored as pairs in Digisphere Gate

Based on Table 14 we can visualize how the lookup from the example above functions based on pairs. The actor received the random generated ID RID1 and was returned AID1.

<table>
<thead>
<tr>
<th>Order</th>
<th>Pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AID1</td>
</tr>
<tr>
<td>2</td>
<td>RID1</td>
</tr>
<tr>
<td>3</td>
<td>BID1</td>
</tr>
</tbody>
</table>

Table 14: Example visualization of lookup in Digisphere Gate

Digisphere Gate does not store additional information other than what is strictly necessary for the functionality it provides. This is a fundamental part of the idea and means that actors cannot store information about the user, other than an ID, in Digisphere Gate.

**Control over data access**

Through its ID binding solution, Digisphere Gate offers control over which actors can communicate about an individual user. By deleting ID bindings, the solution offers precise access control on an actor-by-actor basis. This can be done both by actors and users and will offer different use cases.

- **Binding deleted by an actor.** Actors are offered the possibility to delete bindings they have towards other actors. One actor will usually pass their internal ID to Digisphere
Gate and receive a randomly generated ID with a binding to their internal ID in return. This is described in Table 11: Digisphere Gate example scenario.

- If one actor wishes to stop another actor from looking up the random generated ID, the ID binding can be deleted. This will make it impossible for Digisphere Gate to know which actors the random generated ID belongs to and it can therefore not be used to gain knowledge about a user.

- **Binding deleted by the user.** Digisphere Gate gives the user the option to reserve themselves against being identified by individual actors. This is done by deleting one actor’s ID binding to the user in Digisphere Gate. By deleting the binding one makes sure that the actor cannot identify the user and can therefore not make use of the information received. This is because Digisphere Gate no longer has the actor’s internal ID for the specific user, making the lookup futile. When the actor no longer can lookup random generated IDs and get an internal ID in return the actor is excluded from the chain for that specific actor.

An example of a user deleting bindings in the context of Table 11 could be:

- If Kjell no longer wants Recommendations24/7 to gather information about his use of ToyCars.com, he can log on to Digisphere Gate and delete the binding that Digisphere Gate has between his user profile and Recommendations24/7. This is also updated in the agreement Kjell has with ToyCars.com, which means that Recommendations24/7 cannot create a new binding against Kjell’s profile in Digisphere.

**Security**

By only storing ID pairs, the solution has built-in protection against data loss. If some, or all, of the data should be lost to individuals or actors with malicious intents the information would be worthless without additional information about the involved actors and data from their internal storage solutions. To individuals or actors with no additional information, the data from Digisphere Gate would just be pairs of identifiers not feasible to extract any valuable data from. This is important because it reduces a user’s risk belief with Digisphere Gate and reduces the negative impact of a breach.
The authentication and consent problem
To be able to give the user control of which actors can gain information from their online behavior Digisphere Gate needs to provide a way to access control and confirm the user’s identity. Identification through cookies will not be sufficient because the inaccuracies with cookie identification could make it possible for other users of the same computer to alter another user’s preferences and would not make a coherent view for the same user across multiple devices.

In addition, the user should be provided with information about what Digisphere Gate does, what is stored, and how one can reserve themselves against some or all of these practices. This is one of the three pillars of IUIPC (described in 3.4), awareness of privacy policies. This is also an element of GDPR (described in 2.2.3), stating that the user has the right to know what data is being collected, who it is shared with, and why. This information must be presented to the user by the actor that the user engages with, in the case of online advertising this is usually the publisher. A standardized and natural way of presenting this information would make GDPR compliance and handling user privacy concerns easier for many publishers.

After being informed about privacy policies the user should, ideally, confirm that this is understood and should be given an option to not accept the policy (consent) based on the purpose of data collection. This gives the user an option to stop the collection of personal data. Consent is also a key element of GDPR.

The authentication and consent solution
The problem of access control, privacy policy awareness, and consent can all be solved by creating a login solution as a part of Digisphere Gate. A log in solution would present the user with information regarding privacy policies and give the user the option to either accept the policy by signing in or not accept by not signing in.

The log in solution from Digisphere Gate would be present on the publisher’s webpage but would be provided by Digisphere Gate. The privacy policy presented to the user is not provided by Digisphere Gate but addresses the fact that Digisphere Gate is used and that users can control the usage of their personal by logging on to Digisphere Gate. A mock-up of how this could be presented to the user is shown below in Figure 6.
The example mock-up in Figure 6 shows how a login solution can act as an active consent. Digisphere Gate would be able to provide details about the actors gaining access to information about the user, making it possible to integrate these details into the text.

Authentication with Digisphere Gate would also provide the advantage of Unique ID (described in 2.1.6). This means that the user can be identified across various devices and is not affected by solutions for blocking advertisement. Unique ID requires the user to log in and is, therefore, a natural tracking and identifying method for Digisphere Gate, which is dependent on a login solution for other purposes as well.

**Similarities and differences with cookie matching**

Digisphere Gate bares many resemblances to cookie matching (described in 2.1.7). They solve the same problem and do it in a comparable way. It could be natural to view Digisphere Gate and various providers of cookie matching technology as direct competitors. There are, however, some key differences as shown below in Table 15:
<table>
<thead>
<tr>
<th>Attribute</th>
<th>Cookie matching</th>
<th>Digisphere Gate</th>
</tr>
</thead>
<tbody>
<tr>
<td>User identification</td>
<td>Cookies</td>
<td>Unique ID</td>
</tr>
<tr>
<td>Allows user insight and</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>modification?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Requires consent?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ownership</td>
<td>Every actor owns their own cookie matching table</td>
<td>Owned by a non-profit organization</td>
</tr>
</tbody>
</table>

Table 15: Comparing cookie matching and Digisphere Gate

5.2.2 Digisphere Sprawl and Mind

The product of Digisphere Sprawl and Digisphere Mind results in a solution where actors can store and share data about users. Since Digisphere Gate does not allow for anything else than ID bindings to be stored, other solutions are needed to store additional data about the user.

Mind
Digisphere Mind is the name given to a small data set belonging to one specific user. The data set is “locked” and cannot be “unlocked” or used unless a set of requirements is fulfilled. The criteria for unlocking a specific Mind are based on consent from the user the data relates to and the actor that stored it. This means that the actor that stored the data in the first place can unlock the Mind at any time. This is because the actor had to collect consent from the user when the data was collected.

The data set is kept small on purpose to avoid storing sensitive information related to one user together. One Mind consists only of an ID, a description of the data, and the value. The IDs are stored in Digisphere Gate and are used to make it possible to refer to the Mind at a later point. One example of this data structure could be:

<table>
<thead>
<tr>
<th>ID</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID1</td>
<td>Email</td>
<td><a href="mailto:example.email@gmail.com">example.email@gmail.com</a></td>
</tr>
</tbody>
</table>

Table 16: Example of Digisphere Mind for email

Sprawl
Digisphere Sprawl is used to keep control of the Minds that belong to the same user. Since each Mind consists of the bare minimum of data, the value is extracted through Sprawl.

Sprawl has the ability to combine the information of several Minds into a user profile.
will give actors the ability to look up and find information related to one user, read the data and store new Minds. What data each actor can look up and read is based on whether the actor in question fulfills the criteria for each specific Mind. Actors fulfilling the criteria will be able to look up and read data stored by other actors.

**Sharing user data between actors**
Allowing actors to access data stored by others lays the foundation for an ecosystem and sharing of user data. The use case for this functionality is that actors can benefit from data collected by other actors using Digisphere, as long as it is approved by the actor that collected the data and the user the data belongs to. In a scenario where one actor needs the user’s email, this can be solved without asking the user to provide this information. This functionality means that every actor that joins Digisphere will benefit from actors already enrolled.

**Control and ownership over user data**
Digisphere Mind and Sprawl gives the user full ownership and control over the data. This implies that all storage and access to a user’s data needs to be anchored in the consent from the user in question. If the user should at any point wish to withdraw consent, view information stored about him/her, edit information, transfer information, or delete it this is made continuously available. Digisphere as a whole will offer a service for a user wishing to make alterations or gain insight into what information is stored about them, this can be done through a gateway called Digisphere Portal. This is made possible by the authentication solution in Digisphere Gate. The gateway will combine information from Digisphere Gate, Mind, Sprawl, and Vault to present the user with all information stored about them. In addition to presenting the user with information, the gateway will present the user with the following options:

- Block specific actors from communicating about the user. This is done by deleting the binding the actor has towards the user’s profile in Digisphere Gate, stopping the actor from accessing information and looking up their internal ID for the user in Digisphere Gate.

- Withdraw consent. The user is given an option to withdraw active consents. This stops actors from collecting or storing additional information about the user in the future but does not delete the collected data.
• Edit stored information about the user. If the user notices that some of the stored information is wrong, or has changed since the data was collected, the user is given an option to correct the information.

• Export the information. The user is given an option to transfer the collected data to another solution by exporting it from Digisphere.

• Delete stored information about the user. If the user no longer wishes to have information about them stored in Digisphere, it is given an option to delete some or all the information.

5.2.3 Digisphere Vault

Digisphere Vault will store the contracts used to control access to Digisphere Gate, Mind and Sprawl. Contracts will store which actors can access specific files, based on consent from the collector and the data owner (user). Digisphere Vault will be used to validate permission by looking up existing contracts and return if the access is valid or not. This will be applicable to both Digisphere Mind/Sprawl and Digisphere Gate. Vault will allow for finding contracts, reading contracts, storing contracts and performing the validations.

