Universiteit Leiden

ICT in Business

An Empirical Study into The Reasons of Low Performance of Online Advertisements.

Name: Fen Liu
Student No: s1815490

Date: 25-09-2018
1st Supervisor: Drs. Niels van Weeren
2nd Supervisor: Prof. Aske Plaat

MASTER'S THESIS

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Acknowledgements

This thesis is the final product of my two-year's master program as an ICT in Business student at Leiden University. The launching of this research topic was rough due to the complexity of the areas this topic covered. However, the research process was wonderful and insightful yet sometimes stressful. Here I would like to express my gratitude to the people who gave me guidance, support and encouragement during the process.

First of all, I would like to thank my first supervisor Drs. Niels van Weeren, who helped me a lot from the start to the end. With his expert knowledge on data-driven marketing, he helped me shape my ideas into clear research objectives and guided me to the right direction. Moreover, I highly appreciate that Drs.van Weeren has helped to reach experts through his network.Without his help, it is hard to implement the research. During the process of writing this thesis, his detailed comments and suggestions helped me a lot to finish this work. I would also like to thank my second supervisor Prof. Aske Plaat for his excellent suggestions and for reviewing my thesis.

Second, I would like to thank the people who introduced me to industry experts and gave me the opportunity to do an interview for my research. Needless to say, the experts I have interviewed contribute largely to the results of this research. My sincere thanks to the experts who involved in the interviews. With their help and guidance, I gained a lot of insights into the possible reasons for low performance ads. I would also like to thank the people who participated in the surveys for my research. Without their work, the gap survey cannot be measured in this research.

Finally, I would like to thank my parents for their love and support. I also thank my boyfriend for his great company and comfort whenever I am depressed from stress. Last but not least, I thank all my friends for encouraging me during this long time battle. Without all these lovely people, it is uneasy to go through all the hard times by myself.

Abstract

Personalization was always a hot topic in digital marketing and advertising. But what means personalization for consumers? How accurate an ad needs to be can be counted as highly personalized content? The answer is difficult. There is no 100% accurate targeting since consumers' intention keeps changing. What marketers can do is trying to provide the best offer for their consumers. With the development of various technology in this big data era, personalization becomes much more easier. Meanwhile, the landscape of online marketing also becomes complicated as an increasing number of intermediaries has joined. As a result, the quality of online ads are influenced by multiple parties. In order to find out what factors are still there causing low relevance online advertisements, this study did the first research on this topic that covers a wide range of knowledges related to digital marketing.

Past researches have either focused on methods to improve online marketing performance or algorithms that can produce better recommendations. Since there are not much history researches focused on demonstrating the ineffective problem existed, this research, motivated by the ineffective online advertising phenomenon's existence, aimed to find out what possible reasons influence the performance of online advertisements. Moreover, this research has been conducted during the enforcement of GDPR (25th May, 2018), which is meaningful to see how the landscape will change after its implementation.

The research starts with a gap survey to investigate the gap between customers' expectation and perception about online advertisement. After the analysis of the survey results, three interviews aimed at gathering thoughts from experts in this area to help the design of factors assessment. The assessment design are based on researcher's study on this topic during research period and experts' opinions from interviews. Research findings needed to be validated to reduce bias. In this study, the validation was done by factor assessment among a panel of 10 experts. The result of this research helps to identify what factors existed from five dimensions, namely data management, legal regulation, personalization approach, technology and organization. Among the five dimensions, organization has the biggest influential on the performance of online advertisements. Personalization approach and technology have direct effect on an ad's relevance. However, the decision to use which approach and technology is influenced by organization and legal regulation. Data management is the groundwork for personalization. The result of this research shows data was never the difficult part, the ability to make full use of the data decide the quality of an ad.

Table of Contents

List of Abbreviations

Abbr.	Full Name
AdX	Ad Exchange
AI	Artificial Intelligence
CPC	Cost Per Click
CPM	Cost Per Thousand Impression
CTO	Chief Technology Officer
CTR	Click Through Rate
DMP	Data Management Platform
DSP	Demand Side Platform
EU	European Union
GDPR	Global Data Protection Regulation
IAB	Interactive Advertising Bureau
ISP	Internet Service Provider
JScode	JavaScript Code
OBA	Online Behavior Advertising
KPI	Key Performance Indicator
ML	Machine Learning
PPA	Pay Per Acquisition
PPC	Pay Per Click
ROI	Return on Investment
RQ	Research Question
RTB	Real Time Bidding
SERVQUAL	Service Quality
SME	Small and Medium Enterprise
SSP	Supply Side Platform
TD	Trading Desk

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background

Online advertising is one of the most quickly growing areas in IT industry. Over the past few years, marketing has had to keep up and be content with leaps in technology and our relation to it ever since (*Clodagh O'Brien, 2017*). Dating back to the history of digital marketing, this term was first coined in 1990s (*Clark & Dorie, 2012*). The first clickable banner ad went alive in 1994, which symbolize the change of marketing method (*Joe McCambley, 2013*). It is the first time that marketers could actually know how many people saw and interacted with an ad (*Ryan Singel, 2010*). The legendary first banner ad was purchased by AT&T who paid HotWired $30,000 to place the banner ad above on their site for three months. As a result, they got a click-through-rate (CTR) of 44%, which was a number that's impossible to make today (*Krla Cook, 2018*).

The innovation of this totally different way of marketing attracted marketers to this approach to reach ideal prospects instead of just placing ads wherever space was offered. The increasing demand for targeted ads led to the birth of the ad network, which connects the advertisers and the publishers. They can help digital publishers to increase inventory demand, and help advertisers scale their digital ad buys. Meanwhile, the marketers' need to efficiently manage their ad campaigns across multiple websites gave birth to return on investment (ROI) tracking tools. In the beginning, there was Doubleclick, launched in 1996 (*Ginny Marvin, 2015*). They offered advertisers a new service called Dynamic Advertising Reporting and Targeting, which enabled companies to track how many times an ad was viewed and clicked across multiple websites.

However, pop-up ads rose fast but fell quickly as well. Between 1999 to 2002, marketers turned to paid search and pay-per-click (PPC). In 1999, GoTo.com introduced the first pay-for-placement search engine service. Advertisers could bid for top search engine results on particular keywords. From 2006, digital ads turned to a more targeted way (*Krla Cook, 2016*). Since then, a bunch of techniques can be used to improve the effectiveness of an online ad. For example, real time bidding (RTB), also known as programmatic buying, has become the fastest growing area in online advertising (*Shuai Yuan et al., 2013*). Other techniques like content marketing, native advertising and personalization are all forms of marketing looking for greater efficiency and effectiveness. Marketers use technologies to improve customer experience through coherent messaging. Meanwhile, customers also expect a more consistent and coordinated services nowadays.

1.2 Problem Statement

The biggest problem marketers facing at this moment is that the inefficient marketing brings much less conversions than they expected. According to an Interactive Advertising Bureau (IAB) report, 26 percent of desktop users and 15 percent of mobile consumers use blockers to remove ads from publishers' websites in 2016 (*Lauren Johnson, 2016*). And from a statistic report generated by SmartInsights, across all ad formats and placements, ad CTR is just 0.05% (*Dave Chaffey, 2018*). The fact that there are only 5 clicks per 10000 impressions indicates the low performance of online advertising. For marketers, the fundamental and permanent management question is "how can firms better market their products to potential consumers?" (*Fanjuan Shi, 2017*). All the strategies and techniques they used for online advertising is to gain more conversions. With all the effort and money they put into digital marketing, the results are not as good as marketers expected. There must be reasons why the performance of online advertising is bad. And the ultimate reason would be consumers are not satisfied with the contents marketers send to them.

The factors that influence people's perception about an ad varies from the content, the time, the medium, intention, etc. A consumer's perception can be affected by a lot of different things. The question "When to deliver what content to who through which medium?" should be considered before every marketing decision. The problems on the consumer's side are mainly the following. First of all, the content is the most crucial factor to decide an ad quality. If consumers are not interested in it, it cannot attract any prospects.The personalized content is usually generated based on people's search history (*Aniko Hannak et al., 2017*). However, consumers, no matter intentionally or unintentionally, are in permanent search for marketing information with an objective to substantiate their shopping decisions *(Fanjuan Shi, 2017*). The thing is not all searches represent people's purchase willingness. Moreover, consumer's state of intention keeps changing all the time.

Secondly, time influences people's feeling to a large extent. It is not only about when is the best time to show consumers but also how long should an ad interact with people. People often feel that see an ad for a long time even if they did not show any interest about the offer. The lifetime of an ad should also vary from per industry. For example, people need time to prepare for a trip, thus travel agencies should consider a longer display. On the contrary, people tends to spend less time on consideration when it is a planned or repeat purchase. It is more important delivery ads in time for these purchase relatively.

Thirdly, the personalization technique have some limitations nowadays as well. One problem is cross-device identification. Customer journey usually will involve several devices, but it is hard to identify a customer when he did not log in. This might leave customer an unsynchronized experience. Another problem is recognizing different users using one device. It is weird to see ads intended for your mom but pop up in your computer's web pages only because you share a same internet with your mom. It is also important to identify people's intention to do a search. It is not smart at all to recommend makeup products when someone is browsing for a boyfriend's gift.

Last but not least, all the effort will be a waste if the ads sent through a wrong way. There are channels like search engine result page, web pages, social media platforms, emails, phones, TVs, etc. But which channel is the proper one can influence consumers experience largely. How often do people check their promotion emails? When do people spend time on social media? Do people use ad blocker? All these questions should be taken into consideration.

To summarize, for marketers, in order to apply better techniques to improve customer experiences, it is important to recognize the existing problems and what factors influence the performance of an ad. For consumers, the quality of online advertisements still have a lot need to be improved.There are researches on using smart techniques better understand consumers' online behavior, there are researches on recommending better products based on users history data (search/purchase data), but there is no research provide a comprehensive analysis about the reasons that are influential to the performance of online advertising. Based on all the problems existed, it is necessary to conduct a research on finding the reasons of online advertisements low performance.

1.3 Research Objectives

The main purpose of this thesis is to **research and understand what influences the performance of online advertisements. In other words, what makes people feel online advertisements are not relevant under today's digital marketing situation**. Though online advertising tends to be more tailored, as long as receivers do not appreciate it, it is a waste of efforts. Due to the fact that online ads catch less and less people's attention (*Thales S. Teixuira, 2014*), it is important to know how consumers feel about the content marketers created for them. Furthermore, if there is dissatisfaction, what do people feel unhappy about? Specifically, the aim of this study can be narrowed down into the following objectives:

1. *Estimating the gap between people's expectation and perception about the performance of online advertisements.*

The study starts from the gap between people's expectation and perception about online ads. Thus, it is important to show how large the gap is, what kind of marketing strategies actually do not work well on consumers, and what needs to be improved.

2. *Understanding the mechanism of personalized advertisements design and the delivery.*

In order to recognize possible factors, it is necessary to understand the state of art. The landscape of digital marketing has become a very complex game with multiple players. Every player in the game might contribute to the low relevance of an ad. If we want to look at the essence through the phenomenon, we need to understand the mechanism behind the screens. It includes how people's data have been collected and processed, what algorithms have been used to design a personalized content, and how it is be delivered to the right person.

3. *Measuring the factors that influence the performance of online advertisements.*

After having the knowledge to deliver tailored content, the relationship between every stakeholder and the processes from the start to end, the problems that actually influencing the performance of online advertising will be more clear. The method adopted in this research to measure possible factors is by expert interviews.

4. *Under the circumstance of Global Data Protection Regulation's (GDPR's) enforcement, discussing how the landscape of online advertising will change in the future.*

Since the research is conducted during the time of GDPR's enforcement, and the data protection problem is a very sensitive problem in this field, it would be interesting to see the change and how will it influence the future of online advertising. This was done by the researcher's experience, expert interviews and studies from other researcher's work.

1.4 Research Questions

This thesis project involve both scientific study and in-company interview in order to fulfill the objective mentioned above. In general, the research question for this study is defined as follows:

"What factors influence the personalization level of online ads?"

In conjugation with the specific objectives, the main research questions can be further analyzed in the following sub-questions:

1. ***What does personalization means for consumers?***
   The whole research focused on the word "personalization". Then it is necessary to clarity what is personalization. Based on the problems existed, the researcher think there is a gap between people's expectation and perception of online advertisement. Therefore, a gap survey was designed in this research to get consumers' opinion about personalization. From the survey we can know what do they expect and what is their real experiences.

2. ***What dimensions decide an online advertisement's relevancy level?***
   The delivery of an ad is not decided by one party. There are a few stages that need to be considered. From the data collection, data processing, insight gathering, to the delivery of an ad, every stage might have several factors that's influential to online ads quality. In order to have a clear scope of this research, it is better to see what dimensions are influential first.

3. ***What problems existed in different dimensions?***
   With a clear scope of this research, we can then go deeper into the problems existing in each dimensions. The problems existed in each dimension are the most direct evidence to proof how influential that factor is. And it also an important process to answer the main research question.

The research questions also follow the research process which tries to reach the research objectives step by step. The gap survey answers the first sub-question which aimed at showing the problems from consumer's experience. Moreover, it predicts how the state of art will change the future. Literature review gives a strong academic support to the theoretical groundwork and helps answering the second and third sub-questions. Expert interview is a good way to consider the current situation from a more practical point of view. Further more, it can help to explain possible reasons for some problems from a professional angle with less bias.

1.5 Thesis Structure

This thesis consists of eight parts and conducted by empirical research to gain knowledge by means of literature review, survey and expert interview. The figure below summarizes the composition of this thesis.

Chapter 1 Introduction	Chapter 2 Theoretical Background	Chapter 3 Research Methodology
Background Problem Statement Research Objectives Research Questions Thesis Structure	Online Advertising Landscape Customer Journey Personalization Processes	Literature Review Gap Survey Expert Interviews Validation Assessment
Chapter 4 Literature Review	**Chapter 5 Gap Survey**	**Chapter 6 Expert Interviews**
Data Management Legal Regulation Personalization Technology RTB Ecosystem	Survey Background Dimensions of Online Ads Quality Result Discussion	Interview introduction Interview Processes Result Discussion
Chapter 7 Validation Assessment	**Chapter 8 Conclusion**	
Introduction Result Discussion	Answers to Research Questions Limitations and Future Research	

Figure 1 Structure of the Thesis

Before the research work, related work includes the introduction of this thesis, online marketing's state-of-art and the methodologies used. These are explained in chapter 1, 2 and 3 respectively. The main research work of this thesis involve three parts: survey, expert interviews, and factors assessment. First of all, research works review is the cornerstone of this thesis. A systematic literature review in history research themed in web data collection, personalization approach, legal regulation, technology and RTB help finding obstacles that have been raised and studied in former researches. After that, chapter 5 explains the processes and result of the survey and chapter 6 goes into detail about the interview processes and result analysis. The setting of chapter 7 mainly talking about the validation of findings from expert interviews. The assessment design is based on the findings from expert interviews and knowledge from the literature review. Chapter 8 concludes the whole thesis into four parts which are: answers to research questions, research limitations and future research.

Chapter 2 Theoretical Background

The first thing before going to research is we need to know is the-state-of-the-art. This industry is complex and dynamic. It keeps changing with new technologies arising. Customer journeys become more complicated, with more convenient devices and channels offering shopping online. The relationship between advertisers and publishers has become intricate with a lot of intermediaries doing the transactions. Therefore, it is important to go through the basic theoretical knowledge, the landscape of online advertising and the processes to deliver an ad. This chapter introduces online advertising evolving, types, landscape, customer journey, and the techniques used.

2.1 Online Advertising

Online advertisements are pervasive nowadays. You might see a banner ad with objects you searched while checking today's news, you might find local stores ads when you are checking your favorite bloggers feeds on social platforms, and you might also have an email account with tons of promotion content which you seldom open. We see these ads everyday online, but what purpose do they serve? How does it work? Who plays what role behind it? All these questions can help understanding the mechanism behind.

This type of advertising evolved from the first banner ads into display ads. As one of the oldest form of online ads, display ads appear as everything from banners of all shapes and sizes to text ads relevant to the content of a page, which led to the beginning of performance tracking. With the development of ROI tracking tools, advertisers became more interested in targeting specific consumer demographics, rather than just placing their ads wherever space was offered and hoping the right people would see it (*Krla Cook, 2016*).

After years of development, online advertising means any type of marketing message that shows up with the help of Internet, which means it could appear on search engine result pages, social media platforms, emails or your mobile applications. The way online advertisements look keeps changing with the development of various techniques and the increasing competition. Marketers tend to like online advertising more due to the fact that it can reach more people with low investment. It can be tracked to measure the ROI. Moreover, marketers can personalize the content to a targeted group by analyzing the data gathered from Internet.

As a result, there is a myriad of terms to describe ads nowadays, but there is no strict way to segment ads into different types. You can segment it based on the publish platform which are web pages, social media platforms, emails, phone applications, internet connected TVs, SMS, etc. But you can also group it based on the ad's format. There is banner ads, pop-up ads, interstitial ads, banner swapping, floating banners, flash, video, sky scrappers, PPC, cost per thousand impression (CPM), etc. Or, you can just segment it according to the function of an ad. There are ads designed for retargeting, there are ads intended for recommendation, and there are ads just aimed at brand awareness. Other intentions might be promotion action, improve customer base, purchase persuasion, etc. No matter what is the type, the target, the publish platform of an online ad, the trend is going to a more humanized and personalized way. Unlike traditional unidirectional marketing, the advertisers can now know how many people actually viewed instead of an approximate idea of how many people might see or heard the ads. Moreover, information like age, gender, location, preferences, etc, can all be collected with simple tracking. This kind of advertising, which tracks individual online behavior in order to deliver advertising tailored to his or her interest, is usually called online behavior advertising (OBA).

2.2 Landscape

Concerning how online ads work, there exists variety of models. Due to the fact that there are more than one or two stakeholders, the relation between marketers and publishers is not as direct as before. The classic way today is that there will be one or more middle man who communicate with both publishers and advertisers to diminish the fact that the demands exceed supplies largely. For example, Ad Network acts as a sales representative or broker buying unsold or remnant inventory from the publisher. They apply technologies to understand consumers' needs and sold packaged inventories to the buyers, which makes it easier for advertisers to target the people they want. Ad Exchange (AdX) creates opportunity for the buyers and sellers to trade audience rather than inventory in the thousands sellers would make their audiences available on the platform, buyers can pick their audience and bid on them.

The difference between Ad Network and AdX is that impressions are been sold in different ways. The Ad Network provides aggregated and packaged inventory while the AdX provides opportunity to buy a specific audience. Apart from these two, some agencies invested in demand-side platforms which give them the ability to trade on the AdX efficiently and in real time using data to influence their decision making. Some publishers sold directly to AdX, others invested in a group of company called sell-side platforms, which optimizes the selling points for the publisher. In general, supply side platform (SSP) enable web publishers and digital media owners to manage their inventory space while demand side platform (DSP) allows buyers of digital inventory to manage multiple ad exchange and data exchange. The figure below can show a clear relationship between all players.

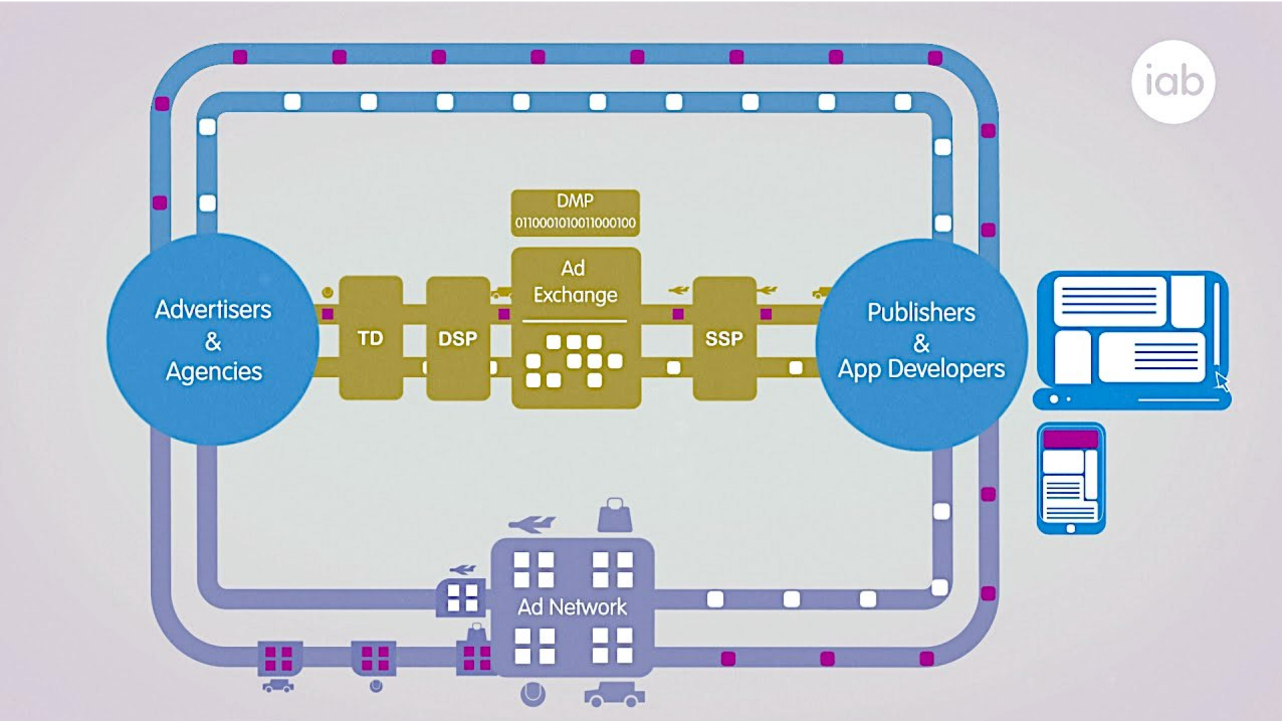


Figure 2 Online Advertising Ecosystem   Source: IAB

A typical ad delivery might start when a publisher declares a few spaces for ads. AdX or Ad Network gathers insight from data they analyzed and hold an auction for advertisers, as long as both publisher criteria for the ads they want met advertisers'. The one who bids the highest price, gets the opportunity to deliver their ads.

