H2020-MSCA-ITN-2018-813497 PROTECT

D2.1

Initial Ethics of Personalisation Case Studies and Literature Review Report

Submission Date: 30 September 2020

Dissemination Level: Confidential

Type: Other

Contributors: Ana Fernandez (VUB - ESR7), Paul Kuyer (DCU - ESR10), Karen Vázquez (UPM - ESR12), Rosalie Waelen (UT - ESR13), Michał Wieczorek (DCU - ESR11), Yashar Saghai (UT - WP Lead)

1 Introduction

New technological developments of the 21st century, such as Artificial Intelligence (AI), Machine Learning (ML) or Human-Computer Interaction (HCI), mainly focus on the optimisation of personalised applications and on accurate algorithmic decisions and prediction models for individuals (Zanker et al. 2019). Personalisation means "a process that changes the functionality, interface, information content, or distinctiveness of a system to increase its personal relevance to an individual" (Blom 2000, p.1). Different areas that use this personalised technology are recommender systems such as Netflix, Youtube or Spotify; web personalisation such as web analytics (Adobe, Google optimizer…) or Customer-Relation Management systems (Salesforce, HubSpot…); information filtering that are for example employed to filter the Internet with algorithms, or even personal e-mail filters, and adaptive texts such as screen readers or keyboard filters (Idem).

Recourse to personalisation, however, is not new and has been constantly used to sell and make products. An example has been exposed in an e-consultancy website, where in the 18th century a customer will arrive to the store and scream "hat" then leave in the carriage, and the retail's workmen will refer to the records of that customer of the hat size and style preference (McCaig 2013). According to the author, a customer-business relationship was always personal, however with the raise of chain shops and products, things started to get impersonal. The beginning of online shopping was similar, companies consider that all customers where the same.

What does the digital revolution add to traditional personalisation? It dramatically increases the level and scrutiny of personalisation. With increasing competition and high consumer expectation, there are multiple digital strategies where personalisation is often used (Fenech & Perkins 2015). 21st century developments have created important changes in our society (Boxever 2015), and with them the improvement and accuracy of personalisation. Its progress can be traced through marketing practices, such as customer-relationship management (CRM) developed in the 1990s, where email marketing campaigns were established to gather emails of individuals to send personalised advertisements. Initially, this implied the use of personal data with only name recognition, but today personalisation is used in emails and advertisements that can be sent based on the individual's activity online; for example, personalised advertisements are now based on your shopping habits, and Netflix recommendations are based from what you have watched in the past.

New analytic tools and technologies facilitate personalisation to a great extent. The introduction of cloud computing in 2007 allowed businesses to gather and analyse data in a much more efficient and accurate way. Moreover, thanks to ML and AI, computers can process huge quantities of data within milliseconds, and are also being used in marketing to improve relevance of products and services to consumers, while digital assistants like Siri and Alexa are offering tailored information and recommendations in our homes, making personalisation a household term and everyday occurrence. Consequently, the majority of businesses understand the usefulness of personalisation and the need to personalise their products and campaigns. Different technologies are increasingly facilitating high levels of personalisation, far from mere personalised advertisements, while many technologies and businesses include personalisation in their designs.

However, this is not exempt from criticism as the practice of personalisation raises ethical and legal issues that must be analysed. Questions arise concerning the limitations of personalisation practices and the protection of consumers against excessive, ubiquitous, or unwanted personalisation. And even in cases where personalisation is explicit, consensual and done as a good working practice, adverse effects can occur as the process is not always effective and can lead to mismatching services and products to users.

It is, therefore, necessary to study the ethical and legal implications of personalisation, in order to highlight and address the adverse effects it can have on our society. This work package consists of 12 case studies discussing different personalised technologies and their ethical and legal implications. Our goal is to emphasise the importance of personalisation today and the impact it will have in the future. We will do so by exhibiting a diverse range of technologies that utilise personalisation, showing how these technologies work and discussing which ethical and legal issues they raise.

1.1 Defining personalisation

Personalisation has a variety of forms, each of the twelve case studies presented here deals with a different instance of personalisation. In order to understand exactly how these different case studies are connected, we need a general definition of personalisation. Before we can formulate such a definition let us answer a few questions.

A first question to answer is: what should be the scope of the definition? As mentioned above, the personalisation of services, as a method of CRM, is already done for ages. When a waiter in a restaurant or the owner of a small shop know their customers, they can indeed personalise the service they offer to them. Or when you get a tailored-made suit, this can be considered a personalised product too. However, it is not these types of personalisation that we aim to address in this work package. What we are concerned with here is what one could call ‘digital personalisation’, that is, when the information on which personalisation is based was gathered automatically – for example personalisation based on data acquired via beacons, personal apps, wearables, browsing history, or video analytics. Hence, we need to formulate a definition that is sufficiently broad – that includes all possible types of digital personalisation – but at the same time not so general that it would include instances of non-digital personalisation too.

A second question is: what is the aim of personalisation? Different existing definitions of personalisation suggest different answers to this question. Possible options are to say that personalisation aims to increase the personal relevance of a system (Blom, 2000), make an experience more relevant (Kasanoff, 2001, p. 15; Kim, 2002, p. 30), create a user model or interface that represents a particular user (Brusilovsky & Maybury, 2002), or contribute to individuation (Suprenant & Solomon, 1987). Another option is to say that the aim of personalisation is to offer content, services (Hagen, 1999), and/or products (Salonen &

Karjaluoto, 2016), which are tailored, customized or simply “relevant” (Corp, 2002, p. 66) to individuals. This last answer seems to be most suiting. The ‘something’ that is personalised is always in some way a kind of content, service or product. Personalisation has the aim of adjusting, tailoring, or suiting different content, services or products to individuals. This can be either specific individuals (Fan & Poole, 2006) or categories of individuals (e.g. teens, sports fans, pregnant women, and so on).

A third and related question is: what is the motive behind personalisation? Of course companies do not personalise their products, services or content just for the sake of it. There can be several possible motives behind personalisation. For example, it can be done to improve the usability of an interface, to understand users or customers better (Karat, Marat & Ukelson, 2000), to improve a customer’s or user’s experience, to keep users or customers coming back for more, or to increase the efficiency of advertisement, pricing or product placement (Fan & Poole, 2006). The motives behind personalisation can differ, they often depend on the nature of the company or institution that offers personalised content, services or products. What is more, matching content and services to specific individuals can be considered to be personalisation, irrespective of the underlying motives. Therefore motives do not need to be included in a general definition of personalisation.

A fourth question is then: how is personalisation done? Or, what processes make personalisation possible? Personalisation cannot be done without any information about the targeted individual. The types and quantity of information about an individual influence how well and how much content and services can be personalised. For example, knowing people’s locations would make it possible for stores to send advertisement to individuals who happen to be nearby. However, if those stores would also have information about the gender, age, purchase history, or preferences of these individuals, they could target their advertisement more efficiently. Different types of information can be retrieved from different data sets. Geographical data reveals where people live or where they are in a given moment; demographic data provides information about gender, age, or marital status; psychological data can tell something about personality types, mental health, or preferences; and behavioural data can be a source of information about one’s previous purchases, or places one often goes to.

So, personalisation takes place by means of data and information about individuals. Acquiring this data and information, analysing it, and connecting it to certain content or services are then necessary processes that are part of personalisation. These processes have been described in several ways. The 2002 White Paper of Vignette Corp. Austin, for example, defines personalisation as “the process of providing relevant content based on individual user preferences or behavior” (p. 66). Pariser very sceptically writes that “In exchange for the service of filtering, you hand large companies an enormous amount of data about your daily life…” (2012, p. 16). Alternatively, one can choose not to indicate the exact processes in a definition of personalisation. However, it is exactly the processes that make personalisation possible that unite the case studies below and, often, these same processes raise ethical and legal concerns. In other words, for the present purposes it is important that the processes behind personalisation are included in the definition. These processes are, broadly, 1) the gathering of data and information about users by means of ‘digital technologies’, such as algorithmic segmentation and decision making, and 2) the use of that data and information to connect content, services, or products to specific individuals or groups of individuals.

Furthermore, content, services or products can be ‘connected’ to individual users in several ways. The content, services or products themselves can be matched to the profile of an individual. For example, users can get different (pre-existing) lifestyle recommendations from their fitness tracker, a voice assistant can adjust its service when used by children, and Netflix’s thumbnails can be changed in order to better appeal to someone’s interest. This way of personalising is *matching* pre-existing lifestyle recommendations, thumbnails, adds, and so on, to specific individuals. In the case of content or services, personalisation can also take place simply by *positioning* the content different. For example, different users are shown different articles when they open their news app and customers see different products when shopping online. Another form of personalisation is *tailoring* content, services, or products that was made to suit the individual’s profile. In this case, products, services, or content are really customized or tailored to an individual. This form of personalisation is more rare than the other two, but should nevertheless be part of the general definition.

In light of these considerations, we define (digital) personalisation as follows:

‘Personalisation’ refers to the joint use of digital data sourced from individuals and (big) data analytics to automatically match, position and/or provide content, services and/or products to specific individuals or categories of individuals.

The personal data used for personalisation can be sourced by the personalised technology itself (e.g. the user’s history) or acquired from third parties. Finally, according to this definition, personalisation does not depend on whether something is experienced as personalised by a user, as users may not be aware that personalisation is occurring.

1.2 Technological details

One of the crucial pieces in personalisation is technology. Nowadays, most of the applications on the Web rely on innovative technological solutions that incorporate Artificial Intelligence (AI). AI is a wide area in computer science composed of different subfields such as Natural Language Processing (NLP), Machine Leaning (ML), Robotics or Visual Recognition. The most relevant one in personalisation is ML, an area that consists in designing efficient and accurate prediction algorithms to make accurate predictions (Mohri et al. 2018). In this section, we describe some of the most relevant AI techniques in personalisation to better understand the case studies described in this deliverable.