5.2.4 Digisphere Bazaar

Digisphere Bazaar is a marketplace that allows actors to purchase views for their commercial goods on other actors’ webpages. Bazaar is a product for actors within e-commerce and creates opportunities for sales between complementary actors. Actors can offer space on their e-commerce webpage for sale based on the individual user that visits the webpage, in a comparable way to how publishers sell their ad impressions in programmatic advertising. Complementary actors to the one selling space can buy the spot for one specific user and offer their own products for sale. The process could look like this for an online bookshop and an online hardware store:
<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>An online bookstore named booksonline.com registers with Digisphere Bazaar.</td>
</tr>
<tr>
<td>1</td>
<td>A user (Lise) visits booksonline.com to purchase the book “Woodworking for dummies”.</td>
</tr>
<tr>
<td>2</td>
<td>After she has purchased the book, a notification about a potential impression is sent to Bazaar along with the information about the purchase and the user.</td>
</tr>
<tr>
<td>3</td>
<td>E-commerce actor Hardwaresupplies.com calculates how much they are willing to pay for this impression. They find that this person is probably in the market for hammers and nails and therefore places a high bid – winning the auction.</td>
</tr>
<tr>
<td>4</td>
<td>Lise is shown the confirmation page and Hardwaresupplies.com presents products on the same page.</td>
</tr>
</tbody>
</table>

Table 17: Digisphere Bazaar example

If there is more than one e-commerce actor interested in presenting products to the user, the winner is decided through bids. By allowing more place for product placement there can be competing actors displaying products next to each other, or one actor could display more products. This would depend on what price each actor would be willing to pay for the product impression. A mock-up of how the confirmation page might look after the example in Table 17 and one additional actor is shown in Figure 7 below.

![Figure 7: Mock-up confirmation page with Digisphere Bazaar](image-url)
**Product impressions**
The item for sale in Digisphere Bazaar is called product impressions. When users have bought a product from one e-commerce actor, they will be shown complementary products from other e-commerce actors. The term product impression describes the moment a user is shown complementary products.

The product impressions must comply to certain standards set by Bazaar and will be presented to the user in a similar fashion. This results in a focus on the product and avoids attention-grabbing text, animations or other unnecessary attributes to the product impression.

**The marketplace**
Digisphere Bazaar is a two-sided marketplace where product impressions are for sale. The two sides of the marketplace are the demand side and supply side. Supply side is the role where the actor offers product impressions for sale. Demand side is defined as the actors wishing to buy product impressions.

Each actor in Digisphere Bazaar could possess a role in both sides of the marketplace. Because the actors are within e-commerce they can provide product impressions for sale on their confirmation page and have products they want to sell through product impressions displayed with other actors.

The relationships between the actors will be complex and mostly governed by the products each actor offers. There are many potential scenarios where Digisphere Bazaar could increase revenue for both supply- and demand side. Most cases in Bazaar work both ways, as the products complement each other. This is made explicit in the first example in Table 18 but is also true for the rest. Some examples of scenarios in Digisphere Bazaar:

<table>
<thead>
<tr>
<th>Supply side actor</th>
<th>Supply side product</th>
<th>Demand side actor</th>
<th>Demand side product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bookstore</td>
<td>The book “Woodworking for dummies”</td>
<td>Hardware store</td>
<td>Hammer and nails</td>
</tr>
<tr>
<td>Hardware store</td>
<td>Hammer and nails</td>
<td>Bookstore</td>
<td>The book “Woodworking for dummies”</td>
</tr>
<tr>
<td>Hardware store</td>
<td>Hammer and nails</td>
<td>Wood supply shop</td>
<td>Wood materials</td>
</tr>
</tbody>
</table>


Manchester United fan store | Manchester United football jersey | Football game ticket provider | Tickets to Manchester United’s next football match
---|---|---|---
Tea store | Herbal tea | Home décor shop | Cozy blanket
Men’s suit shop | A suit | Shoe store | Formal dress shoes
Table 18: Examples of cases for Digisphere Bazaar

In some scenarios, one could imagine that the store selling the supply side product, also can provide the demand side product. One example of this could be the example in Table 18, as most places where one can buy suits, one can also buy shoes. Offering another actor the opportunity to sell these products may damage the supply side actor’s sales. To avoid these cases each actor can provide a list of actors or products they do not allow on their pages (also called blacklisting). This can also be valuable to avoid products or actors that may not be suitable for the users of the supply side actor’s webpages.

Digisphere Bazaar will be a private marketplace in Norway, not open to the public. The initial phases will only include a selected number of actors from the Norwegian e-commerce space. After the solution gains traction and other actors wishes to join there will be a process to accept new members into the marketplace.

**Similarities and differences with programmatic advertising**

Digisphere Bazaar is highly comparable to programmatic advertising. However, programmatic advertising is more a concept than a product or solution. Programmatic advertising consists of DSP, SSP, and ad exchanges. Within each of these, there are several different solutions and vendors. Digisphere Bazaar, on the other hand, is one solution where the different functions of programmatic advertising are included. Bazaar does not offer choices with regards to software providers, the solution is one package where everything is included. This makes the two hard to compare directly. Bazaar could even be argued to be a programmatic advertising solution. While this might be correct, there is an insight to be gained in comparing them. The comparison could provide knowledge about what makes Digisphere Bazaar something potentially innovative and different from programmatic advertising solutions in general.

The main difference between programmatic advertising and Digisphere Bazaar is the focus on product listings in Bazaar, instead of advertisements. It could be argued that this actually is
the same thing because showing the user a product could be regarded as an advertisement. However, Digisphere Bazaar has standards that force the product impressions to look like each other, instead of allowing the actors to freely create their own graphic impression.

There are other similarities and differences between programmatic advertising and Digisphere Bazaar which are summarized below:

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Digisphere Bazaar</th>
<th>Programmatic advertising</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>Product listings</td>
<td>Advertisement</td>
</tr>
<tr>
<td>Type</td>
<td>Product</td>
<td>Concept</td>
</tr>
<tr>
<td>Real-time bidding on impressions</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Key data driver for impression bids</td>
<td>Data about the user’s purchase, i.e. the product</td>
<td>User data</td>
</tr>
<tr>
<td>Importance of user identification</td>
<td>Increases value, but not crucial</td>
<td>Crucial</td>
</tr>
<tr>
<td>Placement of impressions</td>
<td>Confirmation page</td>
<td>No regulations</td>
</tr>
<tr>
<td>Graphic design of impressions</td>
<td>Highly restricted. All created to be similar</td>
<td>Unrestricted</td>
</tr>
<tr>
<td>Actors obtain both roles in the marketplace</td>
<td>Fairly often</td>
<td>Infrequently</td>
</tr>
<tr>
<td>Private marketplace</td>
<td>Yes</td>
<td>Can be both public and private</td>
</tr>
</tbody>
</table>

Table 19: Comparing Digisphere Bazaar and Programmatic Advertising

5.2.5 The solution as a whole

There are attributes that apply to all, or most, parts of Digisphere, these are described here.

Ownership of the solution

Even though Digisphere was conceptualized and brought forward by a commercial actor, Digisphere will be owned by a non-profit organization created with the sole purpose of being a neutral, transparent owner of the final solution without a commercial interest in the data.
Core products
Digisphere Gate, Mind, Sprawl, and Vault make up Digisphere Core. The products in the core are all created to solve problems related to programmatic advertising and are tightly linked by this mission. The indispensable product in Digisphere can be argued to be Gate. This is because it solves a concrete problem for the creators and other components in the core exist as complementary products, increasing the value of Gate.

Publicly available core through API
The components of the Digisphere Core will be made publicly accessible through application programming interfaces (APIs). Everyone that wishes to utilize the components of the Digisphere Core is free to do so without charge. The public APIs to the Digisphere Core is a fundamental idea of Digisphere. This also means that everyone is free to create solutions that utilize the APIs and can use the Digisphere Core to create commercial applications in the interest of actors not associated with Digisphere. Digisphere Bazaar will utilize the public API in the same way as third-party applications. The motivation behind this choice of public APIs is to increase the potential value by maximizing the attraction the solution has towards new potential actors/customers. This approach will be discussed further in chapter 6 Discussion. Figure 8 visualizes how Digisphere Core and the API relates to third-party actors.

![Figure 8: Illustration of the public API of Digisphere Core](image)
Commercialization
With Digisphere Core being open and free to use, Digisphere needs other ways of financing the costs associated with hosting and maintaining the solution. This is where Bazaar comes in as a commercial product. Digisphere Bazaar is separated from the free and open core and connects through the public API. This means that Bazaar can utilize Digisphere Core in the same way as commercial third-party products, giving them the same foundation for competition. Digisphere Bazaar aims to make a profit through its marketplace.

Digisphere also sees an opportunity to create other commercial products which utilize the public API of Digisphere Core. It is possible for the actors involved in Digisphere to leverage their strengths as a company and create other commercial products. One example of this could be a way of commercializing the data in Digisphere Mind & Sprawl by creating a marketplace for user data. Another example is for Startup Alpha to provide a solution for delivering product recommendations for usage together with Bazaar. This, however, remains as a thought at this point and is not conceptualized.

Stacking components and value
The components of Digisphere depend on each other and increase each other’s value. The components have defined interfaces between them and can be viewed as loosely coupled and layered. Dependencies between the components are not necessarily equal between the component, but it can be argued that each component positively affects each other’s value.

The components of Digisphere can be regarded as stackable, with Digisphere Gate as the first element. The rest of the components of Digisphere depend on and use Gate in some way or another. Since Gate provide communication between actors, this is how it is used by the rest of the stack. It is a fundamental component but does not provide considerable value in itself.