2.3 Customer Journey

The customer journey spans a variety of touchpoints by which the customer moves from awareness to engagement and purchase (Definition by FORRESTER). The term touchpoints is seen as the building blocks of customer journey in the sense that customer journeys are defined as a set or sequences of touchpoints (*Asbjørn Følstad & Knut Kvale, 2018*).

One thing that makes online marketing become more complex is the increasing number of touchpoints in a consumer's shopping experiences. Due to the fact that consumers can obtain information from various devices, the customer journey would also span several stages, channels and devices. Different devices have their own strengths and weaknesses on showing information to customers. Consumers might be attracted by an ad shown in his/her social platform, but he might prefer to search more information on a laptop with a larger screen. Consumers change devices for searching, comparing, or waiting for the best offer. Before making decisions, some customers might want try the products/services by him/herself in store, which means there's one more touchpoint causing incoherent experiences.

The reason to include customer journey here is that complicated customer journey is one of the main challenge to overcome. In the Figure below, you can see a typical customer journey with possible touchpoints in both digital and physical.

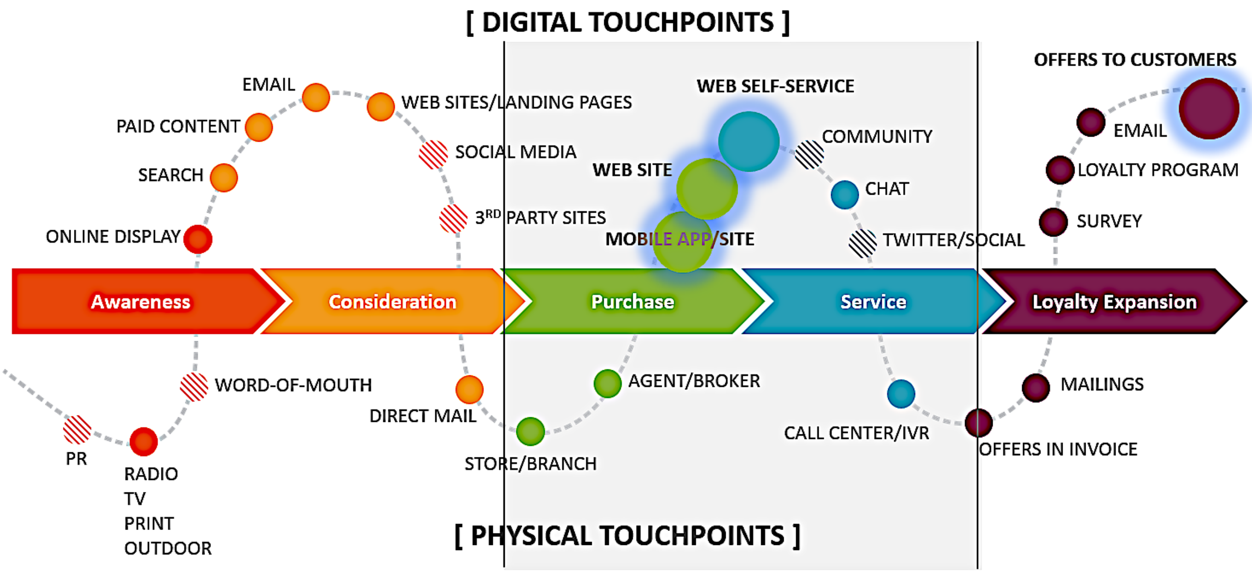


Figure 3 Customer Journey with Possible Touchpoints

Source: eComgrowth.fi

From the figure we can see there are five stages customers might go through. And each stage includes a few touchpoints in different channels. Generally speaking, increasing touchpoints means there are more channels for marketers to attract and communicate with consumers. But in another aspect, it means any mistake during the process can ruin the whole customer experience. It is important to develop a seamless experience that ensures each touchpoint interconnects and contributes to the overall journey.

2.4 Personalization Processes

2.4.1 Data Collection

The start of personalized ads is cross-channel behavior consolidation, data collection and processing. Without data, there is no way to implement personalization. In order to investigate what the possible factors leading to low quality online ads are, it is necessary to understand what kind of data have been tracked and used to target prospects. How do they track a consumer's online behavior? And how they translate data into insights that helps helping marketers make decisions?

First of all, what data will be tracked and used on personalization? Based on the explanation of ad personalization from Google, information like types of websites you visit, mobile apps you have on your device, cookies (a small piece of data sent from a website and stored on a user's computer by a web browser, definition by AddThis Academy) on your browser, websites and apps you have visited that belong to the businesses that advertise with Google, your interactions with Google's ads or advertising services, and your Google account activities will be collected under user's approval. Another industry giant, Facebook, as a social platform, has 2.19 billion monthly active users (*Statista, 2018*), which is a strong proof that they can easily reach prospects that marketers want. According to the privacy policy provided by Facebook, they collect contents, communications and other information users provide when they use their products. It include sign-up information, all the content users create and shared, photos, users networks and connections and users usage (the type of content users view or engage with, the actions users take, the people or accounts users interact with, the time, frequency and duration of users activities). Moreover, device information like device attributes, device operations, device signals, data from device settings, mobile operator or internet service provider (ISP), language, time zone, mobile phone number, IP address, connection speed, even information about other devices that are nearby or on your network can be collected.

Basically, they collect every move from users if users do not switch it off. Except for big companies like Google and Facebook, there are a lot middle size and small size companies also work on data collection and analysis. For example, Segmentify, as an e-commerce personalization platform, helps online retailers optimize their conversion rates by enabling them to deliver a unique shopping experience for each visitor. According to Segmentify, the global e-commerce conversion rate is 2.95%, which is the cause of high customer acquisition costs and low profitability (*From Segmentify official site*). Thus, their aim is to create better personalized online shopping experience.

Companies like Segmentify mainly track data by adding a piece of tracking code (e.g. JS code) on retailer's websites. With this simple code added on your site, you can easily collect information like pages viewed by a contact, duration of views, total number of pages viewed in a visit by a contact, IP address, etc.

In general, there are three types of companies that will track data for different aims. Big tech companies like Google and Facebook, also act as publishers, can get a lot users information since they have a huge amount of users, and also has the best technical team do the data analysis. The difference is the data they collected is different due to the products/services they provide for users are different.. Another type is business companies from various industries themselves, they collect data to better understand their customers. The tables below show all possible data that could be tracked online. Table 1 includes information that's not generated from users, and table 2's data will only appear when there are active user behaviors.

Table 1. General Information Could be Tracked Online

Location Data	Device Data	Connection
Geo Location	Device Type	Public IP
Address Name	Operating System	Local IP
Language	Browser Name	ISP
Local Time	Browser Plugins	Download Speed
Weather	Hardware	
	Display	
	Battery	

Table 2. Dynamic Data that could be Tracked Online

LogIn Data	Web	Social Media	Mobile Phone	Clickstream	
Name	Log In History	Posts	Applications OnDevice	Arrive	High Speed Move
Email Address	Search History	Chat	Phone Number	Search	Low Speed Move
Phone Number	Read History	Likes		Click	Download
Gender	Save History	Saved		Scroll	Remove
Age	Purchase History	Shared		Hover	Purchase
Account ID	Cookies	Connected Account		Add	Quit
Password	LoggedIn Social Account	Followed			
		Commented			

After knowing what data could be collected, it is also important to understand how the data be tracked. For web activities tracking, the main method is using a tracking code inserted in each page, which is usually a small piece of JavaScript code. The most common types of tracking technologies are cookies, beacons, and pixels. Cookies can be sessional or persistent. The difference is persistent cookies will not be deleted when the browser is closed. Websites mainly use cookies for two purposes, either to keep you logged in or to track your behaviors. Beacons are small transmitters that connect to Bluetooth-enabled devices like smart phones. They are commonly used in marketing to send messages to an app based on proximity to the beacon (*Agnieszka Gąsiorek, 2014*). Pixel refers to the code that's placed on website in order to trigger a cookie. Essentially a pixel is an image, but instead of calling an image, it called an application on a media buying platform (e.g, DSP) that will cause the cookie to downloaded to the user's browser (*Hafez Adel, 2012*).
In practice, websites usually track non-logged users by using external third-party software (e.g Google Analytics, AdRoll). Among third-party trackers, there are single- and multi- website trackers like Google Analytics keep the data of their client websites solo and isolated from each other. Multi-website trackers like AdRoll share users's activity across all of their clients (*Robert Heaton, 2017*).

Through tracking codes, when you successfully logged in a website, you will be assigned a long random session ID, which will help the website immediately recognize you without asking for username and password again. This is how cookies help realizing the same users is logged in. With the session ID attached to each request that the website receives, websites can therefore track users' behavior data. In practice, websites can still collect behavior data even the users do not log into their systems. In this case, they cannot know your username or any information you provided to register on a site. But they can still know every movement you did on their site. When a website use third-party cookies help them do the tracking, the information they can collect will extend from the website you currently visiting to any other website that also uses their tracking services (*Robert Heaton, 2017*).

This section mainly answered the questions "What data could be tracked online?" and "How is the date been tracked by which method?". Based on the knowledge we have, it is easier to understand how personalized ads can be produced, which can help find out possible reasons that lead to a low of personalization. The next section will introduce the analysis process, which can answer the question "How is the data translated into meaningful insights?".

2.4.2 Personalization

Personalization is defined as any action that adapts in information or services to the needs of a particular user or a set of users, taking advantage of the knowledge gained from the users' navigational behavior and individual interests and preferences, in combination with the content and the environment (*Eirinaki & Vazirgiannis, 2003*). It is a customer-oriented marketing strategy that aims to deliver the right content to the right person at the right time (*Aguirre et al., 2014*). The strength of this strategy is that it requires a minimum amount of effort by the customer, who relies mostly on the marketer to identify and meet his or her needs (*Joel Jarvinen & Heini Taminen, 2015*). In another words, for customers, they can receive information they want or need as trade of giving personal information like location, search history, email, etc.

In this section, in order to figure out how data be translated into insights like preferences and interests, it is useful to go through the approaches that have been applied on content personalization. Since the personalized content is either through recommendations or retargeting, the main algorithms have been used will be elaborated below.

2.4.2.1 Recommendation

Recommender systems or recommender agents refers to the automatized recommendation of products, services, and contents to users (*Ville Salonen & Heikki Karjaluoto, 2016*). The most four popular personalization algorithms include consumer profiling, collaborative filtering, content-based, and particularly different types of hybrid. Consumer profiling is a portrait telling us the characteristics and the behavior patterns of one customer or a group of them (*Fanjuan Shi, 2017*). There are three method used to deal with different type of data. "Labelling" is usually used on factual information which does not change much. "Rule-based profiling" is used to seek conditional facts about one or a group of consumers, such as user transactional histories data. The last method, "sequence", aims at portraying the behavior patterns that signify the owners of the profile. The content-based personalization mainly analyzes consumer's previous choices combined with acquired information to predict subsequent needs. Collaborative filtering approach aims at predict consumer's rating patterns on items based on how they rated items in the past. The last approach, hybrid approach, uses the multi-approach mentioned above to achieve better personalization. In practice, most firms use hybrid approach most (*Fanjuan Shi, 2017*).

2.4.2.2 Retargeting

Retargeting refers to advertising targeted to customers based on their past action at the advertiser's website (*Navdeep S. Sahni et al., 2017*). A big part of the ads people see online belongs to retargeting. Retargeting is easy to perform. It works just like cookie, by placing a JavaScript tag in the websites, the code can create a list of people who visited the site. As a result, these people will receive display ads or banners on other websites he/she visits for a while.

The mechanism behind retargeting is relatively simple, but the performance of retargeting can be pretty low since consumers are in a permanent search, not every search indicate shopping intention. This means that simple click streams cannot really show a people's interest and willingness to purchase. The tricks it can perform are choosing target users that went to a specific part of the site, or open it up to the whole site. The benefits of this method is that merchants only spend money on people who already have experience with your brand, and probably did some searching or clicking on the website. Usually display media buyers tend to target users in the first stages of customer journey. However, the drawback of retargeting is also enormous. Since retargeting takes a lot of guesswork, too much retargeting can work against you as well (*Joanna Lord, 2011*). For example, when a consumer searched for one brand but feel it is not his/her taste, the retargeting mode will against his/her wishes when he/she decide not going to buy anything from this brand.

2.4.3 Real-time Bidding

RTB is an emerging and promising business model for online computational advertising in the age of big data (*Yong Yuan et al., 2014*). It is a programmatic instantaneous auction, which allows impression buyers to launch their advertising campaigns via multiple ad-networks (*Shalinda Adikari et al., 2015*). It is a technology firstly introduced in 2009 which can help ad networks buy and sell inventory easier (*John Ebbert, 2012*). The implementation of RTB makes that every online ad impression can be evaluated, bought, and sold individually and instantaneously (*Jeff Green*). The rise of RTB comes from the explosion of choice of where display ads can run. With millions of sites accepting display ads, it is too difficult for media buyers to buy the audience they want when buying directly from each individual site. Thus, RTB can helps media buyers find audience at scale.

A RTB ecosystem has two sides, an advertiser side and a publisher side. Each side has its own components and techniques in the bidding processes (*IAB Europe White Paper, 2014*). A typical process of RTB ad delivery starts from a user browsing on a website. When a user opens a webpage, an auction will be triggered. Based on analysis of cookie data and data from Data Management Platform (DMP), the publisher will send user information to SSP, who can forward the information to AdX. The Ad Exchange further sends it to DSPs, who contact with its advertisers and starts an auction. The winner from each DSP auction will enter the second-round auction in the AdX. The highest bidder can finally get this auction and deliver the content to the user on the webpage. Generally speaking, the whole process can be summarized in to three stages: audience identification, auction and ad display. It is said that it only takes 10 to 100 milliseconds to finish the full processes (*Yong Yuan et al., 2014*). The figure below gives a clear view of the relation between each player.

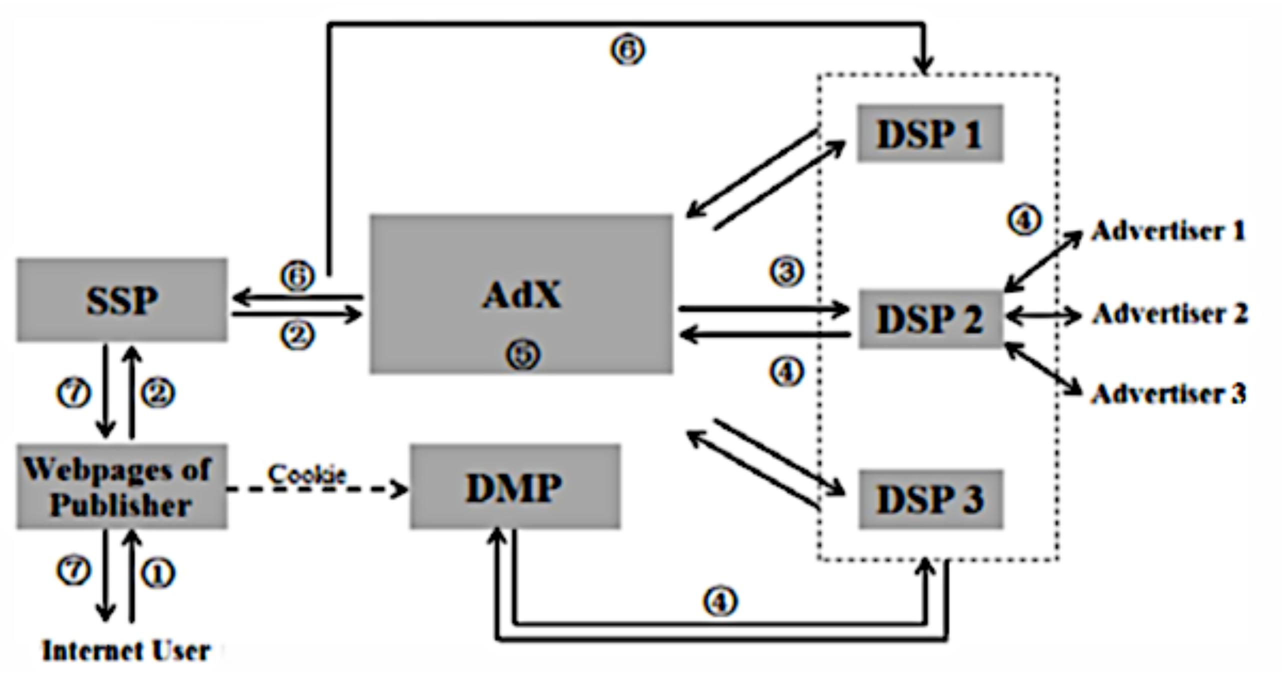


Figure 4: The Business Process of RTB Ad Delivery
Source: Yong Yuan et al., 2014

According to Yong Yuan, a leading DSP company can process the cookie data of more than 570 million Internet users, and characterize every cookie with 3155 attribute labels. More than 3 billion ad impressions are sold by a single DSP every day, and each ad impressions is auctioned within 50 milliseconds. Obviously, RTB plays an important role to deliver online ads. Moreover, the bidding algorithms influence publishers' and advertisers's benefit directly, and also decide whether the impressions sent to the right person or not. Ongoing research on bidding mechanism is trying to find more economical and efficient methods, which is regarded as one of the most important techniques to improve the performance of online ads.

Chapter 3 Research Methodology

This chapter provides information about the methodologies used through this study. There are generally four methods used and each method aimed for different purposes. The results from each research can be used to answer the research questions discussed in chapter 1.

There are two types of methods of research which are normally the most used in the collection of data, these are identified as following: quantitative and qualitative methods (*Wahyuni, S, 2003*). Data is collected and transformed into numbers which are empirically tested to see if a relationship can be found in order to be able to draw conclusion from the results gained. Qualitative methods emphasize on understanding, interpretation, observations in natural setting and closeness to data with a sort of insider view (*Ghauri et al., 1995*). Gunnarson argues that the benefit of applying a qualitative method is that the method takes into consideration the overall picture in a way that the quantified method cannot. A qualitative approach will be more suitable in order to fulfill the purpose of this research, since this thesis is researching what factors are influencing customers engagement. Due to the fact that perceptions, beliefs, ideas, and opinions are difficult to measure in a quantitative way, quantitative research is appropriate for this study. However, quantitative methods are also used in this research in order to measure the gap between consumer's perception and expectation about online advertising. Data are collected through a questionnaire that was completed by 148 people.

Due to the magnitude of the studies, there is a need to collect as much representative data as possible from the selected experts of companies in order to be able to accomplish a deeper understanding of the external and internal factors.

First of all, a quantitative research aimed at measuring the gap between people's expectation and perception about online ads was done. The result of the survey can indicate the problems that existed right now. Secondly, a literature review provides a solid ground of argument for further research, due to the fact that this research topic covers a big part of online advertising instead of a specific technique study. So the literature review was conducted to find clues for potential reasons that cause low relevance ads to be shown to consumers. Investigating past studies concerning varies aspects of designing a personalized content, it can give an overview of not only the landscape now but also historic problems. With all the groundwork, expert interview plays a role to connect the knowledge we have and the real situation. Their professional ideas helps forming the scope of this research, therefore push the research to a deeper level. The final validation was done by an assessment. It is unavoidable that there are bias when it comes to personal opinion. Thus, the validation assessment is essential for this study to arrive the final results. Table 3 represents the methodologies used in regards to the research questions and sub-questions that they are related to.