In the following, we start by reviewing the algorithms and technologies used in e-commerce and social media, specifically focusing on those used in political campaigns by Facebook. We also briefly refer to the beacon technology used for location-based marketing. Next, we

summaries the main algorithms used in video on-demand streaming services, specifically by the Netflix media company. Then, we continue explaining the technical function of personal assistance voice, and the technology designed in persuasive profile and website morphing. Finally, we introduce another type of marketing that relies on the data provided by external devices or smartphones, as is the case of self-tracking devices in health and sports.
In e-commerce, personalised product recommendations help customers find products they would like to purchase by generating a list of recommended products where a product can be a service or even a political campaign. Different personalisation applications are available according to the type of user data collected, that can be classified as content-based (CB); collaborative filtering (CF) and hybrid applications; collaborative filtering being the most popular one (Wei et al. 2017).

Collaborative filtering systems were designed to explicitly provide users with information about items (Schafer et al. 2007). When a user visits a website or application, the site can adapt its content to the user; in other words, it can predict and recommend the potential favourite items for a particular user by leveraging data collected from similar users. These systems are widely used in commercial recommender systems, but also in social media platforms, such as Facebook, for the purpose of political campaigns. In this case, the findings show that politicians report both marketing and dialogue with voters as motives for their social media use.

In these recommender systems, we find Collaborative Filtering technique that includes two methods: memory-based algorithms and model-based algorithms. In the first case, the collaborative filtering problem is approached by using the entire database. It tries to find users that are similar to the active user and uses their preferences to predict ratings for the active user, with the aim of finding the correlation between two users. For this, similarity measures are used. The result of the measure is a value from -1 to 1 which determines how alike two users are: value of 1 means that they both rate in the exact same manner, whereas a value of -1 means that they rate things exactly the opposite. In the case of model-based algorithms, a model is built based on a dataset of ratings, i.e. information from the dataset is extracted and used as a "model" to make recommendations without having to use the complete dataset every time.

In the area of marketing, it is also worth mentioning the so-called “beacon technology”. This technology is mainly used for location-based marketing, targeting messages depending on where the consumer is located. In any given location, when a mobile application receives a signal, it displays a push notification on the screen to trigger the user’s attention. Basically, this type of technology relies on Bluetooth Low Energy and a beacon. A Beacon is a small wireless device that connects with nearby mobile devices through radio signals with a unique identifier (Moody 2015). With such a simple technology, recommender systems can take advantage of customers being located nearby.

The next algorithms we describe are some of the most commonly used in personalisation, and specifically on Netflix. The set of recommendations are generated by taking into account the user’s personal activity (Davidson et al. 2010) (watched, favourited, likes). In this area, there are several types of algorithms, such as the Personalised Video Ranker algorithm, which order the entire catalogue of videos for each member profile (Gomez-Uribe and Hunt 2015), and the resulting ordering is used to select the order of the videos in genre and other rows. Therefore, different member profiles are shown different videos in the same genre row. This algorithm works better when personalised signals (personalised videos for each user) are blended with popular videos (unpersonalised) (Gomez-Uribe and Hunt 2015).We also find the Top N Video Ranker algorithm, whose aim is to find the best few personalised recommendations in the entire catalogue for each member, but only for those videos at the head of the catalogue. The main difference to the previous algorithm is that the ranking is focusing only on the first videos of the ranking, rather than ranking the entire catalogue. The third and last algorithm is Page Generation: Row Selection and Raking, that uses the output of all the algorithms already described to build every single page of recommendations. Inside Netflix, there are other algorithms for the categories of "Search", "Trending Now", "Continue Watching", but those described above are the ones integrated in the recommender system.

Besides these algorithms, nowadays facial recognition is a big discovery for marketing applications. According to the experts, with facial expressions people demonstrate their wishes on shopping (Barreto 2017). Technology in facial recognition takes as input an image or video stream and the output is an identification of the subject that appears in the image. A face recognition system consists of a three-step process: face detection, feature extraction, and face recognition. Several approaches have been proposed for the face recognition step: template matching face, statistical approach, neural network approach. The statistical approach has proven to be the most relevant ones in this area. In this approach, each image is represented in terms of features (De Carrera 2010) i.e. the image is viewed as a point in a dimensional space the goal is to choose and apply the right statistical tool for extraction and analysis. Consequently, many of these statistical tools are extended or modified to get better results, there are bigger systems that includes or there is a part of a recognition algorithm. There are two main statistical algorithms have been proposed: Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).
PCA is a statistical approach used for reducing the number of variables in face recognition when comparing image features. In PCA, a mathematical procedure is used to simplify the complexity of sample spaces with many dimensions while keeping the information (Amat 2017). LDA is a method of supervised classification where the goal is to obtain a function for classifying a new individual from the knowledge of the values of certain discriminating variables (Gutierrez 1994). The main difference between these methods is that, LDA tries to model the difference between data classes and PCA does not have into account any differences in class.

Next, we will analyse how personalisation techniques and innovative technological solutions have even become part of our home through personal voice assistants. Apple´s Siri, Amazon´s Alexa, Microsoft's Cortana and Google's Assistant are all software agents that run on purpose-built speaker devices or smartphones (Hoy 2018). These assistants are very simple to use not difficult to use, and their software is based on matching user voice input to executable commands. The software constantly expects for a key word to wake it up. Once it hears that key word, it records the user's voice and sends it to a specialized server, which processes and interprets it as a command, and depending on the command, the server will supply the voice assistant with appropriate information to be read back to the user, play the media requested by the user, or complete tasks with various connected services and devices. It is important to mention that this kind of assistants continually learn using AI techniques (especially ML), since huge amounts of data are collected from various sources and can be used to train them. However, depending on each company or developer, each system it applies its own specific methods and approaches, which end up affecting the final product.

Another application where personalisation is used is called "persuasive profiling". This involves collections of estimates of the expected effects of different influence principles on a specific individual. In other words, the purpose of the algorithms used here is to indicate which influence strategies are expected to be more effective for each individual. For example, in the first case, if a general target behavior is smartphone buying, then, such a system may select which specific smartphone to present. And in the second case, one could offer the same smartphone to several users, but show the message that the smartphone is recommended by experts for a certain user profile and emphasize that the smartphone is almost out of stock for another. Both messages be may true, but the effect of each differs between those two different types of users. Technology persuasive is in encourage of this application and for this there are two methods explicit and implicit (Kaptein 2015). Explicit profiling, in this case the measure or the information is obtained with questionaries in which users are asked to reflect upon their own traits and then according to the obtained estimates with the user information is adapted with the correct influence principle. In Implicit profiling is not necessary questionary or other type of actively user-generated data, only with interaction with the system for adapt to his or her personal needs.

Following this topic, we also find what has been termed "website morphing". This is a technique that consists in automatically matching the website look to different customer cognitive styles (Hauseret al. 2009). Technically, this technology first appears with a basic algorithm called HULB, where Bayesian inference is combined with dynamic programming, that is, Bayesian inference on a customer's clickstream infers probabilities that the customer belongs to the latent segment (Hauser et al. 2014). Using these probabilities and data from past purchases, the dynamic program automatically selects the best look and feel for the website for each customer. Then, to account for multiple customer visits to the same website, the HULB algorithm is slightly modified. This is a topic that is currently receiving a lot of attention and research.

Although marketing is a very important application in personalisation and recommender systems, the medical field has taken big advantage of this area, what is called "personalised medicine". And behind this, there is an algorithm, Deep Learning (DL) that is a specialized machine learning approach and provides a more effective paradigm, there are four deep learning architectures used in precision medicine for specific tasks (Zhang et. al. 2018). Convolutional Neuronal Network (CNNs), Recurrent Neural Networks (RNNs), Restricted

Boltzmann Machine and AutoEncoders. CNNs comprise one or more layers and show promising prospects in image-based diagnosis such as radiology, pathology, and dermatology. RNNs are networks with cyclical connections and form feedback loops in their hidden layers. Their use is specially to process temporal-related data and make use of sequential information (Schüll 2016). In the case to RBM is other artificial network where their connection only exists in across layers and no intra layers communication. Therefore, is beneficial for drug discovery. The last architecture, AEs is an unsupervised learning model and is designed to capturing more significant information and learning richer representations, are most appropriate to work many unlabeled clinical data and lead to higher prediction accuracy.

As well as, in recent years people use health apps in smartphones for fitness, fertility tracking or inclusive menstrual period in women. There are many types of technology. However, we can find the sensor technology and big data analytics, where each user is their database, each decision is an item for self-database, that's mean the body is not a sensing organ to gains self-knowledge, the body is a data generating device that must be couple to sensor technology and analytic algorithms to be known.

Furthermore, in others researches for fertility tracking there are other models. Natural Cycles is a famous algorithm to fertility tracking, is a bio-statical model, uses user-logged data on daily BBT and menstrual cycles [HD it Work]. The algorithm was intricately designed to account for sperm survival. Also, ovulation day is based on a weighted average of when ovulation occurred in preceding cycles, assigns the 5 days prior to the predicted ovulation day plus approximately three times the number on standard deviation days as fertile days. Definitely, join the personalisation with a good application of technology is a better customer experience. However, this is always in constantly change and the search for a balance between this and personalisation.

1.3 Ethical and legal issues

Although our case studies will cover in more detail selected uses of personalisation, as well as ethical and legal issues connected to it, we will provide a brief overview to introduce the issues and limitations of using personalised technologies. The following remarks are not meant as a complete classification of real-life applications and concerns surrounding personalisation, but should rather serve as an introduction to the problems that will be discussed more extensively in the case studies.