The next component in the stack is Digisphere Mind & Sprawl. These components are regarded as one in this case because they serve a purpose together. Mind & Sprawl serve the purpose of storing user data. They depend heavily on Digisphere Gate for storing and retrieving IDs. Digisphere Mind & Sprawl add additional value on top of Gate.

Digisphere Vault is the next addition to the Digisphere stack with a clear dependency to the component stacked below it, Digisphere Mind & Sprawl. Without the components below it in the stack, Vault would not have anything to govern control over and its value would not exist.
Vault adds additional value to Mind & Sprawl by allowing actors to share collected user data, through the use of contracts. Vault also adds value to Digisphere Gate by adding increased user control and by storing user consent.

Digisphere Bazaar represents the last element of the Digisphere Stack with its dependencies to all Digisphere’s core components. The dependencies Bazaar has towards the rest of the Digisphere stack is not as explicit as the other relationships in the stack because it is outside the core products and connects through the public API. Bazaar is in a unique position in Digisphere because it adds a lot of functionality, complexity, and value in itself. The solution utilizes the rest of the stack but does have functionality which does build directly on Vault, Mind & Sprawl or Gate. Instead, these components provide functions Bazaar relies on. Digisphere Bazaar does provide additional value to the rest of the stack by creating a strong value proposition for Digisphere as a whole.

Figure 9 illustrates how the Digisphere stack depend on each other:
5.2.6 Summary

Digisphere consists of several components which serve different purposes, which is summarized in Table 20. Digisphere Gate is a key component in the solution because other parts build on its functionality.

<table>
<thead>
<tr>
<th>Digisphere component</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate</td>
<td>Gate is a solution with two purposes; (1) an ID binding solution and (2) a solution for authenticating users.</td>
</tr>
<tr>
<td>Sprawl and Mind</td>
<td>Sprawl and Mind results is a solution where actors can store and share data about users.</td>
</tr>
<tr>
<td>Vault</td>
<td>Vault will store the contracts used to control access to Digisphere Gate, Mind and Sprawl.</td>
</tr>
<tr>
<td>Bazaar</td>
<td>Bazaar is a marketplace that allows actors to purchase views for their commercial goods on other actors’ webpages.</td>
</tr>
</tbody>
</table>

Table 20: Summary of the purposes of the components of Digisphere.

The Digisphere components Gate, Sprawl and Mind, and Vault make up Digisphere Core. Digisphere Core is available through a public API. Digisphere Bazaar will connect to this API and is complementing solution to the Digisphere Core. The components of Digisphere have defined interfaces between them and can be viewed as loosely coupled and layered.

5.3 The evolution of Digisphere

Digisphere, presented in chapter 5.2, has evolved a lot over time and has been through many stages before it resulted in software service described above. The various stages will be described here.

5.3.1 Iterations of the core

There were three major iterations of Digisphere Core, they will be briefly described under. The new additions(s) will be described with a short explanation for why there was a need for a new iteration. An illustrative overview of Digisphere’s components are shown in Figure 10: The current architecture of Digisphere.
Initial idea
Digisphere started by being a solution that tracked an ID through multiple services in the programmatic network. In addition to tracking a user, it was supposed to be a “safe haven” for personal information. Instead, this was not the case, as the project lead stated in our initial interview:

“Nobody cares about the user, not even the user themselves... When I started with this I thought it was something everybody would obviously want... The amount of people who cares about this is so marginal; practically nobody cares.

Iteration 1
Even though “nobody” cared about personal information, the GDPR was being enforced by the EU in May 2018, which in turn forces companies to care, by law. The idea was then that Digisphere could increase the attractiveness to other actors by being GDPR compliant, instead of being a “safe haven”. Incorporating GDPR in Digisphere required some new technical features in Digisphere. The new system was called Gate and has two functions; an ID binding solution and a solution for authenticating users. This would allow a user to log in and see what information is stored and remove unwanted bindings.

Iteration 2
Gate as a solution was not sufficient to handle the GDPR, thus some new features were born. Digisphere Sprawl and Mind had the job to store user data because only the ID binding was stored in Gate.

Iteration 3
A problem with Gate, Mind, and Sprawl was that it couldn’t store user consent. The solution was therefore to add a system called Vault. Vault’s function was to store the contracts, and contracts will store which actors can access specific files, based on consent from the collector and the data owner, which is the user.

Bazaar
Bazaar was introduced as a way for actors to profit from Digisphere and was therefore not a part of the iterations that the Core went through. It can be seen in the figure below positioned in the outer layer as a complementary software.
Status
Digisphere is still in the early phases as a product. A minimum viable product (MVP) was planned to launch in October 2017, but without the funding to hire software developers, this goal was not possible to reach. The MVP is still currently under development, but the progress is slow because of the tight budget. The Digisphere concept has not been tested in the market, and it is unknown how Digisphere will be received.

5.3.2 Workshop
The workshop was a milestone in Digisphere’s progress. The outcome of the workshop decided the path for Digisphere. The workshop was held between iteration 1 and 2. Technical requirements and goals were discussed. Our role in the workshop is described in chapter 4.3.3. The outcome of the workshop and its agenda is presented here.

Outcome
Digisphere was regarded to have significant potential for both parties but was considered as an ambitious project. The goal was to build a minimum viable product which consisted of one platform. Scalability and standardization were deemed to be two important factors for long-term success. Since Startup Beta doesn’t support cookies another solution had to be found, this was a big blocker and had to be researched and clarified before Digisphere could
advance. Some high-level system requirements were also mapped out, for example, that the system needs to manage id, delete cookies, use real ID, and have a latency of 300 milliseconds. How these requirements were to be implemented were not discussed.

The workshop concluded that Digisphere will be developed as a free open core with an API. A commercial layer will be built around this core, enabling different actors to commercialize the core. The system will be built on privacy rights for leveraging the value of trust. The trust of the consumer is considered to be vital to the success of Digisphere. Digisphere Gate will have an ecosystem, built on using temp-ID that is shared among the users. It will be scalable and ready for global use.

5.3.3 Summary

Digisphere has evolved a lot since the beginning of our research. The core components of Digisphere has been through three iterations of the initial idea and Digisphere Bazaar has been introduced. Digisphere is still in its early phases and the progress is slow due to a lack of funding. The workshop held by DITIC between iteration 1 and 2 decided Digisphere path and was a milestone for both DITIC and Digisphere.

5.4 Involved actors

DITIC is driven forward by two small Norwegian companies. This chapter describes them, their role and their motivation for participating in the project.

5.4.1 Startup Alpha

Startup Alpha is a startup company established as a software provider within online advertising. They deliver services in the domain of real-time analytics, recommendations, and omnichannel personalization. The company consist of eight people and was started in 2013. The company’s main product is a state of the art Recommender which they deliver to several Norwegian companies. The Recommender delivers recommendations for online shops on what customer will most probably purchase by each product in their inventory based on several data points. This allows the companies to use Startup Alpha to decide what products to show each individual user to maximize the number of visitors that buy something from
their webshop (conversion rate). This process happens in real-time and Startup Alpha is heavily involved in programmatic advertising to make their services possible.

Startup Alpha has a role as lead actor in DITIC. The project originated in Startup Alpha and the individual leading the project is a Startup Alpha employee.

**Motivation**

Startup Alpha is motivated by several attributes of Digisphere. The original motivation comes by the opportunities Digisphere gives for solving concrete problems they have faced. This was the driving force behind Digisphere Gate and the reason for its existence; to solve challenges related to the exchange of IDs within programmatic advertising.

Another motivator was the need for solutions related to information privacy. Startup Alpha noticed that the solution for ID exchange could provide benefits in relation to information privacy.

The motivation behind creating additional products in addition to Digisphere Gate were both political and in Startup Alpha’s best interest. Additional products could be a tool to make the industry easier for smaller actors, such as Startup Alpha. There was also a possibility that the additional product gave an opportunity for Startup Alpha to sell services they already provide and increasing their customer group.

**5.4.2 Startup Beta**

Startup Beta is a company established as a software provider within online advertising. They are a privately-owned company, they have their headquarters in Fredrikstad, Norway, and their sales offices are in Oslo, Norway, and Copenhagen, Denmark. The developer branch is located in Tallinn, Estonia. The company started in 2013. They have now approximately 16 employees.

Startup Beta is a company that focuses on brand and tactical advertising. They offer a creative management platform that is built for marketers within programmatic advertising. The platform provides services for automating creative production for cutting-edge digital campaigns. This is done by enabling brands and advertisers to reach targeted, brand-receptive audiences across a wide range of digital platforms and channels.
The role of Startup Beta in the Digisphere project became prominent after initial discussion with Startup Alpha and the kickoff marked the beginning of a collaboration between the actors. After this Startup Beta had a role as a partner in DITIC.

**Motivation**

Startup Alpha’s motivation is built up by two different ideas. The initial motivation was that DITIC could be a possible tool against actors such as Facebook, Apple, and Amazon. Startup Beta feels that these giants have built up massive amounts of data that is a huge competitive advantage in every possible way, such as funding through advertising and creating good services that people want to use. If it’s not only the giant corporations that have this advantage, but also users of the Digisphere platform, the competition will be more even for actors entering the digital race later. Startup Beta is therefore motivated by the idea that Digisphere Gate, and the technology that Gate is built upon, could be a political tool to even out the playing field.

The project could also help make better standards in the industry, which is another motivator for Startup Beta. They want users to know what their data is used for and how. This can be done by increasing the transparency in the advertising market and implementing a “privacy by design” approach, which is something Startup Beta is passionate about.

**5.4.3 Summary**

DITIC is driven forward by two small startup companies called Startup Alpha and Startup Beta. Both companies are established as software providers within programmatic advertising in the online advertising industry in Norway. The project is led by Startup Alpha with Startup Beta as a partner.