Table 3. Methodology in relation to Research Questions

Methodology	Research Questions	Descriptions
Gap Survey	RQ 1	To show the problems existed with authentic data by estimating the gap. The design of this survey aimed at a clear view of to which extent people feel unsatisfied with what.
Literature Review	RQ 2 & 3	Collect information/knowledges from history research about the data collection, personalization methods, legal regulation, technology and RTB ecosystem in order to generate questions for expert interviews.
Expert Interview	RQ 2 & 3	Experts working in this field can give the latest and professional opinion. The setting of this research offers researcher the opportunity to uncover information that is not accessible.
Validation Assessment	RQ 3	After getting insights from the experts interview, an assessment designed for experts who worked in this industry can help getting measurable results about the findings from previous steps.

3.1 Literature Review

Due to the fact that this research needs knowledge form different areas, this study used a systematic literature review from three angles, namely data collection, personalization algorithms and delivery approach. The reason to set the literature review in three different aspects is that these three themes present the process to deliver an online ad. There are a lot of studies on specific technology used on performance improvement, but no research tried to analyze the overall situation with all the processes involved. Using all types of knowledge from past studies, we can begin to understand possible reasons for today’s low relevancy online ads. The review includes 56 research articles in total, 14 research articles themed on data collection, 16 studies on personalization and studies on legal regulation, technology and RTB Ecosystem. Apart from research articles, this study also reviewed more than 25 reports from tech-companies, which helps connecting the situations right now. The literature search was conducted on several search engines: Science Direct, Elseiver, Google Scholar and Leiden University Library’s database. The main key words used to search relevant literature include: “online behavior tracking”, “cookies”, “web tracking”, “personalization”, “personalization algorithms”, “web personalization”, “real-time bidding”, “RTB algorithms”. Besides using key words looking for literature, references from relevant literature also help reaching related articles. Table 4 shows the amount of literature that was reviewed for this study, grouped by theme.

Table 4. Main literature in the dataset

Theme	Number
Data Collection	14
Personalisation	15
Legal Regulation	5
Technology	4
RTB	7
General	10
Total	55

3.2 Gap Survey

The survey is designed to better illustrate the problems from the view of consumers. Though online advertising is becoming more personalized, users are aware of the benefits and unwilling to assist the improving process. Some people might have the stereotype that marketers steal personal data for their own benefits. Some people try every method to block the ad online and avoid third-party collect their browsing history. And some people just let it go, but do not pay much attention on it. These behaviors stopped the very first step to interact with consumers. Therefore, this research designed a survey to investigate how do people feel about the ads today and what they expect. The survey was conducted from 14th May to 23rd May with 175 people participated. Because of missing values, the number of valid responses varied. The aim of this survey is to show gap between consumer’s expectations and perceptions by ranking the statements related to online ads performance from 0 to 10. The result of this survey can reveal what problems trouble people most, what qualities should an ad have, and the future of online advertising. The detail of this survey process and the results analysis are elaborated in Chapter 5.

3.3 Expert Interviews

There are different approaches to qualitative interviewing, unstructured and semi-structured interviewing. During an unstructured interview the researcher might start the conversation with a question and then actively listens to the respondent who talks freely while a semi-structured interview follow a checklist of issues and questions that the researcher wish to cover during the session (*Darmer, 1995; Bryman & Bell, 2007*).

In this thesis both primary and secondary data are collected. The secondary data used has been critically evaluated and has been collected from relevant literature, databases and internet sources. The primary data is collected by semi-structured qualitative interviews with experts from relevant companies in The Netherlands. There are three experts interview conducted in total. Each interview was analyzed with transcripts using the software Nvivo. Insights concerning the possible reasons can be generated from the expert interview. According to *Blaxter et al. (2006)*, it is worthwhile doing interviews because it offers researchers the opportunity to uncover information that is "probably not accessible using techniques such as questionnaires and observation". Moreover, they add that interview is not merely a data collection tool, it is rather a natural way of interaction that can take place in various situation. Additionally, *Dornyei (2007)* argues that with the presence of the interviewer, mutual understanding can be ensured, as the interviewer may rephrase or simplify questions that were not understood by his/her interviewees. This data can be recorded and reviewed several times by the researcher to help producing an accurate interview report *(Berg, 2007)*. However, there are indeed also some drawbacks as well. Table 5 below shows the advantages and disadvantages of interviewing.

Table 5 Advantages and disadvantages of interviewing

Advantages	**Disadvantages**
High return rate	Time-consuming
Ewer incomplete answers	Small scale study
Can involve reality	Never 100% anonymous
Controlled answering order	Potential for subconscious bias
Relatively flexible	Potential inconsistencies

3.4 Validation Assessment

With regard to the advantages and disadvantages of interviews and to make position clear, although interviewing is a powerful way of getting insights into interviewee's perceptions, it can go hand in hand with other methods "providing in-depth information about participants' inner values and beliefs" (*Hamza Alshenqeeti, 2014*). In this research, an assessment based on the interview results would help obtaining richer data and validating the research findings. After getting insights from the expert interviews, an assessment was generated and sent to 10 experts. The aim of this assessment is to get measurable data concerning the findings from the expert interview.

The design of the assessment was based on the findings from previous work. Participants were asked to grade 30 statements that have been designed by the researcher on their relevance to the performance of online ads. The ranking scale is from 1(very small relevance) to 10 (very big relevance). Participants can give their grade from 1 to 10, depend on how much they think that statement actually influences the performance of online ads. The statements are separated into five types, namely data collection, legal regulation, personalization approach, technology and organization. A detailed explanation of this assessment and the results analysis are stated in chapter 7.

Chapter 4 Literature Review

In chapter 2, a basic knowledge background has been introduced. In this chapter, the focus is more on the findings from past research. Five literature streams related to data management, legal regulation, personalization, technology, and RTB Ecosystem help both researcher and readers understand how online advertising works. Moreover, this review helps finding clues about the problems detected in earlier researches.

4.1 Data Management

The table below shows all literature selected concerning the data collection methods and mechanisms applied on online behavior understanding. Some articles examined the techniques used to collect data and their defects. Some focused on data synchronization approaches and some analyzing the tracking mechanisms. Through these literature, we can find knowledge concerning what data can be collected online, what methods can be used to collect data, what drawbacks each method have and what trouble it might cause for further personalization.

Table 6. Summary of selected main literature on data collection

No.	Citation	Year	Topic
1	Jonathan Stearn	1998	Cookie, Mechanism, Myths of cookie
2	Lazcorreta et al.	2008	Data mining
3	Liao S.H. et al.	2009	Data mining for one-to-one marketing
4	Mika D. Ayenson et al.	2009	Flash cookies
5	Niklas Schmucker	2011	Web tracking
6	Jonathan R. Mayer, John C. Mitchell	2012	Third-party web tracking policy
7	Gunes Acar et al.	2014	Persistent tracking mechanisms
8	Steven Englehardt, Avind Narayanan	2016	Online tracking
9	Muhammad Ahmad Bashir et al.	2017	Ad blockers
10	Anna Kobusiriska et al.	2017	Fingerprinting
11	Johannes Haupt et al.	2017	Email tracking
12	Panagiolis Papadopoulos et al.	2018	Cookie synchronisation
13	Fisk et al.	2018	Users recognisation
14	Arjaldo Karaj et al.	2018	Online tracking landscape

Data can be the most important thing for personalization since it is the a primary resource that marketers can use to understand consumers. It has kept developing since the start of online advertising. There are many approaches to track visitors online behavior. Tracking was never the difficult part, but with the increase of online websites, the customer journey becomes much more complicated. The data volume increased to a great extent, which causes difficulty in data integration and synchronization. The following section will summarize the knowledge and findings concerning the definition of web tracking, the tracking methods, data synchronization and ad blockers from the literature selected.

4.1.1 Web tracking

Online tracking is the name given to the process by which third-party services on websites collect and aggregate data about user's activities and actions on the web. It is presented on multiple different websites with a significant combined traffic and uses cookies or fingerprinting methods in order to transmit user identifiers. Online tracking can also be characterized as the collection of data about user interactions during the course of their web browsing. This can range from simply recording which types of browser access a particular page, to tracking all mouse movements and keystrokes (*Arjaldo Karaj et al., 2018*). Earlier research done by *Niklas Schmucker* has given an introduction to the web tracking and provided an overview over relevant technologies. The major motivations for web tracking are used to tailor individualized advertisements. Instead of showing random ads to users, their profile information, for example, age, sex, and other sites visited in the past, is taken into account to choose content that's relevant to their interests. This simply means trackers can be used for analytics, advertising, conversion tracking, social media, content delivery networks, comments and customer interaction.

When talking about web tracking methods, people usually would think about cookies. But there are more technologies can be used to track browsing behavior on websites. For instance, flash cookies, server logs, widget, web beacons, tags, tracking bugs, pixel trackers or pixel gifs. There are a few academic papers telling the differences between these different tracking methods. But all of them are designed for getting information of users online. The table below gives a description of each term. Through the descriptions, we can have a clear view of the differences between them. The descriptions are adopted from several literature among all selected literature (Niklas Schmucker; Steven Englehardt & Avind Narayanan; Mika D. Ayenson et al.; Arjaldo Karaj et al.).

Table 7. Different Types of Tracking Methods

Tracking Types		Features
Cookies	Session Cookies	A session cookie for a website only exists whilst the user is reading or navigating the website. When the user closes their web browser these cookies are usually removed.
	Permanent Cookies	A persistent cookie for a website exists on a users computer until a future date. For example the cookie expiry date could be set as 1 year, and each time a website is accessed over this period the website could access the cookie.
	Evercookies	Evercookies is a JavaScript-based application which produces zombie cookies in a web browser that are intentionally difficult to delete.
	Zombie Cookies	A zombie cookie is a cookie that is automatically recreated after being deleted.
	First-party Cookies	First-party cookies are cookies set with the same domain (or its subdomain) as your browser's address bar.
	Third-party Cookies	These are installed by third parties with the aim of collecting certain information to carry out various research into behaviour, demographics etc.
	HttpOnly Cookies	A HttpOnly cookie can only be used via HTTP or HTTPS, and therefore cannot be accessed by javascript. This reduces threat of session cookie theft via cross site scripting (XSS).
	Secure Cookies	A secure cookie can only be used via HTTPS. This ensures the cookie data is encrypted, reducing the expose to cookie theft via eavesdropping.
	Supercookies	A supercookies is a cookie with an origin of a top-level domain (such as .com) or a public suffix (such as .co.uk). Ordinary cookies, by contrast, have an origin of a specific domain name, such as example.com.
	Opt-out Cookies	Opt-out cookies are cookies created by a website on your browser folder that enables you to block that same website from installing future cookies.The opt-out cookie tells the website not to install third party advertiser or other cookies on your browser.
Fingerprinting		Fingerprints can be used to fully or partially identify individual users or devices even when persistent cookies (and also zombie cookies) can't be read or stored in the browser.
Pixels	PostBack URLs	It is used for server-to-server cookie-less tracking
	HTML Pixel	It is used for cookie-based browser tracking
Beacon		Beacon is a technology based on the Bluetooth Low Energy (BLE) transmission principle. It enables automated, energy-saving communication between transmitters (so-called beacons) and receivers (e. g. smartphones, tablets or Smart Watches) so that relevant content can be called up and made available by the receiver based on its local geo-data.
Web Beacons/ Web Bugs/Pixel Tags/Clear GIFs		A web beacon is any of a number of techniques used to track who is reading a web page or email, when, and from which computer. They can also be used to see if an email was read or forwarded to someone else, or if a web page was copied to another website.
Widget		Widget is a small program or application which provides or interacts with information in a moveable and customisable way.
Server Logs		A server log is a log file (or several files) automatically created and maintained by a server consisting of a list of activities it performed.

Cookies can be separated by their features like the table above. It can also be grouped based on its function. General speaking, there are technical cookies, analytics cookies, advertising and marketing cookies, and social media cookies. Technical cookies are cookies needed for the proper performance of websites, such as language preference, to be able to login in or to detect fraud with your account. Permanent cookies, session cookies usually belong to technical cookies. Permanent cookies can be removed through the settings feature of browser. Session cookies will be automatically deleted when the web browser closed. Web analytics services usually rely on analytic cookies to evaluate the use of the website (e.g. Google Analytics). The browser chat cookies which offer a chat module on the website also count as an analytic cookie. Concerning advertising and marketing cookies, there are third-party cookies, advertising cookies and re-marketing cookies. Most third-party cookies purpose is to match advertisement to users profile, which makes ads as relevant and interesting to users as possible. Google, Facebook and other advertising platforms use advertising cookies to show relevant ads based on users' recent queries and site visits. Some promotional content uses re-marketing cookies for re-marketing via Google, Facebook or other advertising platforms. The most common social media cookie is share buttons, which only work when sites using scripts and codes that are generated by networks like Facebook and Twitter (*Information based on multi-sites cookies policy, for original sites please check the reference list*).

As a whole, cookies are commonly used for session handling, storage of site preferences, authentication and the identification of clients. Third party cookies are not set for the domain the user is currently viewing, but for the external domain from which additional data, such as images and scripts, was fetched. Most cookies can be removed by users, while Zombie cookies can use client-side JavaScript code to recreate it. Ever-cookie is a well-featured and popular open source Zombie cookie implementation (*Niklas Schmucker, 2011*). Most often, web developers use cookies to store user IDs and passwords on a client's drive, eliminating the need to remember then on return visits to the site. Cookies can also be used to prepare customized web pages. Another popular application is the storage of item's in a client's shopping basket in online ordering systems. Cookies can also help in sequencing of banner ads, making sure a user is not presented with the same banner twice in a row. They also give webmasters a better idea of how many individual users are visiting their site, how often and when (*Jonathan Stearn, 1998*).

Compared with cookie-based tracking, fingerprinting-based tracking can create a unique and persistent identifier for a device or a browser. It differs from cookie-based methods in that this value is a product of the host system, rather than a saved state, and therefore cannot be deleted or cleared by the user (*Steven Englchardt & Avind Naraynam, 2016*).

Other than single tracking codes, web analytics software is also commonly using third-party services. For instances, Google Analytics, One Engagement Hub, Evergage, etc., are used to track the customer journey. These services usually also require webmasters to include a JavaScript code snippet into their websites, which then downloads more tracking code from a third-party server. A different approach does not rely on client-side code, but instead extracts information directly from the web server's log files. An advantage of this approach is that it also works if the client has JavaScript turned off. Moreover, all the data are stored and analyzed on one server, which can provide higher customer privacy. However, client-side technology can collect more information on the local machine than the browser sends out by default (*Niklas Schmucker, 2011*).

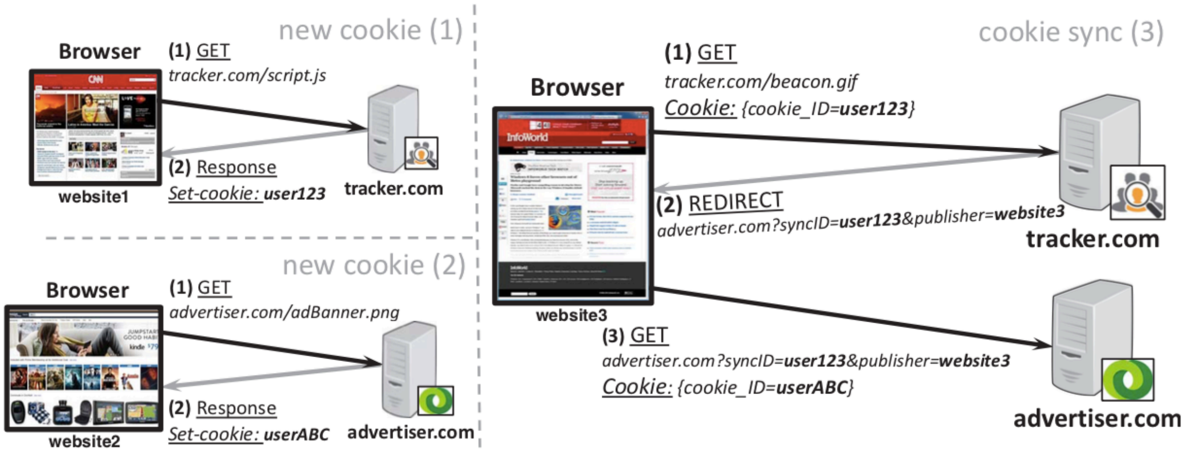
4.1.2 Data Synchronisation

With the development of multi-channel and omni-channel marketing, the customer journey also involves more platforms. However, each tracker knows the same user with a different ID, so it is hard for the collected data to be sold and merged with the associated user data of the buyer. Therefore data synchronization becomes crucial to identify all of a user's activities. Cookie synchronization can facilitate an information sharing channel between third parties that may or may not have direct access to the website the user visits. Cookie syncing allows different trackers to share user identifiers with each other. It is an essential technology that can be used for specific users and merge their database on the background.

At time of writing, a recent study conducted by *Panagiolish Papadopoulos et al.* on May 2018 performed an in-depth study of cookie synchronization. It is said in their research that initially, first-party cookies were used to track users when they repeatedly visited the same site, and later, third-party cookies were invented to track users when they move from one website to another. However, the same-origin policy was invented to forbid the cross-domain tracking of users to protect their privacy. The policy consequently restricts the potential amount of information trackers can collect about a user. Cookie synchronization is designed to overcome exactly this restriction.

Figure 5 adopted from their (*Panagiolish Papadopoulos et al., 2018*) research is a simple example to understand how cookie synchronization works. When a user is browsing several domains, each domain has its own third party trackers. Therefore, it is impossible to know which user is visiting a which website without the third party tracking code. However, let's assume that with cookie synchronization, a user visits website1 and website 2, in which there are third parties tracker.com and advertiser.com respectively. tracker.com knows the user with the ID user123, and advertiser.com knows the same user as userABC. When the user visit website3, which include tracking code from tracker.com but not from advertiser.com. Thus, advertiser.com can not recognize this user as userABC they have in their system.

Figure 5 Data Synchronisation Exemple



How cookie synchronization works is when tracker.com is called by website3, it will also instruct the browser to issue another request to advertiser.com using a specifically URL: GET advertiser.com?syncID=user123&publisher=website3.com; cookie: {cookie.ID=userABC}. Through this way, third party cookies are able to track users across a wide spectrum of websites even if those websites do not collaborate with each other. Of course there are privacy issues concerning cookie synchronization. The analytics from this research is based on 850(Greek) users' web browsing traffic for 12 months in 2016. The GDPR was not implemented yet at that time. However, the finding that 97% of the users are exposed to cookie synchronization at least once can reveal how the situation was before the enforcement of GDPR.

According to another research done by *Steven Englchardt & Avind Naraynam* in 2016, most third parties are involved in cookie syncing. The most prolific cookie-syncing third party is doubleclick.net, sharing 108 different cookies with 118 other third parties.

4.1.3 Ad blocking

According to IAB, ad blocking is technology that consumers use to prevent the download or display of advertising. Browser extensions are the most common forms of ad blocking. Ad blockers are controversial and pose complex issues. On one hand, the IAB has said: “As abetted by for-profit technology companies, ad blocking is robbery, plain and simple—an extortionist scheme that exploits consumer disaffection and risk distorting the economics of democratic capitalism.” On the other hand, the IAB acknowledges that “ad blocking is a crucial wakeup call to brands and all that serve them about their abuse of consumers’ good will.” (*Benjamin Shiller et al., 2018*)

Past research concerning ad blocking topic usually investigated the effects of ad blocker usage on web users’ browsing experience and websites revenue. Research done by *Benjamin Shiller et al.* analyzed the effect of ad blocking on website traffic and quality. They found that each additional percentage point of site visitors blocking ads reduces its traffic by 0.67% over 35 months. The impacts on revenue are compounded: ad blocking reduces visits, and remaining visitors blocking ads do not generate revenue. However, it is hard to say whether 0.67% traffic is significant or not. In another study token by PageFair, it is estimated that over 600M devices worldwide were using ad blocking by the end of 2016, of which over half were mobile (*Matthew Cortland, 2017*). Another research company, comScore, provides estimation of the share of users employing ad blockers for seven countries on June 2015, respectively are Canada (16%), France (27%), Germany (24%), Netherlands (14%), Spain (14%), the United Kingdom (10%), and the United States (%).

In *Benjamin Shiller et al.*’s research, the causal impact of ad blocking on site traffic was measured. It is stated in their research that ad blocking is also likely having substantial effects on the revenue for a variety of sites delivering ad-supported content online. Though they don’t have direct evidence on revenue, their results show that sites with more users who block ads experience reductions in traffic, which they presume arise from the sites’ loss of revenue. Their interpretation of the result is that revenue reductions undermine investment which, in return, compromises site quality, making consumers less interested in visiting in the first place.