Personalisation occurs in a broad variety of technologies and sectors. First, social media could not work without personalisation. Every Facebook-, Twitter- or Instagram feed is unique. Personalisation algorithms select the content that they deem most relevant to specific users. Second, recommendation engines such as those of YouTube and Netflix (for videos) and those of Amazon (for products) offer personalised suggestions to each user. Third, personalised news feeds such as Google news select which news articles are displayed to each user (Thurman & Schifferes, 2012; Haim *et al.*, 2018). Fourth, search engines personalise because they determine which links are relevant to *a given user*. Fifth, prices can be personalised, as is commonly done by airlines (Zhang, 2010). Airlines try to charge their users the highest price they are willing to pay. This leads de facto to individual price discrimination, increasing the profits of the airline. Sixth, online ads are personalised for each user to tailor them specifically to a person's interests and desires (Tucker, 2014). Again, each user receives a personalised set of advertisements, deemed most relevant to them. Consequently the use of personalisation in digital technologies has raised several legal and ethical concerns, including the following four clusters.

1.3.1 Politics and civic life.

An important problem connected with personalisation on news websites and social media (which increasingly function as a substitute for traditional media (Bozdag, 2013)) is that people might no longer be confronted with opposing political views. Personalisation algorithms optimise what people like, or click on, which is mostly things they agree with. In this way, personalisation could lead to 'filter bubbles' (Pariser, 2011) in which people with opposite political views no longer encounter one another. Another political concern connected to personalisation is that political campaigners make different promises to different voters on the basis of data about their preferences. Both mechanisms could ultimately cripple trust in the political system and eradicate a sense of community.

1.3.2 Autonomy and Privacy.

Autonomy is often cited in discussions concerning personalisation. One concern is that personalisation may limit autonomy when the weak spots of the consumers are systematically targeted , which could eventually lead to a decrease in decision making abilities (Ignatidou, 2019). Other autonomy concerns are related to privacy because personalisation techniques may lead to a decrease of control over one's environment, personalisation could be seen as a threat to privacy, even when considered separately from concerns about the disclosure of data (Toch *et al*., 2012; Tucker,2014).

1.3.3 Justice.

Because personalisation enables firms and governments to treat users differently, unfair differences in treatment may arise. This can happen either deliberately, or unconsciously, through bias. One prominent concern is that the less advantaged "are likely to be disadvantaged and disempowered by the turn to mass personalisation" (Yeung, 2017, p 9). To illustrate, O'Neil (2012) discusses an example from the US, where inferior schools use aggressive marketing strategies whereby they focus on people from lower income groups, further disadvantaging the less privileged by making it more likely that they will receive worse education. This example shows that personalisation can be used to unfairly target vulnerable groups.

1.3.4 Economic concerns.

Calo warns about "the capacity of firms to influence consumers at a personal (…) leading to actual and perceived harms that challenge the limits of consumer protection law" (Calo, 2014, p 999). When firms are becoming better at persuading consumers by personalising both the products featured in advertisements, as well as the advertisements techniques to each user, the power of corporations to shape consumer preferences increases. When this occurs, consumers may face an inability to resist advertisement and may suffer a loss of utility derived from their economic activity. For Calo, this challenges fundamental consumer rights. In a similar vein, Yeung (2017) warns that in the consumer context, personalised services may lead to mass manipulation. She argues it is "inherent to the nature of capitalism" that sellers will look for ways to manipulate consumers, by, for example, persuading them to purchase products and services they would never have considered on their own.

1.4 Methodology

We now introduce 12 case studies on digital personalisation technologies to provide a general overview of ethical and legal issues associated with them. We wanted to discuss broad and varied examples of personalisation connected to these technologies in order to present a comprehensive overview of possible applications of digital personalisation. The case studies included in this deliverable are meant to serve as a foundation for the discussion of ethical and legal aspects of digital personalisation. In that sense, we hope that the diversity in the chosen technologies will make it easier for readers to look at other examples in a similar fashion and extend some of our observations to other personalisation technologies.

Where it was feasible, we focused on types of technology (e.g. beacons, fitness trackers or e-learning) and discussed issues connected to personalisation that were applicable across particular products. This was done in order to provide a broad overview of the ethical dimension of personalisation that a reader could then relate to individual applications of a given technology and perhaps supplement with ethical issues arising from that particular application. However, this was not possible in all instances. Some technologies varied across their applications to such an extent, that a case study focusing only on general issues connected to type of technology would not be adequately informative. This was the case, for example in Ana's second case study, which discusses Netflix's personalisation algorithms which greatly differ from personalisation algorithms used by other media companies. On the other hand, Facebook's political ads, as discussed in Ana's third case study, had such a great impact on recent elections that we decided they merited being discussed on their own and not grouped together with other internet advertisement systems providing political content.

The goal of this deliverable is to present the current state of ethical and legal research as related to the discussed technologies. Consequently, we only give an overview of ethical and legal aspects of selected personalisation technologies and do not provide an evaluation of the discussion reflected in the literature. Similarly, at this point, we do not attempt to identify ethical and legal concerns on our own and only summarise those addressed by other scholars – this something that will be done as part of the second deliverable.

In order to find relevant articles discussing ethical and legal issues connected to particular technology types and products, we conducted searches in three databases: Google Scholar, Scopus and HeinOnline. Google Scholar allowed us to retrieve a wide variety of sources from across disciplines and served as a foundation of the search. Scopus was chosen as it is one of the biggest academic databases indexing peer-reviewed journal articles and it was used to retrieve more specialised academic papers. Legal aspects of personalisation technologies

were found mostly through the use of HeinOnline as it is a database exclusively oriented toward legal research. In some instances, the discussion of legal aspects of personalisation was also supplemented by research connected to the GDPR in order to provide a more EU-oriented perspective.
The case studies are connected to different extents with our PhD projects and research interests, and divided into four groups that discuss the ethics of (1) digital surveillance (Ana Fernandez); (2) video analytics (Rosalie Waelen); (3) digital nudging (Paul Kuyer); (4) self-tracking technologies (Michał Wieczorek). They all follow the same structure. First, the objective of the case study is introduced. Then, a technology description is provided, followed by the context outlining the standard use and possible future applications of the technology. Next, we present how the case study is related to the Work Package as a whole and to the author's PhD topic. Finally, we provide an overview of the literature discussing ethical and legal aspects connected to the discussed personalisation technology by first discussing the methodology of the search and then summarising the findings.

1.5 Summary of case studies

Case 1: Beacon technology (by Ana)
Ana's first case study discusses ethical concerns surrounding beacon technology, which is embedded in many personal devices such as smartphones and uses Bluetooth Low Energy signals to track the movement and behaviour patterns of users. These are later collated into a personal profile of the user, which most often serves as a foundation for marketing services or public interest initiatives, as in COVID-19 contact tracing. The great accuracy of beacons allows for the personalisation to occur even on a very small scale, for example, by tracking the movements of a customer inside a single store and guiding them towards products that they should find desirable. However, this raises concerns over unlawful surveillance, intrusiveness, threats to consumer autonomy, as well as the biases and lack of transparency of the used algorithms.

Case 2: Netflix's personalisation algorithm (by Ana)
The second case study written by Ana centres on Netflix's algorithmic personalisation. The media company collects user data in order to maximise user-engagement by suggesting and positioning content that is deemed the most relevant to particular users. The case of Netflix is particularly interesting, as the personalisation extends far beyond only suggesting different titles and arranging them in a different way. Netflix even engages in website morphing and accompanies their content with images that are selected to match a particular user's interests. There is no transparency, however, as to how those recommendations are made and there exists a risk that they might manipulate or confuse users, as well as contain some biases without necessarily providing them with anything useful.

Case 3: Personalised political campaigns on Facebook (by Ana)
Ana's last case study deals with targeted political ads, particularly those presented on Facebook. Following electoral scandals in the UK and the US in 2016, political advertising has been criticised for manipulating voters, leading to the polarisation of opinions and

unfairly influencing the results of the elections. The case study outlines how personal data factors into targeted political ads and highlights the mechanism that determine which ads will be presented to particular user types. While the problems discussed above remain relevant, the use of personalisation mechanisms in political advertising will have serious implications for the future and much attention needs to be brought to these ads' lack of transparency and accountability to the public, potential biases and voter manipulation.

Case 4: Persuasive profiling (by Paul)
Paul devotes his first case study to persuasive profiling, which uses personal data to model consumer characteristics in order to personalise persuasion methods. Persuasive profiles are a powerful marketing tool as they enable companies to target specific individuals and model which marketing strategies and persuasion techniques will be most effective in relation to them. By extracting psychological traits from users' past behaviour, marketers construct persuasive profiles are able to position products and services in a way that maximizes their profitability. However, this can be seen as manipulative, as it is difficult to assess whether the high efficiency of persuasive profiling stems from its ability to adapt marketing strategies to specific users or whether it distorts consumer choice by making them reach for products they would never have purchased without external influence.

Case 5: Website morphing (by Paul)
The second case study written by Paul is devoted to website morphing, a mechanism which changes the appearance of a website based on the personal data of the user that is browsing it. While morphing does not personalise information presented to the users, it still offers website owners opportunity to maximise their revenues by retaining users' interest and using visual cues in order to influence their browsing patterns. A website could change a colour scheme to one that is known to be pleasant to a particular person, or it could change the size of specific links or images in order to draw more attention to them, which might be particularly successful if a user's affinity towards certain type of content is known beforehand. Interestingly, at the time of writing, no paper on the ethical and legal aspects of website morphing has been published, which means that Paul's second case study is charting a new terrain.