Startup Alpha is motivated to create Digisphere because it could solve concrete problems they have faced and solve problems related to information privacy. The big driving force, however, is the possibility of making the online advertising industry in Norway easier for local software providers.

Startup Beta is motivated to create Digisphere because it could even out the playing field for the actors in online advertising in Norway and because it could establish better standards in the industry related to information privacy.
6 Discussion

This chapter will discuss and answer the research question in the light of the chapters in background, theory and case description:

**RQ:** How can digital innovation enable increased competitiveness for local software providers in the Norwegian online advertising market?

We will discuss, in section 6.1 and 6.2, why and how data is required to enable competition and how Digisphere Core can acquire this data. The discussion will then focus on how to increase the amount of data by increasing Digisphere’s attraction. This will be done in chapter 6.3 by discussing the value creation framework in relation to Digisphere. In chapter 6.4 success factors directly related to the case will be presented and discussed. In 6.5, we will extract and generalize our findings and propose a set of design guidelines to maximize the attraction for a software service. Last in 6.6 the limitations to the thesis will be discussed.

6.1 Data is an enabler for competition

To answer how digital innovation can enable increased competitiveness we need to understand what is required for competing as a software provider in the online advertising market. As described in chapter 5.1.3, Norwegian software providers face challenges related to competitiveness because of an insufficient amount of data. This means that creating a solution that would give local actors access to a competitive amount of data might enable them to compete for market shares on other software attributes.

From our interviews, we learned that even for the individual actor it will be hard to quantify the competitive amount. However, it is clear that the competitive amount is more than the local actors have available for them today. The amount will vary from actor to actor, dependent on the software they develop and their segment of the market. In general, we define the amount as enough data such that an actor’s product does not become uncompetitive based solely on a lack of data.

By increasing the amount of data, we refer to an increase in both dimensions of user data. Meaning that an increase will refer to the growing the number of users that there are data about and the amount of data about each user.
Providing local software providers with a competitive amount of data would create opportunities for their products to compete on other attributes. Local software providers would, therefore, have to produce software that is superior in attributes such as technology, price, or user-friendliness. One example of this from our data collection is Startup Alpha and their product recommendation software. By having an equal data foundation to their competitors, Startup Alpha could compete on price or a superior algorithm.

Digisphere Core could have the potential to enable increased competitiveness for local software providers by creating a way for the actors to share and gain access to an increased amount of data. The next chapter will discuss how this can be achieved.

6.2 How to acquire a competitive amount of data

Digisphere Core enables sharing of data between multiple actors through its components Sprawl and Mind, as described in chapter 5.2.2. This functionality means that increasing the number of actors in Digisphere Core will potentially increase the combined amount of data that is shared between the actors. Even though Digisphere Core has the functionality, it is not sufficient to achieve a competitive amount of data. For this amount to be acquired Digisphere will have to attract numerous actors motivated to share their collected data. Figure 11 shows how actors can collaborate to generate a competitive amount of data compared with Google and Facebook. The amount of data in the figure are not based on real data and are only meant to be used as visualization. The colors are used to distinguish the actors.
This discussion is based on a hypothesis that increasing the number of actors using Digisphere Core will increase the amount of data available. This may, however, not be correct. There are several reasons for why this hypothesis might not hold. This will be discussed further in chapter 6.5 Limitations.

6.3 Creating attraction for Digisphere Core

Increasing the amount of data available in Digisphere Core could be achieved by establishing Digisphere Core as an attractive product and something actors in the online advertising industry want to be a part of. How Digisphere could do this will be discussed in this chapter by relating it to our theoretical framework of value creation, presented in chapter 3.

Our value creation framework will in this chapter be used to discuss attraction for Digisphere Core. This is based on the idea that software services that can create value for actors using it will be attractive to be a part of – in other words, that value creation is attractive. We will,
therefore, use theory relating to how value is created in value networks, software-based platforms, and digital innovation to discuss attraction towards Digisphere Core.

6.3.1 Leveraging network effects to increase attraction

The effects of a network could be leveraged by Digisphere to increase its attraction. Together the Digisphere Core components could create the foundation for a value network (described in chapter 3.1 Value network) where the actors benefit from each other. Potentially, every actor that starts using Digisphere Core will benefit from the data of the actors already enrolled. The potential network effects present in Digisphere can be divided into direct and indirect effects. The discussion below is structured with this in mind.

**Direct network effect to increase the amount of data**

The direct effect of an increased number of actors in a network is an increased value for the actors present in the network (Katz & Shapiro, 1985). This can be argued to be present in Digisphere Core. When new actors enroll in Digisphere they bring their data and increase the total amount of data available for actors that are already a part of Digisphere. By enrolling local actors, the network could consist of relevant user data not accessible anywhere else. Actors joining the network could also benefit from the data already added to Digisphere. More data could increase the potential value of Digisphere because there are more to extract value from. This effect could have the potential to increase Digisphere’s attraction for new actors and in turn increase the available amount of data.

From the theory of value networks (described in chapter 3.1) we learn that value networks create value in its facilitation of information exchange between actors and that the value for each actor will vary depending on the other actors in the network. Relating this to Digisphere can give us further insight into how the increase of data might happen. In the context of data, some actors might benefit more from the exchange than others. This might be actors that have less data than the industry average or have a need for a special type of data. The actors already possessing a competitive amount of data may not be attracted by the network structure because it has less value to them. This might result in a network where the demand is higher than the supply of data. To avoid this there might have to be other attractive attributes with the network (discussed further in chapter 6.4).
Value network theory can also provide insight into how the design of the network might affect its value (Stabell & Fjeldstad, 1998). The design of Digisphere Core facilitates relationships where one actor can utilize the data collected by another actor. This, however, must be agreed to by the actor collecting the data and the data subject. Decisions regarding how this will be solved can affect the relations in the network and therefore affect its value. Another similar major decision is whether the facilitating firm will have a moderating role and remove unserious actors and data that is below a certain quality. These decisions will have a potential impact on the amount of data available in Digisphere but will not be discussed further because they are considered outside the scope of this thesis.

Finally, we can investigate Digisphere Core in relation to the three core building blocks of a value network (Kothandaraman & Wilson, 2001) to summarize the potential of direct network effects in Digisphere Core discussed in this chapter. This can be found in Table 21.

<table>
<thead>
<tr>
<th>Building block</th>
<th>Digisphere Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior customer value</td>
<td>Digisphere Core aims to provide superior customer value by giving customers access to a competitive amount of user data.</td>
</tr>
<tr>
<td>Core capabilities</td>
<td>Digisphere Core will deliver superior value by facilitating a network where data collected by each actor in the network can be shared.</td>
</tr>
<tr>
<td>Relationships</td>
<td>The relations in the network will be the exchanges of information in the form of user data. Some decisions regarding the potential relations are not decided but may have an impact.</td>
</tr>
</tbody>
</table>

Table 21: The three core building blocks of the Digisphere value network

**Indirect network effect to increase the amount of data**
The indirect effect of increasing the number of actors in a network is an increased attractiveness for actors outside the network, which in turn will increase the value for actors in the network (Katz & Shapiro, 1985). By increasing the number of actors in Digisphere it could become more attractive for developers of third-party applications. This is because an increased number of actors in the network would give the developers access to a larger potential user group. Increasing the number of third-party applications could have the potential to make Digisphere more attractive for actors to be a part of, which in turn would increase the amount of data available. This is tightly linked to platform theory.
6.3.2 Digisphere as a platform

Exploring if Digisphere could be considered a platform will provide insight into its potential attraction for actors in the online advertising industry.

To become a platform (described in 3.2 Software-based platforms) Digisphere must enable sharing of components or assets across multiple products. It is also a necessity that it facilitates interactions between two sides of a market. Digisphere Core is designed as a software foundation onto which complementary software can be built through its public API. This has two implications:

- Digisphere Core allows its components to be utilized and shared by multiple products.
- Digisphere Core facilitates interactions between two sides of a market by:
  - Providing the actors joining Digisphere with access to third-party applications
  - Providing developers of third-party applications access to the enrolled actors.

Digisphere Core fit the definition of a platform based on the arguments above and can, therefore, be considered a platform. This allows us to apply theory about platform’s competitive advantage, which is leveraging its ecosystem. A thriving platform ecosystem could create products (apps) faster and with greater ingenuity than a single product developer (Tiwana, 2014). This ability could give actors access to services and tools in the form of apps not available outside the platform. Examples of these apps could be new methods for data analysis, offer ways of handling user consent or something completely different, such as Digisphere Bazaar. A thriving platform ecosystem could, therefore, increase the attraction for Digisphere Core.

Examining other attributes of a platform in relation to Digisphere Core gives us:

<table>
<thead>
<tr>
<th>Platform component</th>
<th>Relation to Digisphere Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>Digisphere Core</td>
</tr>
<tr>
<td>Third-party complementary software</td>
<td>Apps (Bazaar is among these)</td>
</tr>
<tr>
<td>Ecosystem</td>
<td>The sum of the enrolled actors and third-party software developers.</td>
</tr>
</tbody>
</table>

96
<table>
<thead>
<tr>
<th>Interface</th>
<th>Public API</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-users</td>
<td>Actors within online advertising</td>
</tr>
<tr>
<td>Rival ecosystems</td>
<td>No direct rivaling ecosystems targeting the same end-users. However, Google and Facebook can be viewed as competitors in terms of data collection.</td>
</tr>
<tr>
<td>Competitive environment</td>
<td>Online advertising in Norway</td>
</tr>
</tbody>
</table>

Table 22: Digisphere Core’s relation to the core components of platforms.