When users usually adopt ad blockers, it usually would lead to two consequences. First, ad blocker usage by a site’s visitors reduces the site’s revenue. Reduced revenue may undermine a site’s ability to invest, which could manifest itself as a diminished site that is less appealing to potential visitors. In the end, it will fall into a vicious spiral, having fewer visitors diminish a site’s value, and less investment leads to fewer visitors. Another consequences is if a site's remaining revenue-generating visitors are most tolerant of ads, it will lead the site to run more ads, therefore increasing the value of the site.

Another study done by *Ben Miroglio et al*. focused more on identifying how different user engagement with the web is for users who installed an ad blocker in their browser from those who did not. They found that people who already spend a lot of time on the internet (advanced/ experienced users) are more likely to install ad blocking extensions. It is validated in their research that ad blocking will diminish user engagement with the web. *Muhammad Ahmas Bashir et al.* researched on how companies circumvent ad blockers in 2017, since users are more aware of their privacy online, use of ad blockers does cut marketers' way to a specific user. In general, researches hold the opinion that ad blockers use do influences the quality and traffic of a site.

Other than blocking ads with ad blocker software, there are other methods, for instance, private browsing modes, opt-out cookies and do not track header. Most browsers provide ways to prevent users from being tracked, either directly or via browser extensions. All major browser vendors include private browsing modes into their browsers. In private mode, typically cookies and other browser persistence mechanisms are disabled. No browser history is recorded, and writing of caching information to disk is prevented. However, sophisticated tracking services might still be able to identify users uniquely. The influences on online ads of using private mode haven't been measured in any research. Compared with millions data generated by users, the loss might be a drop in the ocean. However, when concerning a single person's online experiences, there's no way to provide personalized content for users who use private mode a lot. Some companies that perform web tracking offer users the possibility to opt out of their program by setting a special output cookie, which gets recognized by their tracking servers. However, for users who regularly clear their cookies may also accidentally delete their opt-out cookies (*Niklas Schmucker, 2011*).

As a whole, the research examined in this study all indicate the use of ad blocking might cause trouble in several aspects, namely the quality of sites' content, the visitors experiences, the traffic of the sites, and ultimately, the revenue of the sites. However, due to the fact that the revenue of site is also affected by a lot of things, there is no strong evidence can give a clear answer to the question to which extent will ad blocking will influence websites revenue.

4.2 Legal Regulation

Web browsing history is inextricably linked to personal information. It might cause harm to a consumer from malicious employee, hacker, government agency, etc. The data might be transferred, sold, stolen, misplaced, or accidentally distributed without consumers concern (*Jonathan R. Mayer & John C. Mitchell, 2012*).

Data protection and privacy has been discussed since the rise of the internet. There are two main parties dominating the discussion as to how to protect personal privacy on the Internet. One called for government regulation (*Jared Strauss and Kenneth S. Rogerson, 2002*). It seeks legislation that would set strict limits on how companies collect data online, what types of personal information they can collect, and how they can use it (*Marcia S. Smith et al., 2006*). Another party argues that market and industry self-regulation will yield better results than government rules (*Robert E. Litan, 2001*).

In *Dennis D. Hirsch*'s research, he studied the law and policy of online privacy, and stated the difference between regulation, self-regulation and co-regulation. The European Union's 1995 Data Protection Directive allows E.U. member nations to experiment with a co-regulatory approach to the protection of personal data. In Hirsch's opinion, the European model is neither self-regulation nor pure government regulation. The government retains an important role in reviewing, approving, and enforcing the proposed codes of conduct, but it is the industry associations instead of regulators draft the detailed rules and standards. Therefore, it is a form of "co-regulation" in that government and industry share responsibility for drafting and enforcing regulatory standards (*Dennis D. Hirsch, 2011*). The 2002 ePrivacy Directive, 2002/58/EC, mandated that websites enable users to opt out of having information stored in their browser, except as "strictly necessary" to provide service "explicitly requested" by the user. The 2009 amendment to the ePrivacy Directive, 2009/136/EC, replaced the opt-out rule with an opt-in consent rule. In February 2012 the European Commission proposed a new set of revisions to EU data protection law. Recommended provisions would clarify that consent must be explicit, unambiguously extend the reach of regulations to non-EU companies that track EU residents (*Jonathan R. Mayer & John C. Mitchell, 2012)*.

The EU's general data protection regulation (GDPR) took effect on 25 May, 2018, which arised from the public's attention on privacy issue. The aim of the GDPR is to protect all EU citizens from privacy and data breaches in an increasingly data-driven world that is vastly different from the time in which the 1995 directive was established. Compared with the 1995 directive, there are several key changes. The biggest change to the regulatory landscape of data protection comes with the extended jurisdiction of the GDPR, as it applies to all companies processing the personal data of data subjects residing in the European Union, regardless of the company's location. Moreover, the penalties increased, the conditions for consent have been strengthened and companies will no longer be able to use long illegible terms and conditions full of legalese, as the request for consent must be given in an intelligible and easily accessible form. Apart from these, the data subjects also should be assigned the rights to access, to be forgotten, and well notified *(EUGDPR.org*).

The enforcement of new regulation surely increased data processors' attention and be more cautious about using data. However, the impact on different companies varies. For big companies like Facebook and Google, users are more willing to give their consent since they are using the services. However, for small companies like a startup business, it might be harder to persuade users to give their consent since the service is not irreplaceable.

In a project done by *Arjaldo Karaj et al*., 2000 websites were profiled and compared. The tracking landscape in these sites was taken as a function of the origin of the users visiting. They find that since April 2018, the average number of trackers per page in the EU has dropped by almost 4% while in the US it has increased by 8%. Moreover, the reduction is more prevalent among categories of sites with a lot of trackers. Concerning online advertising, they found that in Europe, most advertisers appear less. Google's advertising services have maintained their market share, while other advertisers across the board have lost reach. This is understandable since Google has significant resources compared to others. In general, GDPR has a measurable impact in reducing the average number of trackers websites put in their pages. The biggest contribution of GDPR is the increased transparency on how personal data is collected and used. Moreover, it has led the online advertising market become more concentrated as the majority of advertiser lose market share (*Arjaldo Karaj et al., 2018*).

4.3 Personalization

Personalization is defined as any action that adapts information or services to the needs of a particular user or a set of users, taking advantage of the knowledge gained from the users' navigational behavior and individual interests, in combination with the content and the environment (*Eirinaki & Vazirgiannis, 2003*). Personalization is used to enhance customer satisfaction, to improve sales conversions, and facilitate purchase decision (*Vesanen, 2005*). The objective of personalization system is to provide users with the information they want or need, without expecting from them to ask for it explicitly (*Mulvenna et al., 2000; Montgomery & Smith, 2009*)

In this section, we make an overview of the personalization methodologies used most by firms in their marketing personalization system.

Table 8. Summary of selected main literature on personalization

No.	Citation	Year	Main Topic
1	Murat Soyal, Ece Guran Schmidt	2009	Performance Evaluation
2	Kwiseok Kwon, Cookhwan Kim	2009	Personalization design
3	Elizabeth Aguirre et al.	2014	Online Advertisement Effectiveness
4	Zohreh Dehghani et al.	2014	Context-aware Recommender System
5	Ming Li et al.	2014	Predict User Interest
6	Wei Wang et al.	2014	Targeting Advertising
7	Maurits Kaptein et al.	2014	Explicit and Implicit Personalization
8	Fanjuan Shi et al.	2015	Context Adaptation for Smart Recommender System
9	Ville Salonen, Heikki Karjaluoto	2015	Web Personalization
10	Yuqian Zhu, Junghua Chang	2015	Key Role of Relevance in Persoanlization
11	Cong Li	2015	Actual and Perceived Personalisation
12	Cong Li, Jiangmeng Liu	2016	Web-based Personalization Effect
13	Jose Estrada-Jimenez et al.	2016	Online Advertising
14	Fanjuan Shi	2018	Content-aware marketing personalization
15	John T. Gironda, Pradeep K. Korgaonkar.	2018	Consumer's Perceptions of Personalized Advertising

The literature selected in this study covered the full aspect of personalization knowledge. Topics are range from the mechanism (*Salonen & Karjaluoto, 2015; Shi, 2018*), the design of personalization (*Kwon & Kim, 2009; Zhu & Chang, 2015*), the recommender system (*Dehghani et al., 2014; Shi et al., 2015*), and the performance of personalization (*Soyal & Schmidt, 2009; Aguirre et al., 2014; Li, 2015; T. Gironda et al., 2018*). Moreover, the selected works were published from 2009 to 2018, which not only shows how the situation changed but also can ensure it is not outdated. There are also some other references adopted from the literature selected in above table. Among all of the literature, there are more technique focused articles, addressing topics such as recommender systems, data collection and processes, or user profiling. Comparatively fewer studies consider the quality and usability of web personalization.

The most recent research that examined how the field of web personalization has evolved in the past 10 years and where the field stands today is *Ville Salonen & Heikki Karjaluoto'*s research concerning the state of art and future avenues of web personalization in 2015. It also offered insight into the most notable gaps identified in the literature and identify important future research directions. It is said that web personalization is appealing as a concept but it is difficult to implement as a business tool.

It is said that "personalization done" and "personalization done well" produces different results (*Fan & Poole, 2006*). However, what constitutes "personalization done well" keeps evolving, as both customer expectations and technological possibilities change. As described by *Hawkins et al.*, when a message is addressed to a particular individual, it is considered as highly personalized. When a message is designed based on some common characteristics of a population subgroup, it can be regarded as moderately personalized. Finally, if a message has no specific target, it is generic or non-personalized (*Cong Li, 2015*).

The success of web personalization relies on accurately detecting and then reacting to current preferences. However, preference finding is difficult (*Chen et al., 2010*). In the web personalization literature, preferences have often been viewed as static (*Tuzhilin, 2009*), while in reality, contextual issues such as timing, location, and phases in the buying process keep preferences in a flux. The complexity of customer preferences and lack of knowledge of the contextual effects make it difficult to establish successful web personalization procedures (*Ville Salonen & Heikki Karjaluoto, 2015*). In order to achieve highly personalized level, there is a lot that needs to be taken into consideration. Other than shopping journey, sales channels, the personalization approach, contextual factors like position, time, state of mind, shopping goal, budget, user interface, legal regulation also need to be taken into consideration.

The personalized content can be produced either by a site itself or third parties. Usually there are specific algorithms designed for different types of data and marketing aim. In *Fanjuan Shi*'s research, four of the most common personalization methodologies were analyzed, namely consumer profiling, content-based personalization, collaborative filtering and hybrid approach.

The consumer profile method uses the characteristics recorded in a profile range from demographic, economic, geographic, and psychographic features to preferences, shopping patterns, purchase history, and financial capability (*Gunter & Furnham, 2014*). It is widely used in personalization tasks to identify "similar neighbors" for active consumers, whose preferences need to be predicted. Content-based personalization seeks to refine consumers' interests in their preceding choices and use acquired knowledge to predict their subsequent needs. This approach is useful to identify items and topics similar to those have been liked by a consumer in the past. However, it requires historical consumer feedback, which makes it vulnerable to the impact of the cold start problem (problem that you start out without any information about consumers). The collaborative filtering approach aims to suggest personalized products to consumers by predicting their rating patterns on item based on their explicit and implicit rating history. There are two types of collaborative filtering methods: memory-based methods and model-based methods. Memory-based methods perform personalization tasks based on item-item or user-user relations. The item-based method predicts an active user's rating of a new item based on his/her ratings of similar items in the past (*Sarwar et al., 2001*). The user-based method discovers a group of users similar to the active user, and uses their rating history to predict the active user's rating of the new item (*Marlin, 2003*). Model-based methods seek to predict an active user's rating of new items by modeling the components and precesses that determine the rating pattern (*Sarwat et al., 2002*). Common model-based methods include Naive Bayes, Associated Rule Mining, Clustering, and latent factor models (*Koren, 2008*).

Unlike the previous three methods, hybrid approach refers to the simultaneous use of more than one personalization method. In general, there are two hybrid approaches to integrating multiple personalization methods. Mixing is referred to as the combination of items predicted by different personalization or non-personalization methods with a purpose to diversify personalization results. Twisting is referred to as the use of several personalization methods in personalization modeling (*Fanjuan Shi, 2018*). In practice, most firms use a hybrid approach to obtain more balanced personalization results which take into account the relevance and novelty (*Burke, 2001; Adomavicius & Tuzhilin, 2005*). The applicability of these personalization methods expands significantly in online advertising. Online retailers like amazon.com and bol.com provide personalized recommendations based on collaborative filtering (*Montgomery & Smith, 2009*). A considerable amount of existing personalization systems follow the process-driven method, which takes consumers' behavior and preferences as input, uses a predefined algorithm or algorithm group to process the data, and then provides output. The figure below illustrates how process-driven personalization work.

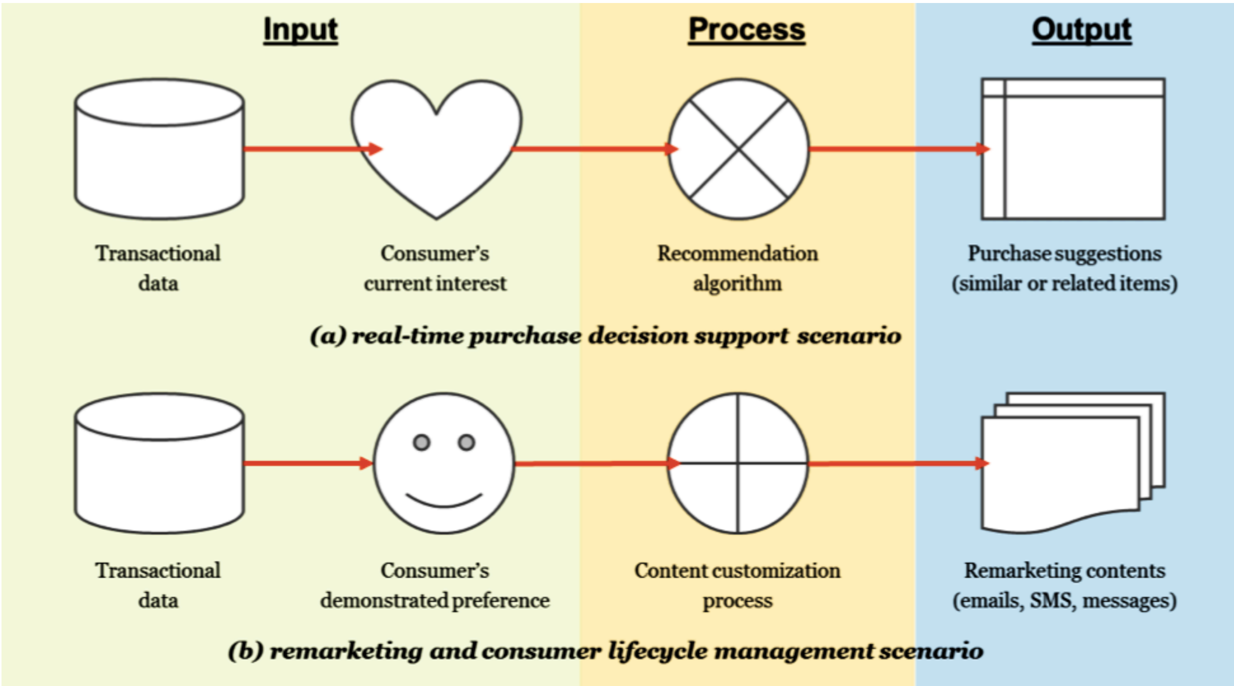


Figure 6: Process-driven Personalization. Resource: *Fanjuan Shi*

The effectiveness of advertising heavily depends on the relevance of the advertisements to a user's interests. Many online advertisers turn to targeted advertising through an ad broker, who is responsible for personalized ad delivery that caters to user's preference and interest (*Wei Wang et al., 2015*). In *Zhu and Chang*'s research, it is stated that the relevance of the advertisement plays a key role. Relevance is defined as the degree to which consumers perceive an object to be self-related or in some way instrumental to achieving their personal goals and values (*Celsi & Olson, 1988*). In the personalized advertising context, relevance is defined as the degree to which consumers perceive a personalized advertisement to be self-related or in some way instrumental in achieving their personal goals and values (*Jiwon Lee et al., 2017*).

4.4 Technology

Marketing is among the most frequent applications of technology. Technology is the ultimate source of personalization. Without innovative technique to translate the data into insights, all the data collected is useless. Along with the development of online shopping, more and more powerful technology has been developed and applied on personalization. According to the hype cycle for digital marketing and advertising made by Gartner (Gartner Hype Cycles provide a graphic representation of the maturity and adoption of technologies and applications, and how they are potentially relevant to solving real business problems), six technologies have a particular use for marketing. Mobile marketing analytics, cross-device identification, multitouch attribution, predictive analytics, artificial intelligence, and customer data platforms in particular can help marketing teams capitalize on customer-centric trends. In this section, these marketing technologies will be discussed based on existing literature.

4.4.1 Predictive Analytics

In advanced analytics, predictive analytics is a branch of advanced analytics which is used to make predictions about future events which are unknown. For marketers, in order to explore churn management, cross-selling, purchase possibility, customer life value prediction, etc, predictive analytics uses many techniques, such as data mining, machine learning, artificial intelligence, to analyze the data they have to make predictions about their customers' next step. The two main objectives of predictive analytics are regression and classification. In general, it can be used to identify risks and opportunities (*Kavya. V & Arumugan. S, 2016*). The prediction process can be divided into four steps: 1) collect and pre-process raw data; 2) transform pre-processed data into a form that can be easily handled by selected machine learning method; 3) create a learning model using the transformed data; 4) report predictions to the user using the previously related learning model. Usually, predictive analytics is bundled with data mining and machine learning, which are used for the extraction of obscure or hidden predictive information (*Nishchol Mishra & Sanjay Silakari, 2012*). It is fair to say predictive analytics plays an important role in helping marketers to know what a consumer's behavior will likely be even before they make the decision.

4.4.2 Machine Learning

The machine learning (ML) method enables computers to learn from data without human intervention or assistance. And it can cope with large datasets which cannot be analyzed by human or human driven analytics tools. Moreover, it can discover the underlying connections, association rules, or patterns which are too complex for human to identify. ML tasks are performed by ML algorithms. An ML algorithm is a set of iterative computational procedures which are designed to find connections, rules, patterns, or anomalies in data. Using iterative algorithm, ML method allows data analytic models to adapt to the variation of new data. ML algorithms can be classified into four kinds (*Fanjuan Shi, 2018*), as shown in figure 7 below.

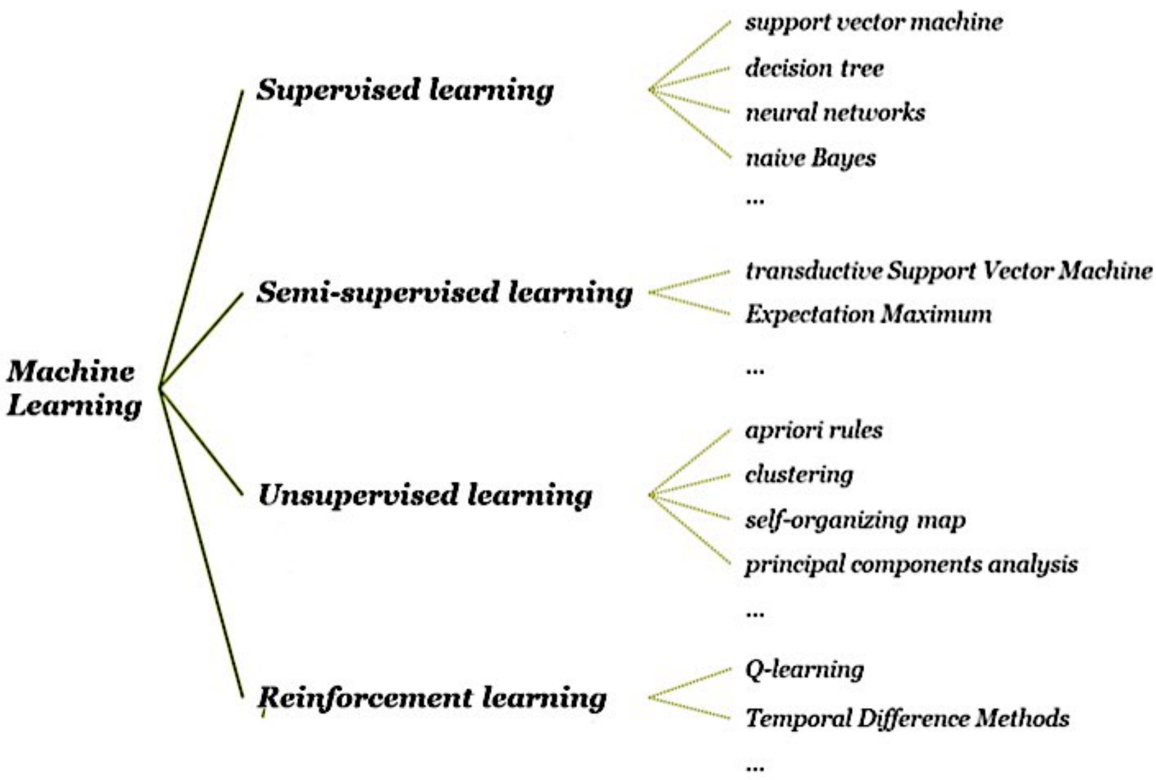


Figure 7: Machine Learning Classification. Resource: *Fanjuan Shi*

Supervised learning is used when a dataset contains examples whose inputs and the corresponding outputs are known to the machines learning algorithms, and when the objective of learning is to find the association rule that connects the inputs and the outputs (*Kotsiantis et al., 2007*). However, sometimes acquiring outputs for data points can be either an expensive and difficult task. In such cases, semi-supervised learning is better for predicting the missing outputs for the data points. Unsupervised learning is used when datasets contain nothing but unlabeled data points, and is helpful for finding some underlying structure or pattern within (*Hastie et al., 2009*).