Case 6: Personalised e-learning (by Paul)
Paul's last case study discusses ethical and legal aspects of personalisation used in e-learning services. Student data in e-learning can serve as a basis for adapting both the content and delivery of teaching, thus increasing student-engagement, while maximising learning outcomes and reducing costs of teaching. In an ideal personalised teaching environment, student would be met with tailor-made examples, which best illustrate the ideas that are most suitable to their current progress in a given field. However, some risks are involved here. The collection of student data for personalisation purposes creates a danger to their privacy and the use of this data by education companies raises concerns over commodification of education and the power private entities hold over public learning institutions. Moreover, the models used in e-learning have also been criticised as their

predictions are not always accurate and they are overdependent on metrics, while their results are often questionable.

Case 7: Personal voice assistants (by Rosalie)
In her first case study, Rosalie deals with digital voice assistants, which are software systems installed in devices such as smartphones and smart speakers in order to provide assistance to users. These assistants process voice commands issued by the user and provide personalised services or recommendations on the basis of the user's behaviour and preferences, for example by recommending them songs that might fit their mood or by helping them navigate through smartphone menus. A voice directed user interface should not only make it easier and more natural to use the system, giving the assistant a voice also gives a personal feel to the service. Thanks to this personalised technology, having a personal assistant is no longer a privilege of the few, but threats to privacy, autonomy of the users and trust, as well as the risk of algorithmic discrimination should make potential adopters cautious.

Case 8: Personalised news (by Rosalie)
For her second case study, Rosalie chose personalised news, that is, the selection and positioning of news items for specific users based on their preferences and interests. This personalisation can occur both at the level of a single newspaper (e.g. different users might see a different order of articles on New York Times' website) and on a level of a news aggregator (e.g. each user's Google News feed will present them different information). While this can benefit users by presenting them stories they are most interested in, many concerns surrounding personalised news can be identified. Among other things, such personalisation can be questionable in reference to its impact on democracy, personal identity, and fairness .

Case 9: Facial recognition (by Rosalie)
Rosalie's third case study discusses the use of facial recognition for personalisation purposes. Although the field of facial recognition has been growing very quickly in recent years, not much research has been done on the ethical and legal aspects of the employment of such technologies in personalisation. As noted in the case study, facial recognition can be used for personalisation in various ways. For example, it can improve personalised marketing and customer experiences, by recognizing known customers or responding to customer's facial expressions. This, and other possible uses of facial recognition for personalisation, raise many issues, including those connected to the biases inherent in the discussed technologies, possible errors, function creep and the erosion of privacy.

Case 10: Personalised medicine (by Michał)
The first case study written by Michał deals with continuous glucose monitors (CGM) - wearable medical sensors that allow diabetic patients to take constant readings of their blood glucose level and have those displayed on a screen of a connected device. Additionally, associated apps often use this data to issue lifestyle recommendations. When coupled with insulin pumps, CGM promises to automate the management of type 1 diabetes by administering glucose whenever the user's blood glucose level falls below a

certain threshold, thus providing personalised treatment on the basis of gathered data. The use of CGM not only has a great impact on the lives of diabetic patients, but also has significant consequences for the functioning of the healthcare system as a whole.

Case 11: Fitness tracking (by Michał)

Michał's second case study outlines the personalisation connected with self-tracking devices used for fitness purposes. Devices such as Fitbit collect users' activity and health data in order to provide personalised recommendations associated with their lifestyle, diet and exercise patterns. However, the models used by the developers of fitness self-tracking technologies often use models that do not reflect the diversity of their users and offer marginalised groups recommendations that are not as useful as the recommendations issued to those fitting the image of a "standard body". Issues connected to the management of user data and privacy are also relevant here as self-tracked activity data can be of particular interest to third parties, who are often able to draw substantial profits from personal activity that is not economically oriented.

Case 12: Fertility tracking applications (by Michał)

The third case study prepared by Michał deals with fertility tracking applications. Predominantly aimed towards women, these apps depend on user-recorded data to predict numerous factors connected with fertility and menstruation. Typically, Users are asked to record their previous periods, basal body temperature and physical activity in order to receive information on how to facilitate or avoid pregnancy, and when the next period might happen (as well as its expected symptoms). Many of these apps have been criticised for their low accuracy and for the lack of transparency of the models they are using. Moreover, their interfaces and the content and form of their predictions have been demonstrated to enforce gender stereotypes and exclude marginalised groups. Privacy issues are also significant here as data connected to fertility is especially valuable to many third parties such as employees, insurers, advertisers and partners.

1.6 References

Blom, J. (2000). Personalisation – A Taxonomy. *CHI'00*. 313-314.

Brusilovsky, P. & Maybury, M.T. (2002). From adaptive hypermedia to the adaptive web. *Comm. of the ACM, 45*(5). 30–33.

Calo, R. (2014). Digital Market Manipulation. *THE GEORGE WASHINGTON LAW REVIEW*, *82*, 58.

Fan, H. & Poole, M.S. (2006). What Is Personalisation? Perspectives on the Design and Implementation of Personalisation in Information Systems. *Journal of Organizational Computing and Electronic Commerce, 16*(3&4). 179-202/

Fenech, C., & Perkins, B. (2015). The Deloitte Consumer Review. Made-to-order: The rise of mass personalisation. Deloitte Development LLC, 1-20.

Hagen, P.R. (1999). *Smart Personalisation*. Cambridge, MA: Forrester Research, Inc.

Ignatidou, S. (2019). *AI-driven Personalisation in Digital Media: Political and Societal Implications*. https://www.chathamhouse.org/publication/ai-driven-personalisation-digital-media-political-and-societal-implications

Karat, J., Marat, C. & Ukelson, J. (2000). Affordances, motivations, and the design of user interfaces. *Comm. of the ACM, 43*(8). 49–51.

Kasanoff, B. (2001). *Making it Personal: How to Profit from Personalisation Without Invading Privacy*. Basic Books.

Kim, W. (2002). Personalisation: Definition, status, and challenges ahead. *Journal of Object Technology, 1*(1). 29–40.

McCaig, I. 2013. The history of personalisation. Econsultancy. Retrieved from: https://econsultancy.com/the-history-of-personalisation/

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Pariser, E. (2012). *The Filter Bubble. What the Internet is Hiding from You*. Penguin Books.

Surprenant, C. & Solomon, M. (1987). Predictability and personalisation in the service encounter. *Journal of Marketing*, *51*(2). 86–96.

Salonen, V., & Karjaluoto, H. (2016). Web Personalisation: The State of the Art and Future Avenues for Research and Practice. *Telematics and Informatics, 33*(4). 1088- 1104. doi:10.1016/j.tele.2016.03.004.

Thurman, N., & Schifferes, S. (2012). The future of personalisation at news websites: Lessons from a longitudinal study. *Journalism Studies*, *13*(5–6), 775–790.

Toch, E., Wang, Y., & Cranor, L. F. (2012). Personalisation and privacy: A survey of privacy risks and remedies in personalisation-based systems. *User Modeling and User-Adapted Interaction*, *22*(1–2), 203–220.

Tucker, C. E. (2014). Social networks, personalised advertising, and privacy controls. *Journal of marketing research*, *51*(5), 546–562.

Vignette Corp. (2002). *Personalisation Strategies-Fit Technology to Business White Paper.* Austin.

Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, *20*(1), 118–136. https://doi.org/10.1080/1369118X.2016.1186713

Yeung, K. (2018). Five Fears About Mass Predictive Personalisation in an Age of Surveillance Capitalism. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3266800

Zanker, M., Rook, L., & Jannach, D. (2019). Measuring the impact of online personalisation: Past, present and future. International Journal of Human-Computer Studies, 131, 160-168.

Zhang, J. (2010). The Perils of Behavior-Based Personalisation. *Marketing Science*, *30*(1), 170–186. https://doi.org/10.1287/mksc.1100.0607

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of machine learning*. MIT press.

Wei, j., he, j., Chen, k., Zhou, y., & tang, z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert systems with applications*, *69*, 29-39.

Schafer, j. B., Frankowski, d., Herlocker, j., & Sen, s. (2007). Collaborative filtering recommender systems. In *the adaptive web* (pp. 291-324). Springer, berlin, Heidelberg.

Moody, M. (2015). Analysis of promising beacon technology for consumers. *Elon Journal of Undergraduate Research in Communications*, *6*(1).

Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., ... & Sampath, D. (2010, September). The YouTube video recommendation system. In *Proceedings of the fourth ACM conference on Recommender systems* (pp. 293-296).

Gomez-Uribe, C. A., & Hunt, N. (2015). The Netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)*, *6*(4), 1-19.

Barreto, a. M. (2017). Application of facial expression studies on the field of marketing. *Emot. Expr. Brain face, no. June*, 163-189.

De Carrera, P. F., & Marques, I. (2010). Face recognition algorithms. *Master's thesis in Computer Science, Universidad Euskal Herriko*, *1*.

Amat R (2017). Análisis de Componentes Principales (Principal Component Analysis, PCA) y t- SNE.

Gutiérrez, R., González, A., Torres, F., & Gallardo, J. A. (1994). Técnicas de análisis de datos multivariable. Tratamiento computacional.

Hoy, m. B. (2018). Alexa, Siri, Cortana, and more: an introduction to voice assistants. *Medical reference services quarterly*, *37*(1), 81-88.

Kaptein, M., Markopoulos, P., De Ruyter, B., & Aarts, E. (2015). Personalising persuasive technologies: Explicit and implicit personalisation using persuasion profiles. *International Journal of Human-Computer Studies*, *77*, 38-51.

Hauser, j. R., urban, g. L., Liberali, g., & Braun, m. (2009). Website morphing. *Marketing science*, *28*(2), 202-223

Hauser, J. R., Liberali, G., & Urban, G. L. (2014). Website morphing 2.0: Switching costs, partial exposure, random exit, and when to morph. *Management science*, *60*(6), 1594-1616

Zhang, S., Bamakan, S. M. H., Qu, Q., & Li, S. (2018). Learning for Personalised Medicine: A Comprehensive Review From a Deep Learning Perspective. *IEEE reviews in biomedical engineering*, *12*, 194-208.