Establishing a platform in a two-sided market is not without its challenges. Since Digisphere is considered a platform this will mean that it will face the “chicken and egg problem”. This means that it will be difficult to attract actors to the platform before the other side of the platform is present.

By applying the design principles in chapter 3.2.1, we can examine how Digisphere tackles the bootstrap problem. The first principle is handled by Digisphere by being GDPR compliant and offering a secure storage for user data. The last two principles do not have a solution and are, therefore, a hindrance to tackling the bootstrap problem.

6.3.3 Digital innovation

Digital innovation creates opportunities for both the breaking down of existing silo structures to create networks and the distribution of innovation between multiple actors. A prerequisite for digital innovation is layered and modular architecture (as described in 3.3 Digital innovation).

Looking at the functionality provided by Digisphere Core in the light of digital innovation could enhance our understanding of how direct and indirect network effects of Digisphere can create attraction. Furthermore, we investigate how the Digisphere components provide a foundation for third-party complementary products and future innovation.

Layered and modular architecture enables digital innovation

The architecture of Digisphere Core consists of digital components with interfaces between them and can be viewed as layered. These qualities enable the digital innovation process to happen and is a key part of establishing both value networks and software-based platforms. The digital innovation theory establishes layered and modular architecture as a base for digital innovation. This theory provides insight into the first steps of enabling attraction towards
Digisphere Core. The layered and modular architecture of Digisphere Core enables the digital innovation process to happen which could increase its attraction. This process will be discussed further below.

**Digital innovation in relation to direct network effects**

Digisphere Sprawl and Mind lays the foundation for sharing user data between actors. The user data is usually stored in the databases of each individual actor. Digisphere Mind & Sprawl will provide the ability for actors to store their user data in a centralized storage structure where it can be utilized by multiple actors. By using Digisphere, data will be moved from each actor’s individual storage into Digisphere’s storage structures. If we view an individual actor’s data storage as a silo it can be argued that Digisphere could contribute towards breaking down the existing silo structures in the industry by providing a central storage where data can be utilized across multiple actors.

Providing a format for storing user data may allow data from different actors to complement each other in Digisphere. User data from different actors could complement each other if there were mechanisms in place for ensuring correct identification of the user in question, ensuring that the data, in fact, belong to the same person. It would also be critical that the data were stored in a shared format, not specific to the actor storing the data, to ensure a successful combination of data from different actors. Digisphere Core could provide both a method for ensuring correct identification and a shared format through its components Mind & Sprawl. Since digital storage of user data would make it a digital component it follows from the arguments above that Digisphere Core can enable different actors to combine complementary digital components (user data) in new ways.

Digital innovation could have the ability to change the relationships involved in innovation and should offer the ability for multiple actors to participate in innovation (Nielsen, 2017). This could, for example, be done by combining digital components in the form of software in new ways, creating new innovative software products. Combining data in new ways could be argued to fit this definition. This is because it has the potential to enhance the value of the data beyond the sum of the individual data components, which could make it applicable for new use cases. The potential in this is that Digisphere Core could function as a venue where actors can innovate by combining data in new ways and by doing this create additional attraction.
One example of this could be if one actor has data about users’ post address and another has data regarding the number of bikes each user owns. Combining these sets of data by matching them on individual users would provide additional value by implying the location in a city it would be most beneficial to advertise a bicycle repair shop.

**Digital innovation in relation to indirect network effects**

Digisphere Core is accessible through a public API and can be viewed as a platform. The public API infer that the core products would have to stay relatively stable, so third-party actors would not have to change their implementation. This could establish a foundation for third-party applications to be created and change rapidly – creating opportunities for innovation in the relations between the two sides of the platform.

The relations between the two sides of the platforms could be another way of distributing innovation. The software providers creating third-party applications to Digisphere Core will be able to create new products faster by utilizing the components of the platform and are provided access to a group of potential users – the actors on the other side of the platform. This will allow actors to create products and test product/market fit quickly. Innovation could happen rapidly and persist as a continuous process where new products are created, evaluated and potentially discontinued.

A continuously improving and evolving ecosystem of third-party applications to Digisphere Core could, therefore, increase its attraction further.

**6.3.4 Summary**

Digisphere aims to provide joining actors with a competitive amount of data to enable them to increase their competitiveness as software providers in the online advertising market in Norway. The aim can be achieved by attracting as many actors as possible to join the Digisphere Core platform. Our theoretical value creation framework provides an understanding with regards to how this could be achieved. This is based on the hypothesis that an increased number of actors leads to an increased amount of available data. The discussion in chapter 6.3 is summarized in the table below.
An increased number of actors can increase the amount of data available for sharing through network effects, thus increasing the attraction to Digisphere. This is because more data attracts more actors, which in turn will attract third-party software providers, which again will attract more actors. This is illustrated in Figure 12.

A thriving platform ecosystem, consisting of third-party software providers and actors using the platform will increase Digisphere’s attraction further. This is because complementary apps can provide actors with services and tools not available outside the platform.

Layered and modular architecture enables digital innovation, value networks, and software-based platforms.

The potential of value creation and innovative relations between the actors using Digisphere Core could further increase its attraction. This is because data could be combined in new ways, potentially creating value that exceeds the sum of its individual parts and value for actors.

Table 23: Summary of theoretical contributions to attraction

In this case, we apply the theory of value network and platforms simultaneously and end up with a topology where the value network establishes one side of the platform and third-party developers the other. Layered and modular architecture enables the digital innovation process which can establish this structure. Digisphere Core provides value to actors enrolling in a value network, even without the presence of third-party software developers. However, the
presence of third-party apps will increase the value of participating in the network. Both these factors will increase the attractiveness of Digisphere. Possibilities for digital innovation for the actors in the network and third-party developers could increase the attraction further.

This is similar to big social network platforms, such as Facebook because being a part of the social network has a value in itself, but third-party app developers can increase this value. The topology of Digisphere is visualized in Figure 13: Summary of the Digisphere platform and network.

An inherent problem of both value networks and platforms is that their initial phase, or rollout, is difficult. This is because their attraction depends on other actors, making it hard to initialize with no actors “onboard”. This problem will have to be overcome for Digisphere to succeed and is addressed in the next chapter.

### 6.4 Factors for success

For Digisphere to succeed the problem related to the initiation of platforms and networks will have to be solved. This problem is found in both value network and software platform theory (Armstrong, 2006; Stabell & Fjeldstad, 1998). The root of this problem for Digisphere is that the initial attraction is low without any actors involved. This is true for both the network and platform part of Digisphere Core and these will affect each other. By establishing a network
as one side of the platform first, Digisphere will become more attractive to third-party developers and the other way around. The network will, however, also have low initial attraction without other actors to benefit from.

The attraction of Digisphere discussed so far is based on network effects. For these effects to create sufficient attraction Digisphere needs to reach critical mass. Critical mass is hard to pinpoint and is not clearly defined for all projects. The project lead of Digisphere defines critical mass for Digisphere like this:

“I do not know when critical mass is reached, but it is when it gains traction. When you are passed a point where people wonder if there is any reason for joining, instead they start fearing that they are missing out on a whole lot of stuff.”

To reach critical mass, Digisphere will have to establish attraction for the initial actors and attributes that could establish this is vital to its success.

6.4.1 Design initially for direct usefulness

Hanseth and Lyytinen (2010) address the problem of initial attraction as the bootstrap problem and present three design principles for tackling it. This chapter will explore how Digisphere can solve the bootstrap problem, building on the first design principle by Hanseth and Lyytinen (2010, p. 9). By designing initially for direct usefulness Digisphere could attract actors to join the network before this is attractive in itself. Digisphere has some attributes that might provide this initial usefulness for actors in the online advertising industry.

Mitigating user concerns

Digisphere and the actors using it will gain a lot by mitigating concerns users might have related to the collection of data. There are two main reasons for this; (1) the actors that have direct contact with the user might find Digisphere Core attractive if it can mitigate concerns for their users, and (2) mitigating user concern will affect the amount of data potentially available in Digisphere Core. We will examine Digisphere Core in relation to the IUIPC framework and discuss the implications of the results.

As described in chapter 3.4, Malhotra et al. (2004) created a framework for measuring Internet Users’ Information Privacy Concerns (IUIPC), based on ten topics regarding the collection, control, and awareness. By doing a survey with questions within the ten topics, a
score based on the IUIPC could be related to Digisphere. If a survey is not feasible to conduct, the ten topics could be used as a general framework to assess Digisphere. By using the framework, we can examine how Digisphere relates to Internet Users’ Information Privacy Concerns, and describe the measures taken to lower the concern. The table below lists the ten topics with a following description of the measures taken by Digisphere, which we believe can reduce the concern that users have.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collection of data</td>
<td>The user will be able to decide who can collect data, and which data can be collected.</td>
</tr>
<tr>
<td>Errors in the data collected</td>
<td>The user will be able to edit any errors</td>
</tr>
<tr>
<td>Unauthorized secondary use of data</td>
<td>The user will have control over everyone who has access and can (at any time) withdraw access.</td>
</tr>
<tr>
<td>Improper use of the data</td>
<td>The user will know what the data will be used for and how with the introduction of consent</td>
</tr>
<tr>
<td>Awareness of privacy practices</td>
<td>The GDPR introduces new privacy policies which will demand that companies provide a clear and understandable privacy policy. Which will also affect Digisphere.</td>
</tr>
<tr>
<td>Control of the data</td>
<td>The user will retain complete control of the data; with possibilities such as editing, moving, deleting, and adding information.</td>
</tr>
<tr>
<td>Global information privacy concern</td>
<td>GDPR will have the potential to reduce the global concern for information privacy concerns.</td>
</tr>
<tr>
<td>Trusting beliefs</td>
<td>Digisphere might increase trust by providing transparency to the user, i.e. the data subject. Also, by having a non-profit owner, Digisphere intends to increase the trust.</td>
</tr>
<tr>
<td>Risk beliefs</td>
<td>Digisphere intends to decrease the risk belief by providing transparency, securely store the data. Digisphere can also increase the belief by being the centralized storage system for data.</td>
</tr>
<tr>
<td>Intention to give personal information</td>
<td>The intention and willingness can, for example, be increased by a “quid pro quo”-system.</td>
</tr>
</tbody>
</table>

Table 24: Digisphere in relation to the IUIPC framework

From Table 24 we can conclude that Digisphere Core has the potential to mitigate some user concerns. This means that this ability can be viewed as a way of creating initial attraction by design for usefulness for actors in direct contact with users.
In addition, the ability to mitigate user concerns is important because it can be viewed as a blocker for increasing the amount of data in Digisphere Core. This is because the users are the owners of their own data, which means that the users can refuse that their data is used inDigisphere Core. Mitigating concerns might reduce the chance that this becomes a problem and is therefore vital to the success of Digisphere.