4.5 RTB Ecosystem

RTB is also known as programmatic buying. It uses per impression context and targets the ads to specific people based on data about them, and hence dramatically increases the effectiveness of display advertising. Before the emergence of RTB in 2009, the display advertising market was mainly divided by premium contracts (since 1994), which took more than 40% of impressions, and ad networks (since 1996), which took the rest of impression that were usually referred to as remnant (*Shuai Yuan et al., 2013*).

However, RTB has experienced an explosive growth since its birth. On the international markets, it is reported that 88% percent of North-American advertisers have switched to RTB when buying ad impressions in 2011. In China, the RTB market starts from the TANX system from taobao.com in 2011. It is reported that in 2013, the amount of RTB ad requests in China has reached 5 billion. European RTB adoption lags that in the US but is now growing fast, led by the UK (*Yong Yuan et al., 2014*). Generally speaking, RTB has taken over a large amount of online impression market. With increasing players (DSP, SSP, DMP, AdX, Ad Network, etc) joining the game, the RTB Ecosystem becomes much more complex than before. In chapter 2, the basic knowledge about RTB has been introduced. In this section, problems and findings from history researched will be discussed. The table below summarized the main literature selected.

Table 9 Summary of selected main literature on RTB

No.	Citation	Year	Main Topic
1	Ye Chen, Pavel Berkhin	2011	Performance-based display ad allocation
2	Hamid Nazarzadeh et al.	2012	Dynamic Pay-per-auction mechanism
3	Shuai Yuan, Jun Wang, Xiaoxue Zhao	2013	RTB for online advertising
4	Yong Yuan et al.	2014	RTB advrtising
5	Weinan Zhang, Shuai Yuan, Jun Wang	2014	Optimal RTB for display advertising
6	Shalinda Adikari, Kaushik Dutta	2015	RTB in online advertising
7	Jun Wang, Shuai Yuan, Weinan Zhang	2016	Mechanism and algorithm

Though RTB has accelerated the speed to deliver relatively tailored ads, there are a lot of limitations as well. The pricing models used in RTB are mostly CPC, PPA, CPM. A drawback of the CPC scheme is that it requires the advertiser to submit their bids before observing the profits generated by the users clicking on their ads. In contrast, PPA allows the advertisers to report their payoff after observing the users' action, therefore eliminates the uncertainty of advertisers and reduces their exposure and computational burden (H*amid Nazerzadeh et al., 2012*). Since advertisers have no control over the inventories or users, it is more difficult to deploy goal-driven campaigns than branding ones (*Shuai Yuan et al., 2013*). As stated in *Shalinda Adikari & Kaushik Dutta*'s study, the complexity and dynamic nature in the RTB process make it difficult to apply forecasting strategies effectively and efficiently. Due to the long industrial chain with various kinds of economic entities, the RTB market is shown to be highly dynamic and far from stabilization and standardization. Apart from the complexity issue, the inherent competitions among the entities in RTB markets might cause information be distorted or hidden when passing from one entity to another, which is the so-called asymmetrical information issue.

A central issue in the performance of display advertising is matching campaigns to ad impressions, which can be formulated as a constrained optimization problem that maximizes revenue subject to constrains such as budget limits and inventory availability (*Ye Chen & Pavel Berkhin, 2011*). The auction mechanisms can influence the performance to a large extent. There are several researches on the bidding algorithm. *Chen & Berkhin* proposed a real-time bidding algorithm that enables fine-grained impression valuation, and adjusts value-based bids according to real-time constraint snapshots. In another study done by *Zhang et al.*, it was thought that fundamental technical challenge is to automate the bidding process based on budget, the campaign objective and various information gathered in both runtime and in the past.

The existing delivery failed to integrate the individual-level behavior analysis and the system-level strategy optimization. In *Yong Yuan et al.*'s survey on RTB, they thought the proposed bidding algorithms do not take into consideration the heterogeneity and diversity of advertisers' behavior. In addition to bidding strategies, budget allocation is also a key decision for advertisers.

Chapter 5 Gap Survey

As described earlier, the gap survey is designed for showing the gap between consumers' expectations and perceptions concerning their online experiences. Since the trigger of conducting this research is from consumers' complaining about the content they received online, it is essential to present how big this gap is, what consumers are mostly unsatisfied about, and what consumers expect to be better. This chapter can answer all these questions.

The questionnaire was distributed online in April, 2018 and ended in May, 2018. It has been sent to the researcher's friends, classmates, and various social network groups. The target sample is people who has received personalized online ads. As the sample covers people aged from 18-40, it represent a diverse online users population. For the people who participated in this survey, everybody was explained what this survey is about before they took the survey. Likert scales ranging from 0 to 10 were used to measure all statements in the survey. To ensure content validity, it has been sent to a few people for testing to get some suggestions on refining the statements before start collecting data. There are a total of 148 valid responses out of 176 replies obtained for the final data analysis. Among these responses, 28 were excluded because of missing or inappropriate data. The respondents were informed of the purpose of the survey and provided with an explanation of the research. The respondents ages ranged from 18-40, and a majority were female (71.62%). All of them have online shopping experiences. The table below illustrates the detail demographic data for this survey.

Table 10 Demographic data of the gap survey (N=148)

Variables		N	%
Gender	Male	42	28.38%
	Female	106	71.62%
Age	18-22	34	22.97%
	23-30	110	74.32%
	31-40	4	2.70%
	>=41	0	0.00%
Online Shopping Frequency	Never	0	0.00%
	Once or less/week	65	44.30%
	2-3 times/week	59	39.60%

	4-6 times/week	14	9.40%
	Everyday	10	6.71%
Receive Online Ads Platform	E-mail	70	47.30%
	Social Media Platform	82	55.41%
	Search Engine Webpage	78	52.70%
	Others(TV, Phones, Messages,etc.)	52	35.14%

5.1 Survey background

The design of this survey is based on SERVQUAL, a multiple-item scale for measuring consumer perceptions of service quality. The quality of online ads or online experiences is unlike goods quality, which can be measured objectively by such indicators as durability and number of defects. Service quality is an abstract and elusive construct because of three features unique to services: intangibility, heterogeneity, and inseparability of production and consumption (*A. Parasurman & Valarie A. Zeithaml, 1988*). Online ads is a type of service marketers offer to customers to provide online shopping instruction. Instead of in-store consult, marketers can guide visitors online to find the items they need. This kind of service is indeed intangible, heterogeneous and always comes with an item (products or services). Therefore, this method can be perfectly used on measuring online ads quality.

Perceived quality is the consumer's judgement about an entity's overall excellence or superiority (*Zeithaml, 1987*). It is a form of attitude, related but not equivalent to satisfaction, and results from a comparison of expectations with perceptions of performance. It is a global judgement, or attitude, relating to the superiority of the service, whereas satisfaction is related to a specific transaction. Expectations are viewed as predictions made by consumers about what is likely to happen during an impending transaction or exchange (*A. Parasurman & Valarie A. Zeithaml, 1988*). In online advertisement, expectations are viewed as desires or wants of consumers, i.e., what they feel an ad provider should offer rather than would offer.

5.2 Dimensions of Online Ads Quality

The original SERVQUAL's five dimensions are "Tangibles", "Reliability", "Responsiveness", "Assurance" and "Empathy". The definitions for each dimension are shown in table below:

Table 11 Definitions of SERVQUAL's dimensions

Dimension	Definition
Tangibles	Physical facilities, equipment, and appearance of personnel
Reliability	Ability to perform the promised service dependably and accurately
Responsiveness	Willingness to help customers and provide prompt service
Assurance	Knowledge and courtesy of employees and their ability to inspire trust and confidence
Empathy	Caring, individualised attention the firm provides its customers

For online ads quality, in this research, we combine SERVQUAL's dimension setting and the conception of real-time marketing. However, there is still no official definition for real-time marketing, here we adopted the definition from Evergage, a professional real-time personalization platform. A real-time personalization should **send the right content to the right person at right time through right channel with good manners** *(Rob Carpenter, 2014)***.** These five dimensions can therefore be the standard to measure an online ads quality. Combined with the dimensions SERVQUAL method proposed, the clear definition of online advertisement quality dimensions are explained below:

Table 11 Definitions of online advertisement quality dimensions

Dimension	Definition
Person	Personalisation at the 1:1 level
Content	Apply appropriate algorithms to determine the most relevant recommendations
Time	Proper time, timestamps, and frequency cap
Journey	Track full journey instead of single clickstream
Manner	Follow the legal regulation, respect consumers privacy

According to the five dimensions above, 10 statements are designed in coordination with these five aspects. With statements cover all dimensions, we can get a clear view of which aspects have larger gap and which dimensions have smaller gap. The table below shows how the 10 statements of the survey are in coordination with these five dimensions. Table 13 presented the statements have been designed.

Table 13 Survey statements in coordination with online ads quality dimensions

Person	Content	Time	Journey	Manner
S5	S4	S3	S1	S8
S6	S7	S9	S2	
S10				

Table 14 Expectation Statements
E1: Online advertisements should update to your current stage in the purchase process. (Your purchase processes might go through "Awareness-Consideration-Decision-Delivery-Use-Loyalty")
E2: Advertisers shouldn't send you advertisements with products you bought not long time ago.
E3: Online advertisements shouldn't be repetitive and redundant.
E4: Advertisers shouldn't send you ads which you even didn't give one second of attention for the first time you saw it. (e.g. the email you never open; the ads you skip in one second)
E5: Online advertisements should take into account your past favoured brands.
E6: Online advertisements should take personal interests into account. (Interests can be generated from your history searching on a search engine and web pages, online purchase history, etc.)
E7: Online advertisements should take shopping context into account. (Location, State of mind; Intention, Budget, Shopping interface, etc)
E8: Advertisers should reach prospects through right methods instead of sending contents to people without consent. (Receivers are well informed of the receiving)
E9: Advertisers shouldn't send you same content through all platforms you went.
E10: Advertisers should take into account the online behaviours during users searching. (Behaviours like clicks, downloads, shares, time and times viewed, etc.)

Table 15 Perception Statements
P1: Online advertisements you received have updated to your current stage in the purchase process. (Your purchase processes might go through "Awareness-Consideration-Decision-Delivery-Use-Loyalty")
P2: The advertisements you received were not with products you bought not long time ago.
P3: Online advertisements you received were not repetitive and redundant. Extremely Disagree Neutral
P4: You did not keep receiving same ads which you even didn't give one second of attention for the first time you saw it. (e.g. the email you never open; the ads you skip in one second)
P5: Online advertisements have taken into account your past favoured brands.
P6: Online advertisements have taken personal interests into account. (Interests can be generated from your history searching on a search engine and web pages, online purchase history, etc.)
P7: Online advertisements have take shopping context into account. (Location, State of mind; Intention, Budget, Shopping interface, etc)
P8: Advertisers did reach you through right methods instead of sending contents to you without consent. (Receivers are well informed of the receiving)
P9: Advertisers did not send you same contents through all platforms you went.
P10: Advertisers have taken into account the online behaviours during users searching. (Behaviours like clicks, downloads, shares, time and times viewed, etc.)

5.3 Results Discussion

Before giving their scores, respondents were asked to give their opinion of the overall impression of their online experience, agree means that they do feel the same way as the questions described. Whilst disagree, on the other hand, claims that they do not feel it the same way at all. Between agree and disagree, respondents can also choose somewhat agree, somewhat disagree, which is less extreme. Or they could choose neutral if they had neither not good nor not bad experiences. Seeing from the general attitude about online advertisement performance ratings, as shown in table 16 below, it is apparent that on average, people hold a "somewhat agree" and "neutral" opinion. It is fair to say consumers indeed can get some help from online advertisements, however, due to the non-transparency, users are unsure about what information they have gave out, and what marketers did with their data. Therefore, the concern of privacy also led to the distrust, making them surf online with great care. According to the mean value of each question result, there's almost half people stand on the neutral to disagree side, which means half of them do not feel the ads they received are highly tailored for them or relevant to their interest. We find that more people feel that the ads they received have helped them making decision during online shopping, which indicate people are positive about the recommendation. And it is something people expect from online ads as well. The data in table 16 and table 17 can show most people hold a neutral or somewhat agree attitude towards online advertisements performance. It can indicate people do aware of the ads they received are specially for them, but low agree rates tell the personalization is not high quality.

Table 16 General attitude about online advertisements performance

Questions	Mean	Std Deviation
1. Do you feel online ads you received are highly tailored for you?	2.68	1.08
2. Do you think online ads you received are relevant to your interest/digital footprints?	2.49	1.10
3. Do you think online ads you received have helped your final decision on shopping?	3.06	1.28

Table 17 Results of general attitude towards online advertisements performance

No.	Agree	Somewhat Agree	Neutral	Somewhat Disagree	Disagree
1	12.75%	34.23%	32.21%	14.09%	6.71%
2	18.67%	38.67%	22.67%	15.33%	4.67%
3	14.77%	16.78%	33.56%	17.45%	17.45%

Table 18 Results towards expectation and perception statements.

No.	Mean E	Mean P	Mean E-P	Std Deviation E	Std Deviation P
1	6.63	6.06	0.58	2.476	2.587
2	7.60	5.28	2.30	2.401	2.903
3	8.32	4.81	3.44	2.082	2.744
4	7.67	5.05	2.63	2.315	2.734
5	7.32	6.36	1.03	2.008	2.254
6	7.44	6.62	0.91	2.302	2.352
7	7.31	5.83	1.51	2.135	2.620
8	8.27	4.72	3.60	2.046	3.182
9	8.02	5.12	2.89	2.171	2.666
10	7.46	5.54	2.06	2.264	2.451
Avg	7.314	5.539	2.095		

Table 18 above gives mean value and standard deviation value of the sample. For the expectation and perception questions, people were asked to rate the extent to which they agree with the statements, from 1 (extremely disagree) to 10 (extremely agree). What we can get from the table is there is gap between each statement concerning the sample. For the expectation statements, consumers gave higher scores, all the mean values for expectation is higher than the mean values for perception. In general, the average mean value for expectation is 7.314, which indicates consumers high expectations. Only statement 1's mean value is lower than average, which means the purchasing stage has lower weight concerning the effect on online ads quality. Number 5 and 7's mean values are comparatively low as well, it means compare to other factors, brands and context are not the most important points for them. On the contrary, consumers expect innovative recommendations instead of an old-fashioned "serve me what I have searched". E3 got the highest mean score among the ten statements. Unsurprisingly, P3 got the lowest mean score compared with others. It is clear from this that the repeated same ads occurrence is the most severe problem.

For all perception results, mostly are under 6 except number 1, 5 and 6. Number 1 and 5 also got scores lower than 7.5 on expectation. Therefore, the gap for these two statements are relatively low. The highest score in perception statement is number 6. It implies consumers are aware of their searching histories being tracked and used on the ads they received. Generally speaking, people's expectation on online advertisement quality is higher than their actual experience. The average gap is 2.095, which indicates that there is indeed a gap existed among the survey sample. Except for direct observation of the data, it is necessary to do a significance test to show how significant the difference is. The terms “significance level” or “level of significance” refer to the likelihood that the difference found by the random sample chosen in this research is not representative of the population, and thus is not based on chance. Significance levels most commonly used in educational research are .05 and .01 levels (.05 is be used in this study). These numbers and signs come from significance test, which begins with the null hypothesis, assumes there are no structural differences in the data. If the chance found through a test is smaller than the chosen significance level, we assume the null hypothesis is incorrect and that there are structural differences.

Concerning this survey, the question needed to be answered is “Is the difference between the means of two samples different (significant) enough to say that there is gap existed between people’s expectation and perception about online experiences?”. A paired sample t-test can help answering this question. This test is appropriate because we are looking at two responses (expectation and perception) grouped together per person. In order to conduct the t-test, IBM SPSS has been used to get the result. The P value is a probability, with a value ranging from zero to one, that answers this question: In an experiment of this size, if the populations really have the same mean, what is the probability of observing at least as large a difference between sample means as was, in fact, observed. If the P value is equal or greater than .05, it is likely to be a result of a chance, the difference is not significant. On the contrary, if this number is less than .05, the difference is significant. From the results shown in the table 19, it is clear that all the P values are less than .05, which means the difference between expectation and perception is statistically significant. In another word, the difference found by this random sample chosen in this survey is representative.

Table 19 Significance of the Test Results[1]

No.	N	Mean Difference Between E&P	Std Deviation Difference	t	P Value
1	125	0.58	2.812	2.290	0.024
2	125	2.30	3.520	7.318	<0.001
3	124	3.44	3.511	10.895	<0.001
4	124	2.63	3.389	8.638	<0.001
5	124	1.03	2.337	4.918	<0.001
6	123	0.91	2.939	3.436	0.001
7	125	1.51	2.928	5.773	<0.001
8	124	3.60	3.886	10.329	<0.001
9	125	2.89	3.504	9.215	<0.001
10	125	2.06	2.986	7.729	<0.001

[1] Due to some respondents miss one or more questions in this survey, in order to ensure the preciseness, the N number used to calculate the significance is slightly different than the N number used on demographic data (there are 148 people answered all the demographic questions but not all the expectations and perceptions questions).

Chapter 6 Expert Interviews

6.1 Interview Introduction

Semi-structured interviews were held with experts who have experience in this field. The interviews were used to uncover the underlying factors causing online customers' low engagement. The benefits of a semi-structured in-depth interview are that open questions enable more creative answers and collect a variety of opinions. Due to the fact that technology keeps changing, in order to give up-to-date findings, it is necessary to get opinions from people who work in this field at this moment.

The researcher has conducted three interviews. Three interviews are conducted at June, July and August in sequence. Before an interview, ten to twenty neutral, open-ended interview questions were prepared. However, flexibility is essential when taking semi-structured interviews. Various aspects of each research question might emerge as a consequence of the conversation. Thus, not every questions prepared have been asked and there were some extra questions during the interview as well. Moreover, taken into consideration that experts are adept at different branches of industry, pre-research about candidate's background and experiences were conducted before the official interview. The questions prepared for each expert focus on different dimensions in coordination with their experiences as well. But in general, all questions relate to five dimensions, namely data management, legal regulation, personalization approach, technology and organization.

These interviews were analyzed by using the a coding tool named Nvivo. The result of these interviews can help the researcher connect the knowledges from this study and literature review with the real situation for companies. Therefore, we can come to a conclusion of factors that might influence the online ads quality. The findings also need to be validated to show its reliability. With the results from these interviews, an assessment has been designed and sent to 10 experts who are experienced with digital marketing. The objective of the expert interviews is to gather a list of features that they deem influential to online ads quality. Questions from five dimensions: Data management, Legal regulation, Personalization approach, Technology and Organization, which were chosen beforehand based on the research of existing studies concerning these five dimensions. The definition and descriptions for each of the dimensions are illustrated in the table below.

Table 20 Dimensions description

Dimension	Description
Data management	Data management refers to every action marketers/advertisers, publishers or companies do with the data stream consumers had online. It includes questions “what data can be collected?”, “How can the data be collected?”, “What can marketers do with the data they gathered?” “How is the data stored?”, “How does data syncing from various sites work?”, etc,. Companies should be clear about the questions above in order to make full use of data.
Legal regulation	Nowadays, the technique developed for data collection can be very powerful. However, privacy is a very sensitive topic when talking about personalization. More people started using an ad blocker or browser blocker to avoid being tracked. Though it does not directly cause loss for marketers, it is still not the best way to deal with it. What consumers need is transparency and the ability to decide what marketers can do with their data. With legal regulation, companies have to respect their customer’s privacy and be clear about their processing of the data. The enforcement of GDPR in European is a start to protect people’s rights, and also force companies transfer to a higher level personalization.
Personalisation Approach	The personalized content online can be a recommendation from websites or applications on an own domain. Usually the content was generated based on users’ search history or what other people also searched. The most common algorithms are content-based personalization, customer profile approach and collaborative filtering. Most companies use a hybrid approach which uses multi-approaches. Another type of personalized content is sent on social media platforms. The delivery usually goes through a RTB system, and the offers usually were made by matching advertisers requirements with the personal profile the third party has. For the other options like SEM, email marketing, the mechanism behind is the same, item-oriented. Some also take into consideration the person’s profile. However, these approaches are not able to consider the real-time context. In another word, the real-time customer journey.
Technology	Technologies like machine learning, predictive analytics, multi-channel synchronization, etc, have been developed for years. But the application on personalization is not mature yet.
Organisation	Organizations can be very decisive concerning the online marketing strategy. Without resources, budget, and people, organizations can do nothing with the data and technologies they have. Moreover, the complex structure of a company, the collaboration between different departments, and even the culture of a company can influence their marketing strategy.