Schüll, N. D. (2016). Data for life: Wearable technology and the design of self-care. *BioSocieties*, *11*(3), 317-333.

it Work, H. D. Independent Expert Fact Sheet on Natural Cycles–a Certified Contraceptive App.

2 Case study 1: Beacon Technology

2.1 Objective

The objective of this case-study is to introduce the ethical and legal issues of Beacon technology, a personalised marketing technology that uses Bluetooth (BLE) to gather location data in order to send relevant and personalised offers to customers. Literature in ethical issues of marketing has been extensively focused in the online aspect of marketing, such as the type of personal information that is retrieved online, how this is intrusive to people's privacy and autonomy (Pariser 2011), or the lack of transparency of how algorithms retrieve information and make decisions[1]. This case study focus on the necessity to study how Beacons modify the already present online tracking, but is an off-line technology that adds location information to the represented virtual person. I argue that an ethical study of Beacon technology is needed, as it brings new ethical and legal issues that go beyond online tracking and online behavioural targeting, and it becomes intrusive to people's real-life movements and spaces. As BLE technology develops, privacy impact assessments must be encouraged to comprehend the influence that this can have to future human-technology interactions.

2.2 Technology description

Beacons are a technology first introduced by Apple in 2013 as iBeacons (Maycotte 2015), a new technology in off-line marketing (Allurwar, Nawale, Patel 2016). It is a Bluetooth low energy (BLE), low frequency and wireless device placed in different stores or streets, that sends signals to other small BLE devices nearby such as mobile phones or tablets. This technology is being adopted rapidly, since 2013 most smartphones contain BLE technology (Moody, 2015), and it's expected to keep growing[2]. The Global Market predicted that Beacon technology will be over 25 billion by 2025 (Wadhwani 2018).

Their size is small, Beacons started as "devices about the circumference of a large apple; today they're mere stickers that can be placed on walls or objects. The smaller and less obtrusive they get, the easier they become to use" (Maycotte 2020). These chips communicate with other Beacon devices and are used by marketers to better personalise the messaging and mobile ads based on customer proximity, while making the tracking location easier and more accurate (personalised). The beacon works as a unique identifier (ID) in the smartphone, broadcasting constantly that identifier (Raiz, 2018).

Marketing is a sector that is increasingly using personalised technologies, whether it is for optimizing advertisements based on customer's interests and web activity, or improving user's engagement online. Generally, marketing technologies refer to online systems that collect large data sets from different sources, for example data management platforms such as Oracle, or customer relationship management like Salesforce (Biegel 2009). These systems are used for marketing online, their goal is to send customers relevant and personalised information by using online browsing data and algorithms to improve sales.

[1] See case studies on Netflix algorithms and Facebook political campaigns by Ana Fernández

[2] WordStream Blog. 5 Things you Need to Know about Beacon Technology. 2020. Retrieved from: https://www.wordstream.com/blog/ws/2018/10/04/beacon-technology

However, the rapid development of marketing technologies urges for the study of new forms of personalised marketing (Toch, Wang, Cranor 2012). Nowadays, marketing is expanding and it is not only an activity restricted to online scrutiny, but it retrieves information from different sources, such as information gathered through Beacons. In this line, offline and online habits have become an interactive activity where Beacons allows to track a person in real time and, for example, to invite people to enter a store nearby or check potential offers based on customer's particular location. This technology influence how people interact with their surroundings by pointing the exact location and therefore allowing users to receive relevant and personalised messages and advertisements from immediate stores, interacting and influencing individuals' connection with their nearby environment. Businesses that have beacons in place can detect where a customer specifically is located at any given moment in their proximity. A Beacon device sends data regularly and its detected by an app or pre-installed service, like Google Nearby[3], on smartphones within reach (Adarsh 2020). The Beacon recognizes the ID number and it is linked to an action, such as receiving an offer, and these messages can be stored on the cloud (Figure 1).

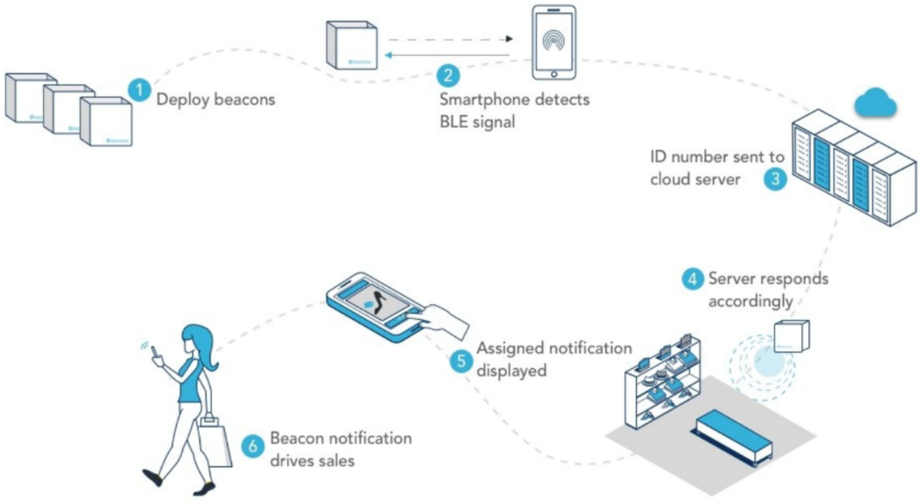

1 Deploy beacons
2 Smartphone detects BLE signal
ID number sent to cloud server 3
4 Server responds accordingly
5 Assigned notification displayed
6 Beacon notification drives sales

Figure 1 retrieved from Beaconstac[4]

Thus, BLE technology also empowers business to interact with customers even after they checked out the store, because the location data is exposed online with Google or Facebook. Beacon technology is promoted by Google and can be integrated to Google Maps, where the most frequent places gathered trough BLE will become more visible

[3] Google Nearby https://developers.google.com/nearby
For more information about the use of Beacons and Google Nearby visit: https://blog.beaconstac.com/2018/10/google-just-killed-android-nearby-notifications-whats-next-for-proximity-marketing-using-beacons/

[4] The image was retrieved from Beaconstact: https://www.beaconstac.com/proximity-marketing

(Blennd Blog 2018). This means that customers are not only tracked by past searches and behaviors online, but their real-life movements are also constantly monitored. Moreover, Apple's iBeacon technology[5] is introduced as a device that alerts an app or website when a customer approaches or enters a location to send useful and personalised information even without the user having to interact with the app. An example is Carrefour, the company installed Beacons in many supermarkets in Romania and sent push notifications to customers before arrival, suggests items based on purchases made previously and display personalised coupons (Hedge 2019). It is a way to engage with customers "in a more personalised way" (Maycotte 2020). BLE also allows to gather information about the exact movements of a customer inside the store, it is even more accurate than the GPS or WiFi, and can be used to send relevant messages to clients inside the store and to send them messages such as promotions or discounts in real-time. It is therefore possible to send targeted adds via BLE, and to send individual and personalised discounts in real-time using beacons depending on the customer's position and habits in the store.

2.3 Context

Beacon Technology is not only used for marketing purposes, but it is expanding and developing new uses. A baseball league is adopting beacons to communicate with people in stadiums and offer seats upgrades, and airlines are using Beacons to communicate with customers in the airport (Maycotte 2020). Evenmore, one of the current uses of BLE technology is to develop Covid-19 tracing apps, such as the one named DP^3T[6], where people can keep track of people they have been in close proximity with. There are different data points that can be gathered from the device, for example, the latitude or longitude of the terminal, the direction of travel of the user, and the time the location was recorded (DPIA report, p. 15). It is argued that the use of Beacons allows for a less privacy invasive technology, the tracing-app has a decentralised system that do not store all the information in one server, and therefore protects the privacy of the individuals that go in line with the GDPR data-protection by design[7]. Although this is now being questioned by several authors (Vaundenay 2020) given the risks to deanonymized data.

Thus, Beacon technology is used for different goals, but in this case study the focus lies on marketing personalisation. This is useful to, for example, send personalised advertisements from the stores a user has visited based on the movements that the customer has done inside the store. If beacon's signals are connected to Google Ads account, it gains a lot of insight into a consumer's offline activity and may even track in store visits (Wadhwani 2018). This means that when a company use Google search ad, it is able to link the ID with the online user that walks into the store.

[5] About iBeacon on your iPhone, iPad, and iPod touch. Retrieved from: https://support.apple.com/en-us/HT202880

[6] DPIA Report – DP^3T, Version: V.01| 01.05.2020, available at https://github.com/DP-3T/documents/blob/master/data_protection/DP-3T%20Model%20DPIA.pdf

[7] Article 25(2) GDPR

2.3.1 1.1.3.1 Potential Use

Although marketing is one of the most famous uses of Beacons, this technology has come a long way since it has been introduced by Apple in 2013. Beacon is a technology that is developing new uses, for example Starwood Hotels is a company that is studying to replace hotel room keys with beacon devices (Maycotte 2020). Beacons' use is increasing in numerous areas of our society, Hedge (2019) mentions that it is being used in food services, travel and tourism, airports, healthcare, sports, banking, analytics, cinema or logistics. In the same article, it is exposed how it has also been used in political campaigns such as Trump's, where they used Beacon IDs to track user's phone, retrieve their interests, and make a portrait of them (Hedge 2019). This activity can be developed further into other political activities in the near future. Furthermore, recent uses are seen in a Vulcania Theme Park, Beacons assist people to find their way into the complexity of the park and show the nearby attractions, pushing the ones closer to the user and with the fewer queue (Idem). Finally, Japan's Nagoya University Hospital, deployed beacons to store patient's vital signs and positions, including the movements from the staff (Idem), for that they use MEDiTAG, a sensor-based beacon wristband to gather patients' vital signs, one of the data retrieved was used to calculate nurse's response to patients calls. These examples allow to foresee a wide variety of uses in the near future, while Beacon sensors continue to grow.