**General Data Protection Regulation**
Digisphere Core could be a solution for actors to comply with several of the articles in the General Data Protection Regulation (GDPR) and could, therefore, be an incentive for actors to use Digisphere. While each actor could use resources to implement the required changes to their services, they could instead use Digisphere Core as a solution to become compliant with the GDPR. This can be a way of providing initial attraction and usefulness.

Many of the articles in the GDPR have been designed to protect and increase control of the data to EU citizens. EU citizens will hereby be called the data subject. Examples of these articles are *The right of erasure (#17)* and *Data portability (#20)*. The articles in GDPR have been summarized in chapter 2.2.3, and will now be used to explain how Digisphere Core can increase its value as a service by upholding those requirements.

**Authentication**
The basic and most important requirement for implementing the articles is authentication, as it is crucial that personal information is not disclosed to the wrong user. As mentioned in chapter 5.2.1, Digisphere Gate is a solution with two purposes; (1) an ID binding solution and (2) a solution for authenticating users. Digisphere Gate would need authentication so that Digisphere Core can function properly, which means that this requirement is not an extra feature to implement. Authentication will be used to log in to Digisphere Portal.

**Contracts**
The concept of contracts is what solves most of the problems presented by GDPR.

The data subject can decide which actors are allowed access to the data. This could be solved with the first purpose of Digisphere Gate, the ID binding solution; this solution is therefore also a key to handle Article 7, Consent. When giving consent, a binding will be made with a contract describing what information the actor can obtain from the data subject. The only action needed to withdraw consent is to delete the binding, and in that way delete the contract.
This will also solve article 20, Data portability. The user can easily give or withdraw access to data to different actors without the need to manually export data.

An important aspect when withdrawing consent is that the data is not deleted. This means that Article 17, Erasure, is not handled and thus would need to be implemented. This implementation would need to take place in Sprawl and Mind as this is the part of Core that stores information.

Data transparency can easily be solved through Digisphere Core. Through the system with contracts, the user will gain full control and overview of who has access to the data as well as what the data is being used for. This information is stored in the contract and is needed for actors to gain access to the data.

Another problem that contracts solve is article 33, Breach Notification. Which data controller and processor to notify could easily be managed by following the contracts.

By implementing solutions to deal with GDPR, Digisphere Core will increase the appeal for more than Software providers. Actors in both supply side and demand side, and also data subjects, will benefit from Digisphere Gate.

**Low cost of joining**

Digisphere Core will have a low cost associated with joining, which could be another way of increasing the initial attraction. Even though there is no charge for Digisphere’s services, there will be a cost of joining for the actors. This cost will be related to implementing the needed changes to their services to get the most out of the service. Creating well-documented guides for this process will be important for the initial attraction. An ideal situation, in this case, would be that actors only see advantages of joining Digisphere which will lower the barriers of joining.

**6.4.2 Bootstrapping the platform**

Digisphere Bazaar is a complementary product and connects to Digisphere Core’s public API. It can be viewed as a way of bootstrapping the platform because it is the first app. Bazaar could be a direct way of showing how the Digisphere Core can be used and demonstrating its capabilities. By doing this Bazaar could increase the platform’s attraction for third-party developers.
Bazaar could serve as an incentive for actors to join Digisphere. Actors that wish to use Bazaar will be a part of Digisphere Core as well. The actors in this situation will not necessarily share their data with the rest of the network, but this could lower the barrier for the actors to share data in the future and establish them as a part of the network.

Digisphere Bazaar could provide initial attraction for both sides of the platform and could, therefore, be a way of handling the bootstrap problem.

6.4.3 Potential challenges and solutions

Attributes with Digisphere that actors might perceive as negative can have a negative impact on the initial attraction if not handled properly. However, handling these correctly could have a positive effect on the bootstrap problem. Some of these factors are discussed below.

App marketplace

Most traditional software platforms, such as Android, Apple iOS, The Chrome web browser, and Facebook, provide a way to easily explore and consume apps through an “app marketplace” (Google Play for Android, App Store for iOS as examples). A marketplace is a place where available apps are published and consumed. The benefit of this is that it provides a place for consumers and developers to connect. A marketplace would make it easier for the actors in the network to explore available apps. This could enable app developers to assess product/market fit faster and contribute to further digital innovation.

This is at the moment not a functionality present in Digisphere. Incorporating such a functionality could be beneficial, and something that we believe should be prioritized to increase the attractiveness for the actors and third-party software developers.

All the data in one place

A concern with Digisphere is the design choice of storing all the data in one place. This could make it a target for people who want to steal the data. If the actors considering to join Digisphere is concerned regarding the security of the solution this could greatly prevent initial attraction.

One way that Digisphere tries to prevent this, is through the security design where they split the data. The personal information is not clearly connected to each other in Digisphere, this is
done through the implementation of Mind & Sprawl as discussed in chapter 5.2.2. Even though the theoretical idea has been established, the security level still depends on the implementation of the idea. One way of increasing trust could be by making the code open source. The next chapter will discuss pros and cons with open source versus closed source.

**Open source**

Open source allows others to verify that the code does what it is supposed to do and will allow others to contribute to the code. But stating that open source is more secure is wrong, because making the source code public does not guarantee review (Lawton, 2002), and therefore does not guarantee that security flaws will be found and fixed. It can, therefore, be said that “it’s the quality, not the number, of the eyes looking at code that counts” (Lawton, 2002). Closed source has the disadvantage that only a few people can look at the code and fix security issues, as well as forcing the users of the software to accept the level of security that the owner chooses (Cowan, 2003).

By making it open sourced the transparency is increased and we believe that it can establish trust and therefore be a factor for success.

**Addressing the bootstrap problem to a larger degree**

Digisphere has attributes that could aid in handling the problems related to bootstrapping the solution. However, more could be done in relation to the design of Digisphere to attempt to increase the initial attraction. Hanseth and Lyttinen (2010) present three design principles, but only the first one is addressed in Digisphere. Designing Digisphere for increased initial attractiveness based on research on the area could be beneficial.

### 6.5 Design guidelines to increase attraction for a software service

We summarize our research by generalizing the findings and propose a set of design guidelines to maximize the attraction for a software service in the context of online advertising. The process of creating the design guidelines (described in chapter 4.4.1) started with extracting relevant parts of the data collected. The extracted data were then analyzed to identify relevant parts of the design guidelines. This was also done with what we had concluded and discussed earlier in our research with relation to theory. The design guidelines
presented by Hanseth and Lyytinen (2010, p. 9) were used as inspiration. The design principles and their basis in theory are presented in Table 25. The background for each of the proposed design guideline is explained below the table. At the end of the chapter, we discuss how a software service can handle the bootstrap problems and discuss the design guidelines applicability outside of online advertising.

The goal of the design guidelines is to maximize the attraction for a software service by making the software service attractive in itself. This means maximizing value for actors by increasing the potential for value creation through the software service. The guidelines are aimed at designers of software services in online advertising looking to increase the attraction of a software service.

<table>
<thead>
<tr>
<th>Design principles</th>
<th>Design guidelines</th>
<th>Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Create a base for digital innovation</td>
<td>DG1. Build upon principles of layered and modular architecture.</td>
<td>Digital innovation</td>
</tr>
<tr>
<td>2. Create network effects</td>
<td>DG2. Enable the actors using the service to benefit from each other to create a network.</td>
<td>Network</td>
</tr>
<tr>
<td>3. Create a platform</td>
<td>DG3. Facilitate interaction with the network for third-party services&lt;br&gt;DG4. Create a marketplace for the network to interact with third-party services.</td>
<td>Platform</td>
</tr>
<tr>
<td>4. Enable digital innovation</td>
<td>DG5. Create a possibility for the actors in the network to innovate based on their relations.</td>
<td>Digital Innovation</td>
</tr>
</tbody>
</table>

Table 25: Design guidelines to increase attraction for a software service

**DG1:** Modular and layered architecture creates a foundation for the digital innovation process which can result in network effects, platform, and further digital innovation. For this reason, we view modular and layered architecture as a prerequisite for the rest the design guidelines.

**DG2:** Enabling actors using the service to benefit from each other will create a structure where increasing the number of actors will increase its value for actors already enrolled and
actors considering joining. Value will be created in the relations between the actors, facilitated by the software service. This is the network effect and can increase the attraction of the service, illustrated in Figure 14.