6.2 Interview Processes

6.2.1 Candidate Selection

The objective of expert interviews is to gather professional opinions about possible reasons for online ads low relevancy. The methods companies use to track and analysis customers are different because the maturity level of each company is different as well. Therefore, the candidate's expert field and experiences can be influential to the results. The results from the expert interviews will only be reliable if experts are chosen carefully since it is the experts' opinions that determine the qualifications of the final research findings.

In order to gather a full view of the problem, it is necessary to choose experts adept in different areas. An expert's profile was built for the purpose of identifying candidates expertise areas, disciplines and skills, as well as the position within an organization that related to the topic. The table below displays the profiles of experts interviewed.

Table 21 Knowledge resource nomination worksheet

No	Position	Expertise Areas	Disciplines/Skills	Industry
1	CTO	Data Science; Legal regulation; Technology; Organization	Data analysis; Programmatic advertising; RTB knowledge; Machine learning; Online tracking.	Information Tech Start-up specifically serve for flight advertising
2	Business Development and Account Director	Data Science Legal Regulation; Personalization; Omni-channel marketing; Technology; Organization.	Customer relationship management; Customer journey tracking; Advanced analytics; Marketing strategy;	Middle size computer software company specifically for customer engagement management
3	Global Online Insight Manager	SEM; Social marketing; Display advertising; Data-driven marketing; Marketing intelligence; Legal regulation; Organization	Data analysis; Customer relationship management;	Large size International Human Resource company

Obviously, these three experts are all experienced with online marketing. But each of them has their own specialities. The first expert is a CTO of a start-up tech company. The company was founded in 2008. With 10 years development and evolution, they are specialized on audience targeting, performance display advertising, data activation, and demand intelligence. The core purpose of them is helping companies reach the right consumers to offer their products or guiding an airline to fill empty seats. Generally speaking, they play a role as a mediator to help airline companies do precision marketing and resource saving. The benefit of an interview with a CTO from a startup can represent most SME's situation. The landscape of digital marketing is unprecedentedly complicated with continuous new players enter and old players leave. The organizations' maturity level, company size, structure, resources, etc, are similar to the startup I interviewed. Thus, this interview can be representative for most small size tech companies in the market.

The second interview was with an expert from a software company focus on high-level customer journey analytics. The mission of the company is to help their customers by providing intelligent technology that harness the power of AI that enables them to build engagement at scale, creating valuable and long lasting customer relationship. As a business development and account director, the interviewee has comprehensive knowledge among online data, personalization approach, technologies and has the insight to address what problems exist from an organization's point of view. The interviewee has more insights on how the maturity of a company can influence their customer engagement, and how to achieve a high level of personalization. The second interview provides insights on all the five dimensions set on 6.1.

The last interview candidate comes from a bigger size company comparatively. The expert is proficient in search engine marketing, social marketing, display marketing, etc,. He also has been working on transferring web data into insights to provide visitors better online experiences. The company the candidate works for is an large international human resource company. With business around the world, the company structure is much bigger than previous two interviews. Thus, it is much harder for them to change, due to the fact that a small change might involve a lot of stakeholders. Moreover, as an international company, the data from their site and the channels they need to synchronize can make their work more complicated than SME.

As a whole, the three candidates can represent what most companies stand for, what obstacles they have, and what are they trying to do in the future.

6.2.2 Data Processing

The research data were collected through three interviews. Each interview has been recorded and later transcribed in order to minimize possible loss of information. When all transcripts were finished, they were analyzed using a coding frame. All codes coming from the interviews and transcripts were brought together in one single coding frame. This coding frame is based on the five dimensions talked about in 6.1. The Nvivo software has been used for the coding. Nvivo is the most powerful software for gaining richer insights from qualitative and mixed-methods data. It can store and organize all the data in one platform, from quantifiable demographic information to qualitative open-ended questions and interviews (*from nvivo.com*).The benefits of using a software for coding are it can help you store all data in one platform and generate reports more easier.

These five dimensions are the top level node. Within the five dimensions, a child node can be generated depends on the type of information. At the sideline of the interview transcriptions, the kind of sub code which categorized this main topic and piece of interview text has been defined in an annotation. After that all the collected codes have been reported in a clarifying scheme, which is easy to recognize. When using the coding frame, important concepts can be identified in the data, such as customer journey, cross-sites tracking, collaborative filtering, etc,. Finally, the findings were linked to each other and to the existing literature. The results of this coding can contribute to the design and content of the final assessment. In the following content, important insights from the coding are presented and analyzed. The results are segmented by the five dimensions.

6.2.3 Data Collection

Data Management

As stated in the dimension description in table 20, data management refers to every action marketers/advertisers, publishers or companies do with the data stream consumers had online. Questions concerning tracking methods, tracking limitations, data types, data sources, data storage, data integration, and things influencing the collection of data have been raised during three interviews. With the help of Nvivo software, 18 insights from experts related to data management were addressed in table 22. The 18 insights were grouped into 7 types of issues (sub-node). The sub-nodes type were based on the content of each insight.

Table 22 Features for data collection dimension

No	Sub-node	Descriptions
1	Accessibility	In the large number of views, we might get little less data, because of the ad blockers.
2	Accessibility	Although if you go to a social platform like Twitter, Facebook, etc. The data is always limited. That's also to do with the fact that people are not always recognizable with the same ID.
3	Accessibility	You can only do personalization based on the reaction on your site, your own stores. A little bit outside domain, but that's not social oriented.
4	Accessibility	An ad being blocked doesn't cause any immediate loss for us, but it's just traffic that we can not buy.
5	Accessibility	If browsers or Apps say 'disable all third party cookies', it's very hard for us to do businesses.
6	Accessibility	There's ad blockers that diminish you, and browser and vendors are more aggressive to be a third party. Also, the favor is going to switching to one-stop shop, like Facebook, Google, Amazon, where they control both supply and demand.
7	Availablity	The challenge is more domains that are out of your control, which makes it sometimes difficult.
8	Computing	There's not enough time to prepare your knowledge around a user until it becomes available for you to show them ads.
9	Collection Methods	For 90%, cookies is what we used to collect data. JavaScript beacon paired with any mix pixel. If javascript is disabled, although this is very small segment, there are mobile SDKs, and third party data.
10	Collection Methods	If we do our collection of data ourselves, certain information is available in the request itself. And certain information we could get by third party. DMPs, etc.
11	Data Governance	DMPs, they receive events from the clients that they work with, typically publishers, and then they group people in segments, and synchronize that information through parties in anonymize way.
12	Data Integration	We do track both clicks and eventually conversions. It is an integrate part of delivering a campaign, most of all, optimizing it.
13	Data Integration	The problem is that you need to combine all the data
14	Data Integration	I think the challenge is to bring the candidates that we have from different platforms to be able to act on a right way.
15	Data Storage	Only if you return within a limited amount of time, then the data can be valuable. Otherwise, cookies are cleaned, then you have to start over again.
16	Reliability	What you see is that the technology is there to store every data you want.
17	Usefulness	We don't usually work with names, addresses, etc,. It is not useful for us to do marketing on such kind of information. We are more interested in the behaviour
18	Usefulness	What's important is that you look at stages that customers going through. The activities they do online and offline.

Legal regulation

The sub-nodes of legal regulation were designed based on who is being influenced. In coordination with the content extracted from interview transcripts relate to legal regulation, the 10 opinions from experts were divided into three types, namely influence on consumers, influence on industry and influence on organizations. Due to the fact that the interviews were done soon after the enforcement of GDPR, there are more thoughts from experts concerning its influence.

Table 23 Features for legal regulation dimension

No	Sub-node	Description
1	Influence on Consumers	The enforcement of GDPR is causing more harm to the publisher rather than the advertiser. For us, it will be less traffic to buy, but to the publishers, the content will not be monetized.
2	Influence on Consumers	There's more transparency for consumer about what's been stored, that's good.
3	Influence on Consumers	It doesn't change much, except that the user is under control weather they want to be tracked or not.
4	Influence on Industry	I don't see it as the death of personalisation because of GDPR. I think it will give you more transparency. It will help people to get insights on how personalisation work.
5	Influence on Industry	Main things slightly strict, but we were already fairly regulated about what we can collect, and how much and how long. Everything is more efficient after GDPR.
6	Influence on Industry	I would say it also hurts the usability as a whole. If every website that you want to go, you need accept the cookie, the policies and the consent. It hurts the publisher, therefore a lot publisher choose to change their business model.
7	Influence on Orgnizations	I think it will help organizations getting better and better. Because companies now really start thinking about how to get on a higher maturity level.
8	Influence on Organizations	The impact on a company from legal regulation depends on how transparent the data was before.
9	Influences on Organizations	As we are forced to operate within the confirms of the law, cookies need to be cleaned after some time of activity, and we can not process any personal identifier.
10	Influence on Organizations	We have a lot of old data, but we cannot use it to train models due to legal regulation.

Personalization Approach

In order to understand why the contents delivered to consumers online are not highly personalized, it is important to know the mechanism of personalization, the algorithms used on generating the tailored contents, the principle and strategy adopted to choose the personalization approach, moreover, to know how the delivered content be monitored and improved.

Table 24 Features for Personalization Approach Dimension

No	Sub-node	Personalisation Approach
1	Algorithm	Collaborative filtering is in the area of finding similar items. It is not very smart. But it can generate revenue. It’s very risky, but a lot companies are still earning money with it.
2	Algorithm	The algorithms is twofold. One is based on business rules, so based on marketing information and knowledge we got in there. And then you have the advanced analytics, and those automated analyses over this visualization of the journeys.
3	Algorithm	The Boolean rule: if they’ve seen A, show them A, perhaps B and C, if they’ve seen B, recommend D.
4	Algorithm	The algorithm that delivers is in a constant feedback loop that gets improved by the interactions with that campaign.
5	Algorithm	We have used the collaborating filtering for product recommendation based on the users behavior we find from similar users and offer the product that they have seen.
6	Algorithm	Collaborative filtering is a flat analytics that has been used by Amazon already for many years. It is purely used on product.
7	Mechanism	What you often see is completely product related
8	Mechanism	The product related way is not a smart way of communication.
9	Mechanism	We are going to prescriptive instead of predictive. Predictive is you are saying they are going to buy this product. Prescriptive is I am going to tell you which parts you are going to take even before you know it.
10	Mechanism	Personalized content goes twofold, there’s the business rules approach, where some customers really want to have their hand on how products we commented. And business rule is something very very simplistic. It’s a boolean tree, if this happens, show this. If user bought this, up-sell this.
11	Monitor	We usually have a very strict frequency cap. We have a low amount usually between 3 and 5 times that we might show an ad. And also paste through out a day.
12	Principle	If the user has expressed their interest, in a certain product, it is very highly likely that we will retarget them.

No	Sub-node	Personalisation Approach
13	Principle	When we start a new campaign, and we define the budgets and the goals, the pricing strategies we want to follow, the click is a very important measure of optimizing how aggressively and how the pricing will be affected.
14	Strategy	Retargeting is when people who express their interest on a product, and we try to find them back, and entice then to buy that product. This is much more smaller pool than branding campaign.

Technology

There are new hot tech-words continuously appearing, and online marketing industry rely heavily on technology. With the trend of a more customer-centric service, technologies are designed for understanding customers as well. The most talked technologies such as predictive analytics, ML, and customer journey tracking have been mentioned during the interviews.

Table 25 Features for Technology Dimension

No	Sub-node	Technology
1	Advanced Analytics	Automated analysis over the visualization of the journeys.
2	Automotive Marketing	The future for us is not service the customers, but giving them the software so they can do everything themselves.
3	Clustering	In other occasions we were also successful on finding their interest and often offering them product from that cluster.
4	Customer Journey Tracking	What we’re doing here is using intent and behavior across different touchpoints that gives you insight on what are going to happen in the next step.
5	Data Storage	What you see is that the technology is there to store all the data you want.
6	Machine Learning	There are some companies that try to find out whether a user is the same within different devices by employing machine learning.
7	Machine Learning	The trend is toward the machine learning, to have as much automation and influence from the data itself.
8	Predictive Analysing	Predictive is you are saying they are going to buy this product. Prescriptive is I am going to tell you which parts you are going to take.

Organization

When we look at the landscape, issues from organizations are easily forgotten. However, there are so many influences on an organization's ability to do personalization. An organization's structure, hierarchy, maturity, diversity, etc, could influence the marketing strategy directly. Uncontrollable factors like resources and budget are limited. The key performance indicator (KPI) that need to be achieved are also hard to break due to several reasons. In order to survive, companies need be be careful with their KPI.

Table 26 Features for Organization Dimension

No	Sub-node	Organization
1	Budget	They don't want to know customers due to a small budget. And they can also not use all the touchpoint.
2	Compatibility	How they build the website was by buying an IT system, which you couldn't change very easily. What our company have is a lot of IT systems around, which were not possible to change.
3	Conflict	The product manager is going for products. And in the end they do not care about the customer.
4	Diversity	Deliver what type of campaign to deliver is something the advertisers themselves need to decide. We can not necessary decide for them. It is their business model.
5	Hierarchy	Big companies have difficulty in business model. And that also impact on people working at specific domain.
6	KPI	They are very often focused on one element within company, the KPI
7	KPI	What you often see is a gap between every KPI.
8	KPI	I think the problem is they are trying to get KPIs done
9	Maturity	If you are going to look at the overall picture, what is missing over here is the maturity level
10	People	If you have the right IT environment, there's also need for people who are able to get the data and use it to see what relation there are.
11	Resources	Companies need to know the budget, the channels, everything that is limited.
12	Structure	What you get is that there's organizational structure that is not inlined with a high-level framework.
13	Structure	There's a big gap between their operation execution and what they really have.

6.3 Result Discussion

The data collection section has extracted the most representative insights from the interview transcripts, which can represent different types of problems. In this section, the analysis goes deeper into the five dimensions. Among the three interviews, the first interviewee (CTO) talked more about the data management, the personalization approach and the relationship between different players in this industry. The second interviewee (Business Development and Account Director) focused more on the organization's dilemma between choosing profit or customer. The third interviewee (Global Online Insight Manager) gave a better view on what the issues would be for big size companies. In general, they all had gave their comments concerning different dimensions. Table 27 summarized the dimensions that have been mentioned during each interview.

Table 27 Dimensions been discussed by experts

Dimensions	Expert A	Expert B	Expert C
Data Collection	YES	NO	YES
Personalisation Approach	YES	YES	NO
Legal Regulations	YES	Somewhat	YES
Technology	Somewhat	YES	YES
Organization	Somewhat	YES	YES

Concerning the data management dimension, all experts have expressed they believe data collection and processing is not a big problem to achieve higher level personalization. The technology is there to store all the data. However, there are some other difficulties caused by uncontrollable factors. For example, some data is difficult or impossible to access. Factors like legal regulation, organizations' maturity level, ad blockers and cross-domain issues make it hard get information from a full angle. Other than the accessibility problem, organizations' scale would also influence the data processing ability. Due to the increasing of channels, data integration is one of the most difficult issues that's troubling marketers. All three interviewees have mentioned data integration is crucial for understanding a user's online behavior. Other problems like data governance, data storage, reliability and usefulness are not as essential as previous issues. But it will still influence the performance more or less.

Legal regulation has caught people's attention in recent years due to the enforcement of GDPR. Experts attitude towards arising attention on it is positive. They all think it will bring more opportunity than loss to companies. For consumers, there is more transparency. Consumers can decide what they want and can take control of their own data. For organizations, it might hurt to some extent, but it depends on how transparent the data was before for that company. Some companies might need to clean all old data, which would cause the quality of online ads to drop sharply at first. But in the long-run, it is good for them to adjust their business model and choose the right prospects to target.

Personalization approach and technology are usually closely connected. Advanced analytics are used to realize hyper-personalized ads. In order to do personalization, techniques like customer journey analytics, predictive analytics, and machine learning can largely help marketers understand consumers in a more efficient way. However, how technology can be adopted and put into practice is another saying. Usually, big size tech companies are more capable on investing in new technologies. Moreover, new technologies need long term scientific research which would take a lot of money and resources while there is little result. Most retail companies are still using old-fashioned personalization algorithms like collaborative filtering, consumer profiling or content-based personalization. The similarity is they are all product based instead of journey based, which is the main reason for non-smart recommendations. Two experts have mentioned personalized content usually goes into two fold, the business rule based and advanced analytics. What we see most is the simple boolean rule which show consumers what they have searched. In summary, personalization approach and technology can decide the quality of ads directly, but it takes time, money and effort for organizations to make full use of them.

All experts agreed that organization is decisive on marketing strategy. Though organization factors do not cause direct impact on the quality of an ad, it has the largest influence on all other four dimensions. A company's position of strength decides what they are capable of. The maturity level decides a company's structure, hierarchy, and diversity. The more mature a company is, the more resources, people, and budget a company have. On the contrary, the less mature a company, the less compatibility due to complex systems and relations. The KPI issue has been discussed a lot during the interviews, very often organizations need to consider the benefit they can get, which is also one of the biggest obstacle need to overcome. In general, organization should be considered in the first place when talking about the factors that influence online ads quality.

Chapter 7 Validation Assessment

7.1 Introduction

The assessment was designed in coordination with the results from expert interview. There are 30 questions grouped into five parts. The principle to design the assessment is to validate the findings from previous research steps. Each participant was asked to give a score to the statements from 1 to 10 based on how influential the problem is to the quality of online ads. The distribution of this survey was also executed through Qualtrics (Qualtrics survey software was launched in 2002 as a way for academics to carry out sophisticated research, it is a tool used to design, send and analyze surveys online. It's the primary method of collecting feedback at scale whether that's a simple questionnaire or a detailed study), and distributed to 10 experts who have related working experiences and knowledge. The detail questions can be found on Appendix V. This survey was distributed from 12 August to 12 September, there are 10 experts who participated in the end. One of the experts also been interviewed in the expert interview stage. All experts are working in different positions, but they all have experience on digital marketing and advertising. The information about each expert's position and industry are stated in table 28.

Table 28 Validation Assessment Experts Panel

	Positions	Industry
1	Product Owner Market Intelligence	Information Technology
2	Marketing Analyst	Marketing and Advertising
3	CTO	Information Technology and Services
4	Online Marketer	Publisher
5	Lead Developer	Information Technology and Services
6	Content Marketer	Media
7	Data and Research Director	Software
8	Customer Success Director	Retail
9	Online Marketing Manager	Retail
10	Product Manager	Advertising

In the validation process, participators were asked to first rate the five dimensions on a scale of one to five. The results are shown in table 29. What we can see is that organization dimension got the highest influencing on online ads quality, while data management got the lowest. Moreover, the Std

Deviation for organization is the lowest as well, which means expert opinions are more centralized compared with other four.

Table 29 Dimension Assessment Result

Dimension	Mean	Std Deviation
Data Management	3.1	1.45
Legal Regulation	3.8	0.98
Personalization Approach	4	0.77
Technology	4.1	0.83
Organization	4.8	0.66

7.2 Results Discussion

As we see from table 30, column 4 gives the mean value of each statement. Organization still received the highest score ranging from 7.4 to 8.5. Because the statements for each dimension are varied, the best way to compare the mean value for each dimensions is using the total average. Since the score ranges from 1 to 10, in order to compare the difference between the dimension scores, the average mean value used in table 30 is the total average value for all statements in one dimension divided by 2. Through this way, we can see the difference between dimension level and problem level (all statements in the assessment are asking how influential on online ads quality that problem is). As a whole, this two values have the same feature. Data management got the lowest score and organization received the highest. However, the technology dimension got relatively low mean value compared with the dimension mean value. What we can conclude here is when is comes to problems caused by technology, it has less influence than experts thought at first. Concerning the mean value within dimension, there are some irregular numbers as well. For data management, No.1 got pretty low value while question 9 in the survey's value is pretty high. With this result, we can say data availability does not have large impact on online advertising, and on the contrary, data integration is a more difficult problem to overcome. No.12 from legal regulation received a relatively low score compared with the other two. It can show that giving consumers more control over their data does not have a big influence on data management. The scores for personalization approach are comparatively centralized, which means it plays an important role on increasing online ads quality but it is not the most decisive factor. Not surprisingly, all the scores for organization are pretty high. It is reliable to say organization influences largely on the performance of online advertising.