2.4 Relationship to Workpackage

Work package 2 focuses on personalisation, technologies that involve gathering personal information and targeting an individual. This case-study shows how Beacon is a personalised technology that is being used to connect with the customer in real-time. It can be also argued that this is not personalised but facilitates personalisation[8]. Nonetheless, this case-study exposes how BLE technology allows for a high level of personalisation. Beacon signals are off-line, helping marketing companies to understand the offline history of a customer's movements. This creates an even more personalised marketing strategy that adds to customer's online habits by following your steps without the need to follow user's browsing data. Thus, allowing to send advertisements but related to user's real movement habits.

2.5 Relationship to ESR PhD topic

My thesis will be directed towards Digital Surveillance and this case-study will be used to explore the area of Beacons as a future development in different surveillance technologies and their ethical consequences. While this case-study has a focus on marketing, the future development of Beacons can become an important factor in Surveillance studies, such as those exposed with the use of Covid-19 tracing-apps and that can be used to track offline movements of communities and countries. Despite this fact, current literature in surveillance have not focus on Beacons, and this case-study can provide a relevant start on this area.

[8] This requires further study that falls out the scope of this research

2.6 Review of ethical and legal issues

2.6.1 Methodology

For this study, a critical analysis was used to identify relevant literature on Beacon Technology. Firstly, Google Search was used as a main source for non-academic literature such as newspapers, blogs and other relevant literature. Secondly, Google Scholar and Scopus was used for academic articles. Finally, the legal research was mainly focused on the GDPR. The terms that were used in this search engines to conduct the research were divided into search words: “beacon technology” “BLM”, “ethics” “moral” “issues”, “privacy” “law”. The findings were prepared on the basis of a critical selection of the most relevant articles and this analysis will provide a comprehensive summary of the ethical and legal issues related to Beacon technology.

2.6.2 Findings

Although businesses are investing significant amount of resources in developing device applications to collect data from beacons, they do not put similar efforts in informing users about their ethical issues and risks. Beacons have both positive and negative ethical effects as it has been shown that BLE technology allows for a variety of uses, specially in marketing, where data retrieved by customer’s movements is used to personalise offers and can help customers find a way inside the store. This WP is limited and it will be included only the negative effects of Beacons[9].

There are new ethical and legal issues that emerge with the adoption of Bluetooth technology, in this section some of them will be summarized, taking into account that this is a relatively new technology and there is not an extensive literature available. Some of them are the invasion of customer’s personal and inter-personal spaces, power imbalances with companies using customer’s walking habits to increase sales, reduction of autonomy, the difficulty for consumers to opt-out from this technology, and other privacy intrusions. In 2014, Buzzfeed published a report[10] stating that a company controlling a number of New York City’s phone booth advertising displays, installed BLE technology in hundreds of phone booths, without any public notice or consultation. Similarly, Target also announced in 2015 the use of Beacons in their stores (Perez 2015). The spread of beacon technology can turn a city into a commercialised map, adding to the surveillance domain in a city that have already grown from security cameras to cell phone towers, and now Beacons. The apps that were

[9] The positive effects of Beacons require further study, while it is implied in this case-study that Beacons are a relatively new intrusive technology and the risks and ethical issues must be studied.

[10] Exclusive: Hundreds Of Devices Hidden Inside New York City Phone Booths. Retrieved from : https://www.buzzfeednews.com/article/josephbernstein/exclusive-hundreds-of-devices-hidden-inside-new-york-city-ph

used to detect Beacons in New York were any app that uses this particular technology, the article mentions that if a user have the app of a Major Baseball League, they will be subject to scrutiny by all these other beacons without even agreeing to it. This led to a number of complaints from citizens and privacy advocates, and the city ordered to remove the beacon devices after the article was published. Hence, **lack of transparency** is one of the concerns, if a customer downloads one app that uses BLE technology, he or she will be subjected to further tracking by other business that contain the same technology; as it is exposed "whether consumers know what method is causing their apps to know when they're in the right spot to be sent a notification appears to be beside the point" (Kaplan 2016). Furthermore, Beacons are small, inexpensive and easily-placed devices that allow to be unnoticed, in contrast to other technologies are more visible like CCTVs. Consumers should be aware when they are in a zone that contains beacons.

Daniel Dimov, a privacy lawyer, has informed about relevant misuses and ethical considerations of Beacons (Dimov 2014). He argues that there are at least two privacy risks associated with BLE technology: risk of unlawful surveillance and risk of receiving undesired targeted advertisements. The **risk of unlawful surveillance** refers to the possibility to target users by finding their location in relation to a Beacon, without them even knowing it. He further explains that there is a growing number of malware programs designed for smartphones that can also be used to retrieved that information. He argues that criminals may use the collected location data to identify behavioral patterns of possible victims, for instance, an individual who often visits the richest part of a city may be more attractive for thieves than an individual who visits poor parts of a city. This is similar with the business that are able to track a customer that went to the richest stores, and can be more attractive to commerce, also influencing the advertisements that they receive. On the other hand**,** the risk of receiving undesired targeted advertisements, although it is not an ethical issue in itself but can lead to discomfort, **harms to autonomy and privacy**, is based on the fact that it is difficult for customers to agree and understand the long list of terms and conditions and the privacy policies, this influences the capacity for customers to understand when they download an app that tracks BLE technology and how this will influence the advertisements. Users might not read how data is collected, nor even know what Beacons are, like Dimov puts it: "You agree that we may collect location data from beacons. You agree that we may pass the collected location data to third parties who may then regularly send communications to you and provide information offers and services that may be of interest to you". When the users agree to the long list of terms and conditions without reading them, hundreds of advertisements can be sent in accordance with the location of the user, even they might not know that Beacon tracking is taking place. Moreover, if the approval is by default, beacon devices can send information to customers in real-time without the necessity for customers to agree (Consumergateway 2016). In addition, DP^3T proposal to use BLE technology rightly points out the issues with this technology, although not used for marketing purposes, it is the first study that exposes privacy issues. They included in the privacy risks, that the use of Bluetooth allows the smartphone device to be tracked (DPIA report, p.35).

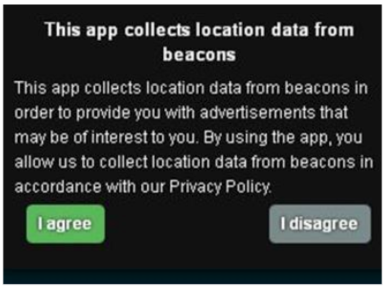
It is also important to highlight that to send too many pop-up messages, offers or discounts can have a negative impact in the individuals who were not interested, but they just were close enough to a beacon, and **intrusiveness** might be triggered by receiving too many undesired messages. Even more, the customer might consider that if they have their phone closed they will not receive these messages, but BLE technology works with Bluetooth and allows to be used even when the customer does not open the app.

Thus, **lack of information and power imbalances** is an issue because BLE is a new technology that most of the people are not aware of and do not know that it exists (Kaplan 2015). Also, the data that is collected must be explained to the customer as business can gather different types of data, including sensitive data, means that the user must be aware and accept the gathering of these different data points, such as location data or personal data, and to understand how the beacon works, if its centralised or not, and what are the chances of being identified. For example, a Beacon can retrieve information on areas of the store that shoppers visit more frequently, how long a customer stays in a given area, and number of movements between areas. However, it is exposed that even with the decentralised system, risks of individuals being identified cannot be completely excluded (DPIA Report p.7). The DPIA report shows that the type of data being collected on Bluetooth beacons could reveal information about families, societies and communities. It is shown that when there are small communities, the source of the individuals that tested positive can be easily identified. It is therefore important to highlight that this is a similar issue for other uses of beacon technology, when even if decentralised, the individual can still be identified.

Reduced autonomy is also part of the issues. If the Beacon is associated with a specific product or store, the customer may engage with it by actively approaching the phone to a Beacon (Consumergateway 2016). On the contrary, Beacons can be sending information about specific products therefore reducing the autonomy of the individual, in this line, the app communicates with the beacons without customers taking a voluntary action. Neither are they in control of the information stored by a Beacon.

Personalised-movement behaviour can trigger **biases**, as beacon technology for marketing purposes in store can be used to gather personalised shopping behaviour and offer certain products based on it, such as coupons and advertisements based on your movements. Products that a customer could not have been interested or rather wrongfully linked to an individual. Similarly, a customer might receive information from a local that is nearby but it is not interested in.

To reduce the risks exposed by beacons, literature has shown that customers should be free to decide when and how to use them. It can be helpful to oblige smartphone companies to obtain user’s consent before collecting BLE location data, as well as apps to specifically mention if they are using BLE technology. If the business in charge of apps are oblige to use pop-up windows when collecting location data, users would be aware of the fact that their location data is being collected such as in Figure 2.


This app collects location data from beacons

This app collects location data from beacons in order to provide you with advertisements that may be of interest to you. By using the app, you allow us to collect location data from beacons in accordance with our Privacy Policy.

Figure 2 Retrieved from (Dimov 2014)

Retailers should let their customers opt-out and be cautious in any attempt to remotely open customer's apps on smartphones and sending messages to them, because imposing and interfering with customer autonomy may get the opposite outcome and trigger the user to remove the app by a feeling of an intrusion of privacy. Finally, laws and regulations such as GDPR focus on the storage of personal data, Beacon technology might retrieve personal IDs while linked to databases such as Google Ads, thus business using Beacon technology have to be aware of these regulations. Following GDPR, considerations about the purpose of retrieving customer's location must be specified and transparency should be clearly stated. In summary, the possibility to gather user's movement behaviours through Beacon technology and personalise advertisements and offers based on this technology rises ethical issues related to people's privacy invasions such as lack of transparency, risk of unlawful surveillance, risk of receiving undesired targeted advertisements, lack of information, power imbalances, reduced autonomy and biases.