![Network effect illustration]

Figure 14: Illustration of network effects

**DG3:** By facilitating interaction with the network for third-party services the service will enable a platform. The network will be established as one side of the platform and potential developers of third-party services will provide the other. Attracting developers could be done by giving developers access to a core set of functionalities of the service through an API. This will enable third-party services to share assets and access a potential user group. Establishing a thriving platform ecosystem can increase the attraction to the network by providing the actors of the network with access to apps not accessible elsewhere.

**DG4:** By providing the network with a marketplace for third-party services these services become more accessible to the actors of the network, which could further increase its attraction for both sides of a platform.

Figure 15 illustrates the effect of DG3 and DG4.

**DG3 and DG4** also enable digital innovation for third-party developers. This is done by providing developers with a relation to the network of enrolled actors which can contribute to faster testing of product/market fit. The reuse of components of the platform through the API can also enable faster development of third-party services, contributing to further innovation. Digital innovation by the third-party developers can increase the attraction to the network further by increasing the usefulness of the available services.
DG5: Facilitating digital innovation in the relations between the actors in the network can be done by creating opportunities for the actors to combine complementary digital components in new ways. What the components are will be dependent on the service, but a shared format will have to be provided by the network to enable this opportunity. The possibility of digital innovation in the network can act as a catalyst for further attraction. This is illustrated in Figure 16.

Handling bootstrap problems
The design guidelines will not create initial attraction by itself. By generalizing our discussion related to how Digisphere can address the bootstrap problem we build on our proposed design guidelines with ideas on how to handle problems related to the low initial attraction for software services.

- **Design initially for direct usefulness.**
  This is design principle is presented by Hanseth and Lyttinen (2010, p. 9) and we add to their design rules that bootstrap problems could be addressed by keeping the cost of joining as low as possible and in this way lower the barriers of joining.
• **Bootstrapping the platform.**
  By creating the first complementary software for the platform one could demonstrate the capabilities of the platform. This could spark attraction for developers of complementary software. Actors considering joining are provided with useful features and increased initial attraction.

• **Handle attributes that might have a negative impact.**
  If the designed software service has attributes which might have a negative impact, these must be identified and handled. This can be anything that can discourage actors from joining. Examples of such attributes can be security or user-friendliness.

**Applicability outside of online advertising**
We have designed these guidelines to be used in the context of online advertising, but we believe they have a potential to be relevant in other contexts. This is something that needs more research and testing because generalization of a theory beyond the case setting can never be achieved from a single case study (Lee & Baskerville, 2003)

### 6.6 Limitations

Limitations of this thesis are discussed in this chapter.

**Validity**
This research is based on a single case study, which is a possible limitation of our research. One case can sometimes be sufficient, but this depends on the uniqueness of the case (Easterbrook et al., 2008). The uniqueness of the case under study can be discussed, but it is possible to imagine that there exist similar cases. Studying multiple cases that are trying to achieve the same would have given us a better data foundation for this thesis. This was, however, not feasible for us because of the difficulty in locating such cases. The extensive time consumption of a case study was also a factor here.

We have not been able to use both qualitative and quantitative research approaches during this study because of the time restraint of a master’s thesis. This could be a threat to the validity of our research because the combination of these approaches could provide better validation during data collection (Robson, 2002).
**Data collection influenced by discussions with co-supervisor**
The data collection regarding the case under study has been influenced by the project lead which also had the role of our co-supervisor. Our data collection has been done through interviews with the “underdogs” of the industry. Most of our interview objects have been startups or companies that are not a part of the “big players”. Although this has allowed us to understand what challenges the Norwegian actors struggle with, we could have benefited from the perspective of the “big players”. We believe that this could have made us biased when prioritizing data collection.

**No video recording during interviews**
We recorded audio and collected documents presented during the interviews. In some cases, this was not enough, as the interview object expressed information through pointing or body language. This posed a challenge during transcription and some information may have been lost for this reason. On the other hand, we were two observers during all the interviews, which helped us recall the information that was presented during the interviews.

**The project’s feasibility is not considered**
A limitation directly connected to the project case is that Digisphere might not be feasible to develop. Time and money are in the IT industry alpha and omega when it comes to software projects. Without financing, Digisphere will probably not be feasible from a startup-up point of view, because money is already such a valuable resource.

In the case that Digisphere has a successful startup, expenses for operation and maintenance will still need to be provided for the entire life cycle. This might mean that the economics of the non-profit free platform described in this thesis will not be possible and might just be a utopian idea used as a desirable end goal.

**Digisphere remains theoretical**
This research is limited to the theoretical description of Digisphere because it has not yet been realized. This is the nature of projects like these and with the time constraints of a master’s thesis, another approach would not be possible. This is, however, a weakness with our research. We would have benefited from the possibility to examine Digisphere as a finished product and the effects it might have had on the online advertising industry and its actors.
**GDPR is not yet implemented**
As the deadline for implementing the requirements for the GDPR is after our deadline for delivering the master’s thesis, reviewing the practical effects this has on the advertising market is not possible. This thesis has therefore only looked at the regulation through a theoretical lens, and our prediction is influenced by this fact. Potential topics for future research regarding GDPR have been mentioned in chapter 7.3.

**New additional information**
Our hypothesis builds on the fact that local software providers in Norway lack the amount of data to be competitive. Because of the lack of transparency in the online advertising industry, it is hard to verify information collected and information can come to light that was not expected. A new report made by The Norwegian Data Protection Authority, published late March 2018, states that the Norwegian advertising company Schibsted has logged over seven million data points on a single test person (Datatilsynet, 2018). Schibsted has a role as a software provider in the Norwegian online advertising industry and this could imply that some Norwegian actors have more data than revealed during data collection.

Because of the timing of the report from Datatilsynet (2018) we have not been able to investigate why Schibsted might be in a stronger position than first assumed, but could have benefited from doing so.

**The hypothesis of “increased actors = increased data” might not hold**
As mentioned earlier some discussions in this thesis are based on a hypothesis that increasing the number of actors in the Digisphere network increases the amount of available data. This might not be the case for several reasons. The end user is given the power to stop actors from sharing their data and might therefore not support the hypothesis. The same can be said for the actors’ willingness to share data among themselves. This might be something that needs additional motivating factors or might not be wanted at all. These factors might invalidate the hypothesis.

In any case, the real world will behave differently than the foundation that we have based some discussions on and will complicate the picture we have described. That being said, the discussions are based on theoretical concepts and are not an attempt at mirroring the real world. This means that the discussions are still valid, and our contribution stands.
Other frameworks for Internet Users’ Information Privacy Concerns
There are other frameworks for measuring the privacy concern of internet users. We choose to use Malhotra et al. (2004). Alternatives, such as Smith, Milberg, and Burke (1996) could also have been used to serve the same purpose. In our view, this would not have changed our research significantly as most of these frameworks consist of the same main elements. IUIPC by Malhotra et al. (2004) builds on Smith et al. (1996) and was the preferred choices for us because it is newer and therefore seemed more relevant to today’s situation. It is also based on an extensive literary review of the research within the area.

6.7 Summary
Data is required for competing as a software provider in the online advertising market in Norway. Digisphere Core could have the potential to enable increased competitiveness for local software providers by creating a way for the actors to share and gain access to an increased amount of data. This functionality means that increasing the number of actors in Digisphere Core will potentially increase the combined amount of data that is shared between the actors.

Increasing the amount of data available in Digisphere Core could be achieved by establishing Digisphere Core as an attractive service and something actors in the online advertising industry want to be a part of. Our theoretical framework will be used to understand how Digisphere could be established as an attractive software service and in that way increase the competitiveness of local software providers. We understand this by applying the theory of value network and platforms simultaneously to end up with a topology where the value network establishes one side of the platform and third-party developers the other. Layered and modular architecture enables the digital innovation process which could establish this structure.

Digisphere Core provides value to actors enrolling in a value network and therefore has a value even without the presence of third-party software developers. However, the presence of third-party apps will increase the value of participating in the network. Both these factors will increase the attractiveness of Digisphere. Possibilities for digital innovation for the actors in the network and third-party developers could increase the attraction further.
For Digisphere to succeed the problem related to bootstrapping networks and platforms will have to be solved. The root of this problem for Digisphere is that the initial attraction is low without any actors involved. For network effects to create sufficient attraction Digisphere needs to reach critical mass. To reach critical mass Digisphere will have to establish attraction for the initial actors. Attributes that could establish this is vital to its success.

By designing initially for direct usefulness Digisphere could bootstrap the network by creating a solution that mitigates users’ concerns, handles problems related to GDPR and keeps the cost of joining low. Bootstrapping the platform is addressed in Digisphere by creating the first complementary app, Bazaar. Digisphere will also have to address potential challenges with its design to ensure that these do not have a negative impact on the bootstrapping process.

We summarize our research by generalizing the findings and propose a set of design guidelines to maximize the attraction for a software service by making the service attractive in itself. The guidelines can be found in Table 25.
7 Conclusion and future research

We will in chapter 7.1 conclude on our finding in accordance with the research question. In chapter 7.2 we outline our contribution. In chapter 7.3 clarify topics which we find interesting that have arisen from our study and could use some more focus than we have had the opportunity to give.

7.1 Conclusion

Our research question asks how digital innovation can enable increased competitiveness for local software providers.

The position that data has in this industry means that having less data than other actors is a weakness. Big actors, such as Facebook and Google, have systems for harvesting data that actors in Norway are unable to match. The result of our data collection shows what we initially thought; that local software providers lack data to compete.

A solution to the lack of data can be to enable local actors to share data between each other and in that way increase the total amount of data available for everyone. To reach a competitive amount of data with this solution as many actors as possible will have to participate in the data sharing. We have in this thesis discussed how a software service can attract actors.