Table 30 Result for Validation Assessment

Dimension	Dimension Mean	Avg./2	Mean	Std Deviation	No	Feature
Data Management	3.1	2.759	3.88	1.83	1	Availability
			5	1.48	2	Collection Methods
			5.9	1.45	3	Computing
			5.7	1.68	4	Accessibility
			5.7	1.27	5	Usefulness
			5.7	2.05	6	Data Storage
			5.7	2.19	7	Reliability
			4.7	1.68	8	Data Governance
			7.3	1.27	9	Data Integration
			5.6	1.11	10	Reliability
Legal Regulation	3.8	2.858	6.4	1.69	11	Industry
			4.25	1.09	12	Consumers
			6.5	1.32	13	Organizations
Personalization Approach	4	3.35	6.8	1.83	14	Strategy
			7.7	0.78	15	Algorithm
			5.3	0.9	16	Principle
			6.9	0.7	17	Mechanism
			7.1	1.22	18	Monitor
			6.4	1.2	19	Algorithm
Technology	4.1	2.973	5.89	0.99	20	Advanced Analytics
			6.4	1.74	21	Practicability
			5.2	1.94	22	Advanced Analytics
			6.3	0.9	23	Practicability
Organization	4.8	3.886	8.1	1.14	24	Maturity
			8.1	0.7	25	Structure
			7.5	0.67	26	People
			7.3	0.9	27	KPI
			8.5	1.02	28	Resources
			7.4	1.28	29	Compatibility
			7.5	0.81	30	Hierarchy

Chapter 8 Conclusions

This research project was performed to investigate and measure the factors that influence online advertisements performance from a comprehensive view. There is a gap between consumers' expectation and perception concerning the quality of online ads. Therefore, it is necessary to find out what causing the gap. In another word, what factors make it hard to achieve consumers expectation. The answer to the research question and sub-questions are provided in the following section. Limitations and future research are also illustrated in this chapter in section 8.2.

8.1 Answers to Research Questions

After the entire research processes, the research sub-questions can be answered here:

RQ1 What does personalization mean for consumers?

The gap survey done in this study can answer this question. From the result analysis of the survey, it is clear that the gap between consumers' expectation and perception is significant. When we look at the scores for each statements, we can know what consumers prefer and what they don't like. Take into consideration of the highest 5 scores for expectation statements, we can conclude consumers expect less repetitive and more innovative contents. Moreover, transparency and contextual are important for consumers as well. In brief, personalization means smart, innovative, transparent and unobtrusive for consumers.

RQ2 What dimensions decide an online advertisement's relevance level?

Literature review and expert interview help answering this question. During the literature review process, researcher found that most literature concerning digital marketing can be grouped into the following dimensions: data management, personalization approach and RTB Ecosystem. Technologies are involved in most literature more or less. Legal regulation for online advertising is less studied comparatively. After the expert interview, organization has been discussed a lot compared with other dimensions. Combined with all the information and how influential of each dimension, five dimensions were selected in the end, namely data management, legal regulation, personalization approach, technology and organization.

RQ3 What problems existed in different dimensions?

Both expert interview and the validation assessment contribute to the answers of this question. Based on the opinions from three experts, it is believed that good performance comes form a combination of factors. For the data management dimension, even though the technology is mature enough to gather necessary data, the collection process is influenced by other factors like ad blockers, cross-domain issues, data storage, data governance and reliability issues. From assessment result, data integration is the most difficult problems. Concerning legal regulation dimension, it is agreed that it indeed limit the ability to gather more information about consumers, but it can improve online ads quality while the transparency between companies and consumers goes up as well. The personalization approach and technology companies use nowadays are old-fashioned as a whole. It is not the technologies are not mature yet, but the adoption of new things are influenced by various factors (budget, resource, compatibility, etc.). There are a lot of problems in organization that are hard to overcome. The scores for organization problems in the final assessment are pretty high as well. Resources, structure and maturity level got relatively high score among others.

After answering the above three questions, we are able to answer the main research question for this study: ***What factors influence the personalization level of online ads?*** The answers to the three research questions can form the answer of the main question. RQ1 indicate what should personalization be like, and what problems existed from consumers view. RQ2 gave clear guidance to the research direction and scope. RQ3 helps us arrive the answer to the main research question. The table 31 gives the answer to the main research question. It is validated that all the five dimensions are influential. And organization is the most influential one. Concerning the factors within each dimension, the table listed the factors that get comparatively higher scores compared with others (each question in the assessment is in coordination with one of the factors). Due to the fact that each organization's situation is different, the number of each factors do not represent the how influential that factor is.

Table 31 Factors that influence the personalization level of online ads.

Dimension	Factors	
Data Management	1. Accessibility	2. Availability
	3. Data Governance	4. Data Integration
	5. Data Storage	6. Reliability
	7. Usefulness	

Legal Regulation	1. Limitation on Data Collection	2. Requirements on Data Transparency
	3. Influence on Organization Business Model	
Personalization Approach	1. Algorithm	2. Mechanism
	3. Principle	
Technology	1. Advanced level	2. Practicability
Organization	1. Budget	2. Compatibility
	3. Diversity	4. Hierarchy
	5. KPI	6. Maturity
	7. People	8. Resource
	9. Structure	

8.2 Limitations and Future Research

Like most empirical studies, there are some limitations of this study. First of all, due to the broadness of this research topic, there are a lot of knowledge involved, which makes it hard to find similar research literature. Instead, this study did literature reviews on relevant topics in order to find problems that existed in different areas. Lacking literature on similar research causing less support on the contribution of this research. Secondly, as the opinions collected from experts are based on each expert's personal experience. The position and industry of each expert might also influence their attitude towards the factors that influence online ads. Bias is a problem of this study which possibly cause measurement errors that misleading conclusions. Third, gap survey of this research was distributed through social media which may lead to participant's bias and limit the diversity of the sample. Last but not least, in terms of the scope of this study, the five dimensions were selected based on history study field and expert opinion, which means there is also risk of bias.

Regarding the limitation described above, there are several directions to which our research can be extended. Since this research is the first study on finding the factors for online advertisements' low performance, the resource and time are limited. If this research can be conducted with more experts involved, the result can be more validity. Besides, the factors that influence the relevancy of an ad also varies from different industry and organizations. Future research can narrow down the research scope to just one industry or organization, which is easier to do the research with specific background.

References

Acar, G., Eubank, C., Englelardt, S. & Juarez, M. (2014, Nov.). The web never forgets: Persistent tracking mechanisms in the wild. *2014 ACM SIGSAC Conference on Computer and Communications Security*, 674-689.

Adikari, S. & Dutta, K. (2015). Real Time Bidding in Online Digital Advertisement. *B. Donnellan et al. (Eds.): DESRIST 2015, LNCS 9073*, 19–38. DOI: 10.1007/978-3-319-18714-3_2.

Aguirre, E. Mahr, D., Grewal, D., Ruyter, K. & Wetzels (2015). Unraveling the personalization paradox: The effect of information collection and trust- building strategies on online advertisement effectiveness. *Journal of Retailing*, 91(1), pp. 34- 49. doi: 10.1016/j.jretai. 2014.09.005.

Alshenqeeti, H. (2014, March 31). Interviewing as data collection method: A critical review. *English Linguistics Research*, Vol. 3, No. 1; 2014.

Ayenson, M. D., Wambach, D. J., Soltani, A., Good, N. & Hoofnagle, C. J. (2011). Flash cookies and privacy ii: Now with html5 and etag respawning. *World Wide Web Internet And Web Information Systems.*

Bashir, M. A., Arshad, S., Kirda, E., Robertson, W. & Wilson, C. (2018). How tracking companies circumvent ad blocker using websockets. *Northeastern University.*

Berg, B. L. (2007). Qualitative research methods for the social sciences. London: Pearson.

Blaxter, L., Hughes, C. & Tight, M. (2006). How to Research. (3rd Ed.) New York: McGraw-Hill Education.

Bradlow, E. T., Gangwar, M., Kopalle, P. & Voleti, S. (2017). The Role of Big Data and Predictive Analytics in Retailing. *Journal of Retailing* 93 (1, 2017), pp. 79–95.

Breuer, R. & Brettel, M. (2012). Short- and Long-term effects of online advertising: Differences between new and existing customers. *Journal of Interactive Marketing* 26 (2012) 155-166.

Bryman, A. & Bell, E. (2007). Business Research Methods, 2nd ed., Oxford: Oxford University Press.

Carpenter, R. (2014, June 15). Real-time marketing isn't what you think it is. Retrieved from https://www.evergage.com/blog/real-time-marketing-isnt-what-you-think-it-is/

Chaffey, D. (2018, Aug 19). Average display ads click through rate. Retrieved from https://www.smartinsights.com/internet-advertising/internet-advertising-analytics/display-advertising-clickthrough-rates/

Champiri, Z. D., Shahamiri, S. R. & Salim, S. S. B. (2014, Sep. 28). A systematic review of of scholar context-aware recommender systems. *Expert Systems with Applications* 42 (2015) 1743–1758.

Chen, Y. & Berkhin, P. (2011). Real-time bidding algorithms for performance-based display ad collocation. *Microsoft research*.

Clark & Dorie (2012, November 11). The End of the Expert: Why No One in Marketing Knows What They're Doing. Retrieved from https://web.archive.org/web/20131104235646/http://www.forbes.com/sites/dorieclark/2012/11/11/the-end-of-the-expert-why-no-one-in-marketing-knows-what-theyre-doing/

Cookie types. Retrieved from http://www.aboutcookies.org.uk/cookie-types.

Cook, K. (2018, April 26). *A brief history of online advertising*. Retrieved from https://blog.hubspot.com/marketing/history-of-online-advertising

Cortland, M. (2017, Feb 1). 2017 Adblock Report. PageFair. Retrieved from https://pagefair.com/blog/2017/adblockreport/

Darmer, P. (1995). Qualitative interviews. In: Darmer, P. and Freytag, P. V. (eds), (1995). *Business Research Methodology, Lund: Student Literature*, pp. 252-272.

Dörnyei, Z. (2007). Research Methods in Applied Linguistics: Quantitative Qualitative, and Mixed Methodologies. Oxford: Oxford University Press.

Ebbert, J. (2012, Nov. 27). What is real-time bidding? Retrieved from https://adexchanger.com/online-advertising/real-time-bidding/

Eirinaki. M & Vazirgiannis, M. (2013, Feb 1). Web mining of web personalization. *ACM Transactions on Internet Technology* (2003): 1-27. doi:10.1145/643477.643478.

Englehardt. S. & Narayanan, A.. Online tracking: A 1-million-site measurement and analysis. *Proceedings of 2016 ACM SIGSAC CCS '16*.

Estrada-Jimenez, J., Parra-Amau, J., Rodriguez-Hoyos, A. & Forne, J. (2017). Online advertising: Analysis of privacy threats and protection approaches. *Computer Communications* 100 (2017) 32-51.

Følstad, A. & Kvale, K. (2018). Customer journeys: a systematic literature review. *Journal of service theory and practice*. https://doi.org/10.1108/JSTP-11-2014-0261

Foss & Clive (1991). Beacon. *In Kazhdan, Alexander. The Oxford Dictionary of Byzantium. Oxford University Press.* pp. 273–274. ISBN 978-0-19-504652-6.

Gasiorek, A. (2014, Sep 25). What are beacons and what are they do? Retrieved from https://kontakt.io/blog/infographic-beacons/

Ghauri, P., Grønhaug, K. & Kristianslund, I. (1995). Research methods in business studies: A practical study, New York: Prentice Hall.

Gironda, J. T. & Korgaonkar, P. K. (2011). iSpy? Tailored versus invasive ads and consumers' perceptions of personalized advertising. *Electronic Commerce Research and Applications* 29 (2011) 64-77.

Green. J. (2014). Real-time bidding and programmatic progress. Retrieved from https://retargeter.com/blog/real-time-bidding-and-programmatic-progress/

Hannak, A., Sapiezynski, P., Molavi Kakhki, A. & Lazer, D. (2017, June 15). Measuring personalization of web search. *National Science, CNS-1054233*.

Heaton, R. (2017). How does online tracking actually work. Retrieved from https://robertheaton.com/2017/11/20/how-does-online-tracking-actually-work/

Hirsch, D. D. (2011). Protecting the Inner Environment: What Privacy Regulation Can Learn from Environmental Law, 41 GA. L. REV. 1, pp. 28–30.

Jankowski, Jarosław & Kazienko, Przemysław & Wątróbski, Jarosław & Tomaszewska, Anna & Ziemba, Paweł & Zioło, Magdalena(2016). Fuzzy Multi-objective Modeling of Effectiveness and User Experience in Online Advertising. *Expert Systems with Applications*. 65. 10.1016/j.eswa.2016.08.049.

Jarvinen, J. & Taminen, H. (2015, Aug 5). Harnessing marketing automation for B2B content marketing. *Industry Marketing Management*, 54 (2016) 164–175.

Johnson, L. (2016, July 26). IAB Study says 26% of desktop users turn on ad blockers. Retrieved from http://www.adweek.com/digital/iab-study-says-26-desktop-users-turn-ad-blockers-172665/.

Kaptein, M., Poulos, P. M., Ruyter, B. & Aarts, E. (2015). Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles. *Int J.Human-computer Studies* 77 (2015) 38-51.

Karaj, A., Macbeth, S., Berson, R. & Pujol, J. M. (2014, April 24). WhoTracks.me: Monitoring the online tracking landscape at scale. *International World Wide Web Conference*, http://dx.doi.org/10.1145/2872427.2883028.

Kavya. V & Arumugam. S (2016). A review of predictive analytics in data mining. *International Journal of Chaos, Control, Modelling and Simulation (IJCCMS)* Vol.5, No.1/2/3.

Kim, C., Kwon, K. & Chang, W. (2011, April). How to measure the effectiveness of online advertising in online marketplaces. *Journal Expert Systems with Applications Volume 38 Issue 4*, Pages 4234-4243.

Kwon, K. & Kim, C. (2012). How to design personalization in a context of customer retention: Who personalizes what and to what extent?. *Electronic Commerce Research and Applications* 11 (2012) 101-116.

Lewis, P. (2016, Aug 19). How Beacons Can Reshape Retail Marketing – Think with Google. Retrieved from https://www.thinkwithgoogle.com/articles/retail-marketing-beacon-technology.html

Li, C. (2016). When does web-based personalization really work? The distinction between actual personalization and perceived personalization. *Computers in Human Behavior* 54 (2016) 25-33.

Li, C. & Liu, J. (2017). A name along is not enough: A reexamination of web-based personalization effect. *Computers in Human Behavior 72,* Pages 152-139.

Li, J., Yuan, Y., Zhao, X. & Wang, F. (2017). Research on information structure of programmatic advertising markets. *IFAC PapersOnline* 50-1 (2017) 13587-13592.

Li, M., Sagl, G., Mburu, L. & Fan, H. (2016). A contextualized and personalized model to predict user interest using location-based social networks. *Computers Environment and Urban Systems* 58 (2016) 97-105.

Li, W. & Huang, Z.Y. (2016). The Research of Influence Factors of Online Behavioral Advertising Avoidance. *American Journal of Indus-trial and Business Management*, 6, 947- 957.

Litan, R.E. (2011). Law and Policy in the Age of the Internet, 50 Duke L. J. 1045, 1045.

Lord, J. (2011, April 5). Retargeting: What it is and how to use it. Retrieved from https://moz.com/blog/retargeting-basics-what-it-is-how-to-use-it

Lu, Y., Kruger, R., Thom, D., Wang, F., Koch, S., Ertl, T. & Maciejewski, R. (2014). Integrating Predictive Analytics and Social Media. *IEEE*.

Marvin, G. (2015, December 23). What is an Ad Network. Retrieved from https://martechtoday.com/martech-landscape-what-is-an-ad-network-157618

Mayer, J. R. & Mitchell, J. C. (2012). Third-party web tracking: Policy and Technology. *2012 IEEE Symposium on Security and Privacy*, DOI 10.1109/SP.2012.47.

McCambley, J. (2018, March 14). *The first ever banner ad: why did it work so well?* Retrieved from www.theguardian.com

Miroglio, B., Zeber, D., Kaye, J. & Weiss, R. (2018, April 23). The effect of ad blocking on user engagement with the web. *In proceeding of the 2018 World Wide Web Conference*, page 813-821.

Mishra, N. & Silakari, S. (2012). Predictive Analytics: A Survey, Trends, Applications, Oppurtunities & Challenges. *(IJCSIT) International Journal of Computer Science and Information Technologies*, Vol. 3 (3) , 2012, 4434- 4438.

Montgomery, A. & Smith, M. D. (2008, July 1). Prospects for personalization on the Internet. *Carnegie Mellon University Research Showcase.*

Morris, S. (2018, May 15). Definition of tracking pixel. Retrieved from https://skillcrush.com/2012/07/19/tracking-pixel/

Namdev, N., Agrawal, S. & Sikari, S. (2015). Recent advancement in machine learning based internet traffic classification. *Procedia Computer Science* 60 (2015) 784-791.

Nazerzadeh, Hamid & Saberi, Amin & Vohra, Rakesh (2007). Dynamic Pay-Per-Action Mechanisms and Applications to Online Advertising. *Northwestern University, Center for Mathematical Studies in Economics and Management Science, Discussion Papers*. 61. 10.1145/1367497.1367522.

Newsome, C., Wylie, G., Greene, M., Neubert, O., McKinnon, F., Appello, M., Cohen, N. & Reale, E. (July 2014). Programmatic trading. Retrieved from https://www.iabeurope.eu/files/8614/0776/0957/IAB_Europe_Programmatic_Trading_White_Paper_July_2014_v2.pdf

O' Brien, C. (2017). The evolution of digital marketing. Retrieved from https://digitalmarketinginstitute.com/en-eu/the-insider/05-10-16-the-evolution-of-digital-marketing-30-years-in-the-past-and-future

Papadopoulos, P., Kourtellis, N. & Markatos, E. P. (2018, May 26). Cookie synchronization: Everything you wanted to know but afraid to ask. *In Proceedings of the 27th International Conference on World Wide Web*, WWW'18, 2018.

Qin, R., Ni, X., Yuan, Y., Li, J. & Wang, F. (2017). Revenue models for demand side platforms in real time bidding advertising. *Systems Man and Cybernetics (SMC) 2017 IEEE International Conference*, 438-443.

Sahni, S., Navdeep & Narayanan, Sridhar & Kalyanam, Kirthi. (2017). An Experimental Investigation of the Effects of Retargeted Advertising: The Role of Frequency and Timing. SSRN Electronic Journal. 10.2139/ssrn.2852484.

Salonen, V. & Karjaluoto, H. (2016). Web personalization: The state of the art and future avenues for research and practice. *Telematics and Informatics*, 33 (2016) 1088-1104.

Schmucker, N. (2011). Web tracking. Berlin University of Technology, SNET Seminar2 Paper.

Sheth, J. N. & Parvatiyar, A. (1995). The evolution of relationship marketing. *International Business Review* Vol.4, 397-418.

Shiller, B. , Waldfogel, J. and Ryan, J. (2018), The effect of ad blocking on website traffic and quality. The RAND Journal of Economics, 49: 43-63. doi:10.1111/1756-2171.12218

Shi, F. (2017). A Smart Marketing Personalization Methodology Based On Real-Time Contextualisation In The Era of Big Data. *Sous la direction de Chirine Ghedira, Catherine Pivot et Jean-Jack Cegarra, Université Jean Moulin (Lyon 3). Thèse soutenue le.* Retrieved from http://www.theses.fr/2017LYSE3019/document.

Shi, F., Ghedira, C. & Marini, J. (2015). Context Adaptation for Smart Recommender Systems," in *IT Professional*, vol. 17, no. 6, pp. 18-26. doi: 10.1109/MITP.2015.96

Singel, R. (2010, November 27), Web gives birth to banner ads. Retrieved from https://www.wired.com/2010/10/1027hotwired-banner-ads/

Smith, M. S. & Cong (2006). Research Serv., RL 31408, Internet Privacy: Overview and Legislation in the 109th Congress, 1st Session.

Soysal, M. (2018). Global e-commerce conversion rate. Segmentify. Retrieved from https://www.segmentify.com/about-segmentify/

Soysal, M. & Schmidt, E. G. (2010). Machine learning algorithms for accurate flow-based network traffic classification: Evaluation and comparison. *Performance Evaluation* 67 (2010) 461-467.

Statista (2018). Number of monthly active Facebook user worldwide 2008-2018. Retrieved from https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Stearn, J. (1999, Dec.). The 10 common myths of Cookies. *Elsevier science*.