2.7 References

Adarsh, M. 2020. Bluetooth Low Energy (BLE) beacon technology made simple: A complete guide to Bluetooth Low Energy Beacons. Beaconstact. Retrieved from: https://blog.beaconstac.com/2018/08/ble-made-simple-a-complete-guide-to-ble-bluetooth-beacons/

Allurwar, N., Nawale, B., & Patel, S. (2016). Beacon for proximity target marketing. Int. J. Eng. Comput. Sci, 15(5), 16359-16364.

Biegel, B. (2009). The current view and outlook for the future of marketing automation. Journal of Direct, Data and Digital Marketing Practice, 10(3), 201-213.

Blennd Blog.2018. WHAT IS GOOGLE BEACON: HOW GOOGLE BEACON PROJECT AFFECTS SEO. Retrieved from: https://blennd.com/what-is-google-beacon-technology-seo/

Consumergateway. 2016. On the Use of Beacons in Retail: Practice and Research. A Gateway to Consumer and Customer Behaviour. Retrieved from: https://consumergateway.org/2016/02/26/on-the-use-of-beacons-in-retail-practice-and-research/

DPIA Report – DP^3T, Version: V.01| 01.05.2020, available at https://github.com/DP-3T/documents/blob/master/data_protection/DP-3T%20Model%20DPIA.pdf

Dimov, D. 2014. Privacy Risks of Beacons. Infosec. Retrieved from: https://resources.infosecinstitute.com/privacy-risks-beacons/#gref

Frey, R. M., Vuckovac, D., & Ilic, A. (2016, September). A Secure Shopping Experience Based on Blockchain and Beacon Technology. In RecSys Posters.

Hedge, A. 2019. 5 powerful beacon use-cases for 2020: Takeaways from 2019. Beaconstac. Retrieved from: https://blog.beaconstac.com/2019/11/powerful-takeaways-from-beacon-use-cases/

Kaplan, D. 2015. Most Consumers Don't Know Beacons Exist — Should That Matter To Retailers?. GeoMarketing. Retrieved from: https://geomarketing.com/most-consumers-dont-know-beacons-exist-should-that-matter-to-retailers

Lewis, P. (2016) . How beacons technology can reshape retail marketing. Retrieved from: https://www.thinkwithgoogle.com/marketing-resources/retail-marketing-beacon-technology/

Maycotte, H.O. (2020). Beacon Technology: The Where, What, Who, How and Why. Forbes. Retrieved from: https://www.forbes.com/sites/homaycotte/2015/09/01/beacon-technology-the-what-who-how-why-and-where/

Moody, M. (2015). Analysis of Promising Beacon Technology for Consumers. Elon Journal of Undergraduate Research in Communications, 6(1). Retrieved from http://www.inquiriesjournal.com/a?id=1136

Pariser, E. (2011). The filter bubble: What the Internet is hiding from you. Penguin UK.

Perez, S. 2015. Target launches Beacon Test In 50 Stores, Will expand Nationwide Later This Year. Techcrunch. Retrieved from: https://techcrunch.com/2015/08/05/target-launches-beacon-test-in-50-stores-with-expanded-rollout-later-this-year/

Raiz, G. 2018. What is Beacon Technology and how are Businesses Benefiting?. Retrieved from: https://www.rightpoint.com/thought/2018/11/16/beacon-technology

Toch, E., Wang, Y., & Cranor, L. F. (2012). Personalisation and privacy: a survey of privacy risks and remedies in personalisation-based systems. User Modeling and User-Adapted Interaction, 22(1-2), 203-220.

Vaudenay, S. (2020). Centralized or decentralized. *The contact tracing dilemma*.

Wadhwani, P. 2018. Beacon Technology Market Set to Surpass $25 Billion by 2024. RFID Journal. Retrieved from: https://www.rfidjournal.com/beacon-technology-market-set-to-surpass-25-billion-by-2024

3 Case study 2: Algorithmic personalisation – the case of Netflix

3.1 Objective

Algorithmic personalisation is used by a wide range of businesses to improve content and user-experience. This case study focuses on the algorithms used by Netflix, a media company that has algorithmic personalisation as their main model to show relevant content and optimize their user-engagement. This case-study offers an introduction of the most relevant ethical and legal issues of algorithmic personalisation towards consumers, and invites to consider the issues and limitations of these technologies.

3.2 Technology description

Netflix is a well-known media service provider, the company's main business is a subscription-based streaming service which offers online streaming for movies and TV-shows, including their own productions, and it also distributes content to all over the world. Netflix has many technological features, for example search algorithms that allows to discover new connections with natural language processing and text analytics. This case study focuses on one of the most important Netflix's technologies, the personalised algorithms, also known as recommendation systems. The importance of this personalisation is enormous, where 80% of hours streamed by Netflix' user is determined by their recommendation algorithms (Gorgoglione et al., 2019). Netflix seeks to create a fully personalised page experience, one that offers movies and TV shows that are relevant to each user's preferences, first by recommending videos that are likely to match, and the company also organise those videos into personalised rows that can fit current mood and context, and even personalised images that represent each video. In an interview, Xavier Amatriain, the vice president of personalisation algorithms, told the magazine Wired:

We know what you played, searched for, or rated, as well as the time, date, and device. We even track user interactions such as browsing or scrolling behavior. All that data is fed into several algorithms, each optimized for a different purpose. In a broad sense, most of our algorithms are based on the assumption that similar viewing patterns represent similar user tastes. We can use the behavior of similar users to infer your preferences (Vanderbilt 2013)

This is done by machine learning' algorithms that learn from their users. Netflix argues that they use personalisation for "helping members discover content they'll love"[11], and that personalisation is one of the "pillars" of their business because it allows each user to have a particular view of their interests and it even helps them expand their interests over time:

Personalisation starts on the homepage but also extends out across the product and beyond, such as deciding what messages to send our members to keep them informed and engaged. We want our members to spend less time looking for something to watch and

[11] Netflix website: https://research.netflix.com/business-area/personalisation-and-search

more time watching something they truly enjoy: an old favorite to rewatch or a new pick from our growing portfolio of original content (Netflix Website)

The technology behind personalisation is based on a combination of different algorithmic approaches that include machine learning (ML) and recommendation algorithms. These are often being improved through online but also offline experiments, the company even holds a contest to boost the company's algorithmic personalisation (Hallinan & Striphas 2016). ML algorithms on Netflix are numerous and complex, where they mention the following types of algorithms: online, collaborative, neural, Bayesian, bandit, and combinatorial[12] (Netflix Research Blog). This study is limited for the complexity of them, and will mainly give an overview of the most relevant aspects of these algorithms.

First, **"Jump starting"** (idem) is the initial analysis that a user encounters when opening a Netflix account. The user has to choose between different titles that will be analysed by the algorithm to send personalised recommendations. However, the algorithm develops with the user over time, and the most recent watches have more weight over the initial preferences, thus making the most recent watches drive the recommendations.

Rows and rankings are personalised based on complex systems of algorithms that present an ordering based on three layers: the choice of the row (continue watching, trending now, award-winning comedies…), then titles that appear on the row and the ranking of these titles, the strongest are at the top, and are based on studies that argues that the first on the left will be the most noticeable (unless the language selected is Arabic or Hebrew).

Artwork personalisation[13] refers to personalised algorithms for visual displays of the movies. The visual representation is extremely important for the user as it steers the attention of Netflix recommendations, or in their words, it "gives you some visual "evidence" for why the title might be good for you… If we present that perfect image on your homepage… then maybe, just maybe, you will give it a try" (Netflix Technology Blog). The company personalises the experience based on user's viewing history[14], for example, the algorithm analyses someone who has watched many romantic movies can be more interested in the movie Good Will Hunting if the artwork displays a couple (Figure 3), whereas a user who watched comedies might be more attracted to one that show a comedian.

[12] Netflix Research Blog: https://research.netflix.com/business-area/personalisation-and-search

[13] This is related to Paul Kuyer's case study of website morphing
For more information, visit: Netflix Technology Blog. Artwork Personalisation at Netflix. https://netflixtechblog.com/artwork-personalisation-c589f074ad76

[14] More information on Netflix Blog: https://help.netflix.com/en/node/100639

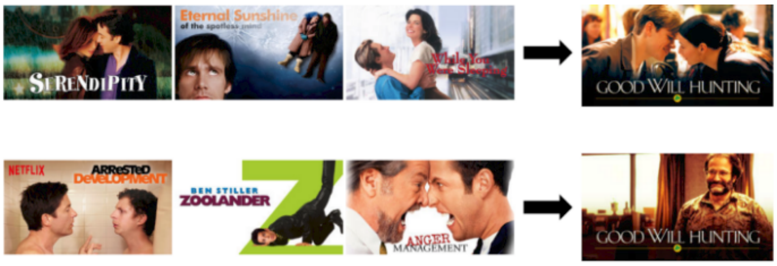


Figure 3 retrieved from Netflix Website – on the left are presented artworks that a user watched previously, the right of the arrow represents the artwork chosen by the algorithm for a new recommendation

For artwork personalisation, contextual bandits are a class of online learning algorithms that are trained for retrieving data from the user such as the member, title, image, and if the selection resulted in a play of the title or not, they also distinguish between how it results in a quality engagement (not only clicks but actual screen time). The data retrieved is usually gathered by the inputs previously mentioned, but also their country, language preferences, device that are using, time and day of the week. Furthermore, the algorithms also test offline whether a user will be engaging with certain titles, and they compare it with the online behaviour.