Our research has explored one way that digital innovation can be used to enable increased competitiveness for local software providers by creating attraction for a software service. Digital innovation enables the breaking down of existing data silo structures in the industry to create a value network where data is shared. This network increases its attraction with the increasing number of actors enrolled. The network establishes one side of a software-based platform which attracts a second complementary side; third-party software developers. This could further increase the attraction of the network through access to apps not available anywhere else. Both the network and the platform facilitate further digital innovation, which again increases the attraction of the network. The software service could address the bootstrap problem by designing for initial direct usefulness, create the first app and solve potential challenges in its design.
7.2 Contribution

This thesis contributes in three ways. This section outlines the contributions, who they are aimed at and where in our thesis they can be found.

First, we create a clear description and overview of the online advertising industry in Norway and how this relates to online information privacy. This is aimed at those looking for insight into the practices of the online advertising industry which is criticized for a lack of transparency. The overview can be found in chapter 2.

Secondly, we contribute by describing the status of the online advertising industry in Norway. The challenging situation for local software providers is presented and we describe how these challenges can be met, with the use of a conceptual framework of value creation. This is aimed at IT professionals in the online advertising industry and can be found in Chapter 5 and 6.

Lastly, we propose a set of design guidelines for increasing the attraction for a software solution. These guidelines are created as a contribution to software service design. The guidelines can be found in Chapter 6.5 and are aimed at designers of software services in online advertising.

7.3 Future research

Some topics have appeared during our research which we think could benefit from being studied further.

Additional research on the design guidelines
Because generalization of a theory beyond the case setting can never be achieved from a single case study, more research on our proposed design guidelines should be pursued. This could explore if they can be applied outside the context of online advertising and strengthen their value.

GDPR
More research into the benefits and drawbacks of the GDPR for the actors in the online advertising industry should be pursued. During our research, we have spoken to individuals
concerned that the GDPR will strengthen the big players and increase their competitive advantage. It could also be explored if the GDPR is good or bad for the industry, and what effects it might have.

**Blockchain**
Blockchain has emerged as an interesting technology in relation to cryptocurrency. This could be an interesting prospect for the online advertising industry. Some online news articles are already addressing this opportunity (Moss, 2018; Muthegere, 2018; Olenski, 2018) and we have listed some examples of use cases:

- The immutable distributed ledger found in blockchain could be used to enhance trust and transparency.

- Explore the possibilities of smart contracts such as enabling consumers to own and sell their information

- Explore if blockchain technology has solutions for preventing ad fraud.

That being said, blockchain technology is still immature and cannot solve everything. Further research into this area and the possibilities for the online advertising industry should be pursued.
Bibliography


Russell, M. A. (2013). *Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google, GitHub, and More*; O'Reilly Media, Inc.


Appendix

Attachment A: Interview Guide Part A

Introduksjon
1. Hensikten med prosjektet?
   a. Hvorfor trengs løsningen med tanke på et globalt perspektiv?
   b. Hvorfor trengs løsningen for dere?
   c. Hvorfor har ikke dette blitt gjort før?
      i. finnes det tilsvarende løsninger
         1. hva er eventuelt styrker/svakheter ved disse?
2. Hva er løsningen?
   d. Kan du beskrive løsningen?
   e. Hva lager løsningen om brukeren?
      i. Hvordan identifiserer løsningen brukeren?
   f. Hvordan begrenser løsningen aktørenes tilgang på data om konsumenten?
   g. Hvem skal bruke løsningen?
3. Aktører
   h. Hvilke aktører er involvert? (brukere, myndigheter, andre aktører)
      i. Hvordan har dere blitt kjent?
         i. Hvordan ble det sånn at nettopp dere ble partnere rundt dette?
4. Prosessen for å innføre løsningen
   j. Hvordan skal løsningen lages?
      i. Hvordan finansieres dette?
         1. Budsjett?
   ii. Hvem skal utvikle løsningen?
   k. Hva er strategien for å oppnå kritisk masse?
      i. Hva er kritisk masse for dere?
         1. Hvem er en del av den?
   l. Hvordan vil antall konsumenter påvirke antall aktører, og motsatt?
5. Prosessen så langt
   m. Kan du fortelle litt om prosessen så langt?
   n. Sentrale muligheter og utfordringer?
   o. Hva er det usikkerhet om?
   p. Hva har dere lært så langt?
   q. Hva er planen videre?
Attachment B: Interview Guide Part B

Først litt om oss og oppgaven

Generelt:

- Kan du fortelle kort om hvem dere er og hva dere jobber med?
- Hvordan løser dere salg av deres reklameplasser?
- Hvordan fordeles reklameinntektene?
  - Finnes det andre modeller?
- Tjener dere penger på andre måter?
- Hva brukere dere trackerne til?
  - Har dere tatt et beviset valg på antall og type trackere?
- I hvilken grad tilpasser dere nettsiden for individuelle brukere?
- Er deres reklameplasser (ad impressions) mulig å kjøpe for alle annonsører?
- Hvordan unngår dere upassende reklame?
- Hva betyr GDPR for dere?

Utfordringer:

- Er det noe med dagens marked som dere opplever som problematisk?
- Hvilke utfordringer vil dere få mtp. GDPR?
- Hvoran håndterer dere persondata?
- Hvordan ser dere for dere å håndtere GDPR?
- Hvilke alternativer har dere til DoubleClick?

Bransjen:

- Hva skiller måten dere håndterer reklame fra resten av bransjen?
- Har dere merket noe forandring i markedet mtp GDPR?
Attachment C: Example of data reduction

Workshop on the ID matching project

Startup Alpha/Startup Beta
Location: Posthuset Evolve

Involved parties
The project involves Startup Alpha, Startup Beta Real-Time Marketing and two master students from UiO (Jostein Hellerud and Karl Thomas Hauglid).

What is the project?
The Id matching project is making a service that enables the exchange of id’s between parties that do not want to reveal their internal id’s. The system must be able to match user/id with collected data, and it must be able to follow and match the user/id between services. The service will also solve the privacy challenged caused by the new EU regulation.

Consequences for the industry
The consequence of having an id matching services will benefit several of the involved parties. Securing ID will solve the privacy challenge enforced by the new EU regulation, increasing the transparency in user data collected. This means that the consumer will have full control over the data that is collected and shared between the services. Seen from a marketing and sale perspective the Id’s is the most valuable data on the world wide web. The IMP opens for trade of functionality and data, and it permits commercial collaboration and it can be expanded with metadata. The IMP opens up for smart product data engine for real time dynamic product data. This means personal recommendations of content based on behaviour

Users of the matching service
● ?

Similar existing services
Some similar services were mentioned and needs looking into. These were:
● BankId
● MinId
● SPID
● Open ID (used by Facebook, Google, Microsoft)
● OpenPDS
● Enigma
● Hyperledger
● Masterpass
How the project differentiates from existing services
The ID matching project will develop a free open core with an API. A commercial layer will be built around this core, enabling different actors to commercialize the core.

The system will be built on privacy rights for leveraging the value of trust. The trust of the consumer will be vital to ensure the success of the project.

The IMP will have an ecosystem built on using temp-ID that is shared between the users. It will be scalable and ready for global use.

High level system requirements
The user should be able to create/edit/delete/publish adverts with dynamic data feed content that match specific profile ID.

What the system can do:
● Manage id
● Delete cookies
  ○ Move away from cookies (deprecated technology)
● Use real ID

Tech req:
Latency max 300ms
Retrieve a product based on profileID within the DC

The way forward
For the technical side of the project the following three phases is defined:

1. a. When the banner is loaded Startup Beta will send a request type recommendation and brand/shopID
b. Startup Alpha will deliver Most popular (and similar) together with metadata.

2. a. Startup Beta sends the cookieID together with the banner request
   i. Startup Alpha will deliver a personalized recommendation and product metadata.
   ii. This requires a profiling service with a binding between Startup Beta ID, Startup Alpha ID, and ABC ID.
   b. The profiling service can be extended to contain data and functionality which can be used to tailor the message in addition to product.

The project does now plan on starting on phase 2, instead of 1, because of market demands.
During the workshop a potential show stopper was found. This is the ability to retrieve the userid from the DHTML Lib.

More research into this issue will be a priority in the way forward.

**The business side of the project will look into these challenges:**

- Develop a clear and better value proposition
- Develop and agree on common price model
- Specify customer target group for Proof of concept
- Test value proposition and price model
- Get proof of concept with some brands
- Case study
- Get some PR
- Evaluation of POC and business case
- Release as feature with normal process if successful

**Questions remaining after the workshop**

Who owns the core?
How open is the commercial layer?

What will our role be in the project?
What can we write about in this project?
Attachment D: Examples of data display

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Get token for provider</td>
<td>Creates a temporary authenticity token</td>
</tr>
<tr>
<td>send credentials, TTL</td>
<td>Returns data for token generation</td>
</tr>
<tr>
<td>Get id for id at provider</td>
<td>Registers sent id, if no bindings then create bound global id</td>
</tr>
<tr>
<td>send user/cookie id - provider id</td>
<td>Returns id bound to global id</td>
</tr>
<tr>
<td>Bind id to id for provider</td>
<td>Registers a binding to a global id for the two supplied ids</td>
</tr>
<tr>
<td>send user/cookie id - user/cookie id - provider id</td>
<td></td>
</tr>
<tr>
<td>Get generated id</td>
<td>Registers a new global id and binds generated id</td>
</tr>
<tr>
<td></td>
<td>Returns generated id</td>
</tr>
</tbody>
</table>

![Whiteboard with diagrams](image-url)
Attachment E: Examples of conclusion drawing