Strauss, J. & Rogerson, K. S. (2002). *Policies for Online Privacy in the United States and the European Union*, 19 Telematics and Informatics 173, 188.

Teixuira, T. S. (2014, Jan 17). The rising cost of consumer attention: why you should care and what you can do about it. Harvard Business School.

Toubiana, V., Narayanan, A., Boneh, D., Nissenbaum, H. & Barocas, S. (2011) Adnostic: Privacy preserving targeted advertising. *NSF PORTIA and MURI PRESIDIO projects.*

Vesanen, J. (2005). What is Personalization? - A Literature Review and Framework. *Helsinki School of Economics Working Papers.*

Wahyuni, S. (2003). Strategic alliance development: a study on alliances between competing firms s.n.

Wang, J., Yuan, S., & Zhang, W. (2016). Real-time bidding based display advertising: Mechanisms and algorithms. *University College London, arXiv:1610.03013.*

Wang, W., Yang, L., Chen, Y. & Zhang, Q. (2015). A privacy-aware framework for targeted advertising. *Computer Networks* 79 (2015) 17-29.

Yuan, S. & Abidin, Z., Ahmad & Sloan, Marc & Wang, Jun (2012). Internet Advertising: An Interplay among Advertisers, Online Publishers, Ad Exchanges and Web Users. *Information Processing and Management*, arXiv:1206.1754.

Yuan, S., Wang, J. & Zhao, X. (2013, June 27). Real-time bidding for online advertising: Measurement and analysis. *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*, DOI 10.1145/2501040.2501980.

Yuan, Y., Wang, F., Li, J. & Qin, R. (2014). A Survey on Real Time Bidding Advertising. *2014 IEEE International Conference on Service Operations and Logistics, and Informatics,* 418-423. doi: 10.1109/SOLI.2014.6960761.

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1988). Communication and Control Processes in the Delivery of Service Quality. Journal of Marketing, 52(2), 2-22.

Zhang, J., Mao, E. (2016, March). From online motivations to ad clicks and behavioral intentions: An empirical study of consumer response to social media advertising. *Psychology & Marketing* Vol. 33(3) 155-164.

Zhang, W., Yuan, S. & Wang, J. (2014). Optimal real-time bidding for display advertising. *ACM SIGKDD international conference on Knowledge discovery and data mining*, 1077-1086, doi 10.1145/2623330.2623633.

Zhu, Y. & Chang, J. (2016). The key role of relevance in personalized advertisement: Examining its impact on perceptions of privacy invasion, self-awareness, and continuous use intentions. *Computers in Human Behavior* 65 (2016) 442-447.

Appendix I Expectation and Perception of Online Advertisements Performance Survey

This survey aims at analyzing the gap between customer's expectations and perceptions about the online advertisements. Today, when we talk about online advertisement, we expect a highly personalized offer due to the fact that marketers can collect information from our digital footprint(A digital footprint is a trail of data you create while using the Internet. It includes the websites you visit, emails you send, and information you submit to online services). By using this information, marketers are trying to optimize the accuracy of personalized ads. However, whether an advertisement is useful or not only depends on the people who see it. Thus, please feel free to show your attitude about online ads through this survey. Every sound is helpful for the researcher to better investigate this online advertisement imprecise phenomenon. Note: There're 5 general questions at first to collect basic background. All the information collected is kept anonymous and confidential, and will only be used for the purpose of this research. The main questions are divided into 2 part and will take a maximum of 10 minutes. The first part focuses on your expectation and the second focuses on your perception of online advertisement performance.

General Questions :

1. Age

18-22, 23-30, 31-40, 41+

2. Gender

Male Female

3. How often do you do shopping online on average?

Everyday; 4-6 times/week; 2-3 times/week once or less/week Never

4. On which platform have you received online ads based on your digital footprints? (You can choose one or more)

E-mail; Social Media Platform; Search Engine Webpage; Others (TV, Phone Message, etc.)

5. General attitude about online advertisement performance. (Agree-Somewhat Agree-Neutral-Somewhat Disagree-Disagree

Do you feel online ads you received are highly tailored for you

Do you think online ads you received are relevant to your interest/digital footprints?

Do you think online ads you received have helped you to make your final decision on buying something.

Expectation Questions: The following Expectation set of 10 statements relating to your feeling to which extent you think online advertisements should possess the features described by each statement. If you strongly agree that online advertisements should possess a feature, circle the number 10. If you strongly disagree that online advertisement should possess a feature, circle 1. If your feelings are not strong, circle one of the numbers in the middle. There are no right or wrong answers, all we are interested in is a number that best shows your expectations about firms offering advertisements.

E1: Online advertisements should update to your current stage in the purchase process. (Your purchase processes might go through "Awareness-Consideration-Decision-Delivery-Use-Loyalty")
E2: Advertisers shouldn't send you advertisements with products you bought not long time ago.
E3: Online advertisements shouldn't be repetitive and redundant.
E4: Advertisers shouldn't send you ads which you even didn't give one second of attention for the first time you saw it. (e.g. the email you never open; the ads you skip in one second)
E5: Online advertisements should take into account your past favored brands.
E6: Online advertisements should take personal interests into account. (Interests can be generated from your history searching on a search engine and web pages, online purchase history, etc.)
E7: Online advertisements should take shopping context into account. (Location, State of mind; Intention, Budget, Shopping interface, etc)
E8: Advertisers should reach prospects through right methods instead of sending contents to people without consent. (Receivers are well informed of the receiving)
E9: Advertisers shouldn't send you same content through all platforms you went.
E10: Advertisers should take into account the online behaviors during users searching. (Behaviors like clicks, downloads, shares, time and times viewed, etc.)

Perceptions Questions: The following Perception set of 12 statements related to your feelings about the online advertisement. For each statement, please show the extent to which you believe online advertisement has the feature described. Once again, circling 10 means that you strongly agree that online advertisement has that feature, and circling a 1 means that you strongly disagree. You may circle any of the numbers in the middle that show how strong your feelings are. There are no right or wrong answers, all we are interested in is a number that best shows your perceptions about the online advertisement.

P1: Online advertisements you received have updated to your current stage in the purchase process. (Your purchase processes might go through “Awareness-Consideration-Decision-Delivery-Use-Loyalty")

P2: The advertisements you received were not with products you bought not long time ago.

P3: Online advertisements you received were not repetitive and redundant. Extremely Disagree Neutral

P4: You did not keep receiving same ads which you even didn’t give one second of attention for the first time you saw it. (e.g. the email you never open; the ads you skip in one second)

P5: Online advertisements have taken into account your past favored brands.

P6: Online advertisements have taken personal interests into account. (Interests can be generated from your history searching on a search engine and web pages, online purchase history, etc.)

P7:

P8: Advertisers did reach you through right methods instead of sending contents to you without consent. (Receivers are well informed of the receiving)

P9: Advertisers did not send you same contents through all platforms you went.

P10: Advertisers have taken into account the online behaviors during users searching. (Behaviors like clicks, downloads, shares, time and times viewed, etc.)

Appendix II Online Advertisements Quality Interview with Company A’s Expert

Due to the fact that there is a huge gap between people’s expectation and perception about online advertisements quality, thus, it means there is necessity to researching on finding out “what factors influence the quality of an online ad”. In order to understand the mechanism behind screen, an interview with expert from either DSP, SSP, Ad Network or Ad Exchange, who plays a role between advertisers and publishers, can contribute to the research’s further work. The design of this interview is aimed at understanding the mechanism behind the delivery of online advertisements. In another word, the relationship between advertisers and publishers. With knowledge about how they work with each other, how a content been produced, how an ad delivered to which group, it can help the researcher better design the experiments for following work.

Company A is a software as a service company bringing efficiency and increased profitability to aviation industry with its marketing automation technology. Their business include help companies track customer behavior, gather information from all sources, integrate third party data. Moreover, using technologies to understand audience, visualize demand spoilage and spillage, track objectives in real time. General speaking, Company A plays an role to help companies to connect with customers by targeting specific users across various channels in real time and offer personalized recommendations. The interview with expert from Company A can contribute on understanding the how a DSP help advertisers target customers, what technologies they used, what are their biggest obstacles, etc.

Part 1-【General Questions】

Q1: Can you briefly introduce what do you mainly work on at Company A?

Q2: How do you think about today's ROI of online advertising? (E.g. PPC, CTR, CPM, etc.)

Q3: What factors do you think influence online advertisement's ROI ? (E.g. Ad blocker, incomplete online users information, complex customer journey, data integration, cross-platform tracking, etc.)

Part 2-【Questions About Data】

Q4: As a DSP, can you tell me what obstacles make it hard to produce and deliver tailored ads for targeted people? (E.g. Ad blocker, Data, Advertising giant, etc.)

Q5: What is the biggest challenges to collect data from different channels? (E.g. data integration, cross-platform tracking, cooperation, privacy, etc.)

Q6: What tools can you use to collect data? (E.g. cookie, beacon, pixel, etc.)

Q7: What sources will you use to gather data/insight? (E.g. Third-party, ad network, SSP, Publisher, etc.)

Q8: What kind of data do you collect and analyze? (E.g. Location, search history, email address, etc.)

Q9: How do you gain insight from the data/customer behavior collected online?

Q10: Does the enforcement of GDPR influence the data collection since the data exchange regulation becomes much more strict than before?

Part 3-【Questions about Personalization】

Q11: How do you think about the quality of today's online advertisement? In another word, the relevancy level of online advertisements. (Dimension: Time, Content, Person, Medium, Privacy).

Q12: What kind of customer behavior will you track? (Two type of customer behavior: Searching behavior, Interaction behavior. E.g. Keywords searching, Save, Add to cart, Delete, Impression, Click, Search, Purchase, Download, Register, Log in, Abandon,)

Q13: What customer behavior will make them be recognized as prospects?

Q14: How do you use people's history online data to produce personalized recommendation?

Q15: Can you introduce the technology/algorithm used to decide the personalized content?

Part 4- 【Questions About Real Time Marketing】

Q16: How do you perform dynamic adjustments at the right time?

Part 5- 【Questions About Future Development】

Q17: What do you think can be utilized to improve online advertisements quality? (E.g. Technology, Cooperation, Consumer awareness, etc.)

Part 6- 【Questions About Cooperation】

Q18: How do you think about the landscape of online marketing? (DSP, SSP, Ad Network, Publishers, Advertisers, DMP, etc)

Q19: As a DSP, how do you cooperate with other players?

Q20: How do you think about the landscape will be like in the future? What influence will GDPR bring to the agencies between advertisers and publishers?

Appendix III Online Advertisements Quality Interview with Company B's Expert

Part 1—General Questions

1. Can you tell me something about your work at Company B first?
2. How do you think about the landscape of online advertising today? Is the content online really personalized or relevant enough?
3. How do you think about the phenomenon of "follow me everywhere" ads once you did some search online? From my observation, consumers would receive same content in different platform for several times, which is one of the biggest reason for people who use ad blocker.
4. What issues you have been faced during all the projects you have done? Technology issues? Other non-technology issues?
5. What do you think will be improved in recent years?

Part 2—Main Questions

a). Data Management and Legal Regulation

1. This question is kind of broad, I have prepared few aspects that I think have issues. First of all,

the data problem. Data is the start of everything. However, it is the start of problems as well. Problems like multi-channel causing difficulties for data synchronization, legal regulation makes it harder for data collection, consumers frequently cookies cleaning causing trouble on user identification, ad blockers stop the way to reach customer data, the data can be collected are limited instead of a 360 degree view. All these problems making it uneasy to conduct further personalization. So my question is how do you think about these problems from data?

2. I noticed that Company B's tracking software provide bespoke solutions for companies. And one of the strength is the software has many flexible integration options for pull data from different systems. Can you tell me something about how Company B collect data from a relatively full angle. For example, a retail company using the software to track their customers, except data from their website, consumers might also visit their social media account, email subscription, other intermediaries (DSPs, SSPs, Ad Exchanges, Publishers). How will the software deal with data from different sources?
3. From my point of view, incomplete data, data synchronization, fiction data from accidental click or intermediaries are the biggest reasons existed at the first stage. How do you think about it? What else do you think concerning data issues?
4. We know consumers intention is continuously changing, how do you think the data you collected and analyzed can not indicate consumers real time intent?
5. Consumers are in permanent searching, which means not all searching history represent shopping intent, how to deal with this problem?
6. What influences have the enforcement of GDPR bring to online advertising?

b). Personalization

The second stage would be personalization, as far as I know, there are various mechanism for personalization in different systems. It might be a pre-designed marketing automation system using consumers profiles for recommendation or personalized ads. It might be just a boolean function conducted as "if...then...".

1. While for Company B, I saw the software's decision-making engine will take into consideration the consumer context, journey context, and the knowledge already knew about the consumer to choose the best conversation. Can you tell me something about the personalization approach used for more relevant content?

2. I have read personalization algorithms like consumer profile approach, content-based approach, collaborative filtering. But it still feels pretty vague to me. I know the personalization every organization used is different, but the result shouldn't be much difference since the data they can collect is the same. Usually, what kind of personalization approach will be adopted. Can you give me an examples about personalization algorithm. Or you can just briefly tell me something about the personalization mechanism.
3. From my own experience, I kind of feel the recommendation or personalized content for me is not smart enough. True experience is after I did a search of Tableau few month ago, I am keep receiving this ads until now. And another terrible example is people will receive ads with the products/services they just purchased. How do you think of this kind of problems?
4. I noticed that the software combines customers real-time behaviors with history data, I wonder how this personalization process works? Is it like the software detect behaviors indicate possible purchasing based on the algorithms designed, and then will trying to deliver a relevant content through RTB?
5. I know machine learning, predictive analytics, artificial intelligence , a lot technologies have been applied on personalization, can you tell me how these technologies influence the relevance?
6. As a whole, what do you think has been embedded for adopting personalization?

C). RTB Ecosystem

1. Concerning the delivery process, the RTB ecosystem is quite complicated. As far as I knew, there are some research on RTB algorithms optimization. Aimed at better resource allocation. However, due to the non-transparency, problems like DSPs knows the user information while advertisers do not know, which means DSPs might make decisions to maximize their own revenue instead of advertisers and consumers benefits, would also cause wrong delivery. How do you think about this kind of problems existed here?
2. The RTB nature needs to decide offer less than 120 milliseconds, do you think it is hard to practice highly complex and time consuming techniques for the decision making?
3. The inherent competitions among entities in RTB markets, information might be distorted or hidden when passing from one entity to another, causing the asymmetry information issues. Do you think this influences the relevance?
4. Except for RTB delivery approach, offline contract, is there any other delivery approach?

D). Organization

1. One thing I can not really find much information is the organization issues. Can you tell me something about how organizations influences consumer engagement? To what extent?

Except for the dimensions we discussed above, the data collection issues, the legal regulations, the personalization approach, the delivery method and organization issues. Do you have anything else in mind that you think influence consumers engagement?

Appendix IV Online Advertisements Quality Interview with Company C's Expert

1. First of all, can you briefly introduce what is your work mainly about?
2. What kind of data do you collect to understand visitors?
3. What challenges or problems have you met during the processes to collect data? (e.g. legal regulation, GDPR, ad blocker)
4. Do you have trouble with collecting data after the enforcement of GDPR?
5. What methods have you used to understand visitors digital footprints? I mean, after you get all the data, how will you make use of it? Concerning the work you do, it is not to provide personalized ads for customers, but to provide recommendation for visitors to help them find matching jobs. Can you give some examples of how you recognize a visitor a candidate?
6. What techniques do you use during the process to gain insights from online data? Machine learning? Predictive Analysis? Artificial Intelligence?
7. What problems do you think make it hard to understand visitors real-time intention?
8. What can do you to avoid prospects loss on the top of the funnel.
9. Actually I have go through Company C's website. What I experience is like normal job search platform. I can choose job type, locations, industries. I feel it is just like LinkedIn, can you tell me is there any difference compared with other job searching engine?
10. I notice that under each job introduction page, there's some recommendation, "others also search" something. What algorithm has applied on this recommendation result? User-based collaborative filtering approach? Do you think it works well? I feel it do helped to some extent. But it is not highly personalized.
11. Where else does the website offer personalized experience for visitors?

Appendix V Customer Engagement Influencing Factors Assessment
(This survey was distributed from 12 August to 12 September, 2018)

Part 1—Rating for General Dimensions

1. From 1-5, to which extent do you think the dimensions below influence the performance of online advertisements/customer engagement? (Data Management; Legal Regulation; Personalization Approach; Technology; Organization)
2. What is your comments concerning above five dimensions? Do you agree that the problem mainly comes from these dimensions? Please write your comments down.

Part 2—Rating for Problems Within Dimensions. On a scale from 1-10, to which extent do you think the statements below actually influence the performance of online advertisements?

a). Data Management

1. Sometimes, data are not available due to several reasons. E.g. Ad blockers, consent issues, search blocking, etc.
2. The personal profile building might not be accurate due to interference. E.g. users are in permanent searching, however, not every searched item indicates purchase intention.
3. Missing one or more of customer journey result in incomplete data. Insufficient customer journey tracking might cause difficulty on deciding the right time to serve personalized content for consumers.
4. The fact that a tracker can only track a consumer's behavior on their own sites, causing difficulty to get a complete view of consumer's behavior.
5. The fact that a consumer's intention keeps changing, without real-time tracking, the personalized content could be delayed.
6. DMP can provide large amounts of data, but the quality and usefulness of it is uncertain.
7. Data obtained through a business transaction (e.g. purchased data) have various problems, such as incoherence, grouped data, unreliability, etc.
8. Discontinuous tracking which miss one or more touchpoint lower the relevancy level.
9. Multi-channel touchpoints making data aggregation more difficult.
10. A complicated landscape means complex relationship between participators. However, under informational asymmetry and conflicts of interest, the low data/information transparency might lead to marketers making a wrong decision.

b). Legal Regulation

11. Regulation by law on the type of data can be collected, the way it should be used, stored and deleted limit the ability to get a broader view of customer behavior.
12. The fact that the data have been used by companies should be transparent to customers should be given full control over their data may cut marketers way to specific consumers.
13. The time of personal data can be stored by one party is regulated, which means history data may not able to be stored for a long-term analyze.

c). Personalization Approach

14. A big part of online ads are retargeting. However, it is an approach based on history search, which do not involve much analysis. This results in “follow me everywhere” contents and dissatisfaction.
15. Most recommendation algorithms are item oriented instead of customers oriented. As a result, a consumer might keep receiving content for a product that can be used for a long time after purchase, receiving ads for products with similar features, or what other people also searched. This item oriented mechanism limits the ability to understand a consumer's real needs.
16. The traditional "fishing net method" to reach as many people as possible lowers the level of personalization. Though the targeted people will be segmented into group with similar interest (Today’s branding campaign, campaign segmentation), the CTR(click through rate) is still quite low, which means it is not highly personalized.
17. Social media platform ads usually will be designed and delivered through RTB ecosystem. However, the quality of personalization is jagged. Moreover, information might be distorted or hidden when passing from one entity to another, causing issues with asymmetry information.
18. Lack of frequency capping (limiting the number of times that the same content is shown to one person in a specific period). This results in same ads showing too many times for a single person in a short period.
19. The existed bidding algorithms did not take into consideration of the heterogeneity and diversity of consumers’ behavior. Therefore, it fails to integrate the individual-level behavior analysis and the system-level strategy optimization.

d). Technology

20. How advanced the analytics tools are that the organization uses.
21. Due to the complex context, in order to do personalization, it is necessary to apply different technique on different cases and different people. However, it means more effort and investment, which bring less benefit than not doing it.
22. A lack of technically oriented person working on personalized ads. Some companies prefer outsourcing the online marketing to a third party, which is less expensive. However, outsourcing means there might be conflicts of interest between parties.
23. Large-scale computing takes more time, while the RTB (real-time bidding) nature should be done in short time (less than 120 milliseconds). The fact that large-scale computing can't keep up means higher level analysis for real-time decision making is more difficult.

e). Organization

24. The organization's maturity level, which determines their strategy on online marketing.
25. The complex structure of a company (do department work in silos or integrally). Complex structure means it is harder to adopt new technologies, IT systems.
26. The influence of the quality of cooperation between a company's IT and business department.
27. The fact that marketers have KPI target that they need to achieve. These can be stressful, and can force them to put KPI first instead of customers engagement.
28. The fact that most companies can only use a limited amount of resources. Big tech companies have much more power because of the large amount traffic. Therefore, a SME (small-middle enterprise) can either choose to cooperate with them or do it in-house with fewer resources.
29. The influence of a company's adaptability to new technologies and marketing strategy. For example, some companies take less time and effort adapt to newest technology while some companies are struggling with budget, resources, people, etc,.
30. The influence of the hierarchical structure of an organization. This can influence things like decision making, marketing strategy, budget, etc., and consequently influence the performance of their online advertisements.