Netflix also contains **marketing personalisation**, the company relies on algorithms to send personalised advertisements and news to people, including personalisation of budget. Here algorithms are applied to send relevant messages to users, for example they send emails and notifications about new recommendations. Even more, they also promote budget allocated algorithms that decide “what to advertise, to whom, and for how much”. To do this, they employ a variety of ML algorithms and statistical techniques such as “Causal Models, Contextual Bandits, and Neural Networks”.

Overall, the company argues that their recommendation systems do not include demographic information (such as age and gender) as part of the decision making process[15] and that their recommendation systems are constantly re-trained to improve accuracy. The recommendations are based on user’s data, but also includes ML which studies user’s similarities and fits them into categories or “test groups” around the world (Plummer 2017).

3.3 Context

There are many examples of personalisation and recommendation’ algorithms used by different companies such as Twitter, Facebook, Spotify, Instagram or Amazon. In the previous section was also mentioned different uses of personalisation. The type of personalisation differs between companies, thus the model of business matters from which the data is retrieved and there is a wide range of possibilities. This study focuses on a specific company, Netflix, to investigate the most pressing issues related to personalised algorithms. Understanding the methods and techniques that Netflix uses to personalise

[15] Netflix Blog : https://help.netflix.com/en/node/100639

user-experience, will help to unveil ethical and legal concerns, the limitations and problems that this type of personalisation brings.

3.3.1 Potential Use

The possibilities of the algorithmic personalisation lead to new uses in Artificial Intelligence (AI) and recommendation algorithms. The pattern to increase the level of personalisation becomes obvious, and in the future it will be more personalised, if possible, with the arrival of AI. This technology offers the possibility to find new patterns of behaviour and improve personalisation (Yu 2019). In their Blog, Netflix argues that there are many possibilities to continue expanding this personalised approach in the future, such as also include algorithms in synopses, metadata and trailers (Netflix Technology Blog). Ironically, in a Buzzfeed article, it is argued that “Netflix has no chill, so you can” (Nguyen 2018), meaning that Netflix is always looking for ways to expand and improve their algorithms. Similarly, other authors argued that it is expanding from behavioural targeting to deep learning, content personalisation and user engagement exponentially (Khandelwal 2020).

3.4 Relationship to Workpackage

This workpackage (WP) focuses on personalisation, the use of personalisation is an essential part of Netflix, shown in their business model as a personalised streaming service, which makes it a relevant case study in this WP. I use the case of Netflix to investigate the influence of algorithmic personalisation in business practices and to study their ethical and legal implications. There are limitations to these personalisation practices and issues that must be analysed.

3.5 Relationship to ESR PhD topic

My PhD topic focuses on digital surveillance and will study different forms of surveillance. This case study focuses on algorithmic personalisation that is a form of business surveillance.

3.6 Review of ethical and legal issues

3.6.1 Methodology

For this study, a critical analysis was used to identify relevant literature on Netflix Technology. Firstly, Google Search was used as a main source for non-academic literature such as newspapers, blogs and other relevant literature. Secondly, Google Scholar and Scopus was used for academic articles. Finally, the legal research consists on the GDPR and lawsuits that were relevant in California (US) where Netflix started, these were also retrieved by Google Search. The terms that were used in this search engines to conduct the research were divided into search words: “Netflix” “Netflix Algorithms”, “Netflix Personalisation”, “ethics” “moral” “issues”, “privacy” “law”. The findings were prepared on the basis of a critical selection of the most relevant articles and this analysis will provide a comprehensive summary of the ethical and legal issues related to Netflix personalised algorithms.

3.6.2 Findings

This technology has many benefits, among them, Netflix uses personalised algorithms to give recommendations to people that have a high probability of being relevant. It provides members with a content that they will enjoy and it helps them to find new movies or TV-shows to watch. It has become evident that Netflix algorithms have positive value for the company and for the user, as Spiegelman, the VP of Netflix, argued “If you make people find more things that they want to watch, then they will get more via the service and they’ll be inclined to stay,” (Nguyen 2018). On the contrary, this case-study looks into the issues pertaining these personalised algorithms, while acknowledging that it has positive value. Personalised algorithms are used by multiple companies, and it is urgent to look at the ethical and legal issues related to this technology.

Netflix strated in 1997 as a DVD provider, however it is not until 2007 that Netflix has become a streaming service online, and only in 2012 it becomes a world-service[16]. The novelty of this company and the constant change of the algorithms does not allow for is no extensive literature written about the ethical issues of the algorithms. Here it will be shown some of the most prominent issues.

Netflix not only personalises contents, but also keeps records of all choices and send specific messages. Thus, **privacy issues** become at the top of the list, where personalised algorithms based on user’s data retrieval can lead to privacy invasions. An example of this is that even if Netflix does not retrieve data such as race or gender, the company still knows something far more personal, what your actual taste in movies and TV is, what is called “aggregated data”. Netflix’s twitter account wrote last year “To the 53 people who’ve watched A Christmas Prince every day for the past 18 days: Who hurt you?”[17] (Nguyen 2018). Companies know plenty of sensitive information from aggregated data, and that data can be also hacked and leaked (Tingley 2019). Even more, data can be sold to third parties without user’s consent (Wang et al. 2018). It is essential to develop privacy impact assessments to avoid such violations. The data used by these systems can be divided into implicit and explicit, where the implicit data is the one feed to the system by the user, however the explicit is the one that is deduced by algorithms, for example a user might not be aware that is prone to movies that portrait strong women, but the algorithm *understands it* from the movies that have been watched. This type of data requires special attention and intrudes into people’s privacy rights (GDPR), even if the data is anonymized, the danger to de-anonymization exists (Narayanan & Shmatikov 2008). Narayanan and Shmatikov identified several Netflix users by comparing their "anonymous" reviews in the Netflix data to ones posted on the Internet Movie Database website, a third party that contained sensitive information, identifying their political leanings and sexual orientation, and Netflix obtained a lawsuit because of this reason in California (Singel 2017).

[16] See full Netflix history here : https://media.netflix.com/es/about-netflix

[17] Netflix’s twitter link: https://twitter.com/netflix/status/940051734650503168?s=20

Related to privacy issues, there is a **lack of explanation and transparency** about data gathering. Although Netflix's blog is open and contains some information about personalisation, an average user cannot fully understand how personalisation is done based on their privacy policy. One of the most famous scandals from Netflix is that the company kept records of the choices the users made while watching the interactive movie Bandersnatch, a Black Mirror film that supported a new scenario where users decide from different choices how the adventure must go, what is called "choose your own adventure". The policy researcher Michael Veale requested Netflix data[18] under the GDPR right of access rules, in order to educate people on how to request their data access, and he exposed that the company never asked for permission to store user's data neither the users knew about this (Epstein 2019). Collecting data is a common activity for companies, but more efforts should be made to ensure that users know what is happening, and Veale also argued they should be able to opt-out of the practice if users desired to do so. There is something paradoxical about this[19], because the movie Bandersnatch is about a person that becomes paranoid for being tracked and watched. This can be extrapolated to any type of data retrieved by algorithms, because users do not know how their information is stored and further personalised into this website. In this line, more efforts could be made about informing users about their rights and the possibility to delete their history from Netflix website (Keller 2018).

However, there are also issues about the **complexity of algorithms to have transparency**, as algorithms are complex systems containing blackboxes that are difficult, if not impossible, to explain. Netflix uses neural networks, which are an example of the complexity of these systems, even more to explain them users that are not experts in the field. The neural networks consist of algorithms that mimic human brains for pattern recognition, clustering and classification (Nicholson 2019), it results into millions of patterns that are interpreted by the system but that are difficult for a human being to understand.

Algorithmic personalisation can be a **persuasive technology** (can be also called addictiveness or coercion), how rows are tailored to your viewing habits, and how algorithms learn from the user, the more up to date an algorithm is, and the more it is tailored to you and the more you watch (DeLeon 2016). Moreover, the magazine Wired also mentions how ML algorithms are used to break user's preconceived ideas and show titles that the user might initially never thought of watching (Plummer 2017). Netflix persuades users into watching more shows and episodes by using personalised offers, therefore the company takes advantage of a technology that make users more addicted. A bad star ratings, for example, can no longer dissuade users from watching, given that Netflix removed its global five-star rating system and decades of user reviews: "Now, users rate

[18] See the Twitter threat where Veale exposes Netflix: https://twitter.com/mikarv/status/1095110948908662784

[19] https://www.theverge.com/2019/2/13/18223071/netflix-bandersnatch-gdpr-request-choice-data

content only themselves with thumbs up/down icons, and a very specific, personalised "match rating" — 98%, 81%, 62% — has taken the star ratings' place" (Nguyen 2018). The algorithmic personalisation of the artwork display can be misleading and there is an issue with the **necessity of the algorithms**, the usefulness of an algorithm for personalisation cannot be taken for granted, e.g. that the personalised algorithms are the reason why a user clicks at a movie, if a member clicks a title it can only come from the one displayed by the algorithm, what it is also called "chicken and egg" problem (Burruss 2020), there is no other option that the one the algorithm have displayed. Netflix also argues in their blog that in order to learn how to personalised artwork they will need to retrieve a lot of data to prove one piece of artwork is significantly better for a user. Furthermore, confusion[20] is also an issue that come with the development and rapidly-changeable algorithmic technology. Artwork can be changed in between sessions based on the algorithmic decisions that a user might prefer a different image, however the continuous changeability can confuse people; this issue can be related to the necessity of algorithms, because the constant change impact the effect of the displays in clicking the titles. Algorithm artwork display cannot be always followed, an article by the Wall Street Journal, an algorithmic study for the comedy *Grace and Frankie* found that users clicked more on the movie if there was no images of the actress Jane Fonda on the artwork. However, Netflix was under pressure to maintain good relationship with the actress and the content team asked the leaders in charge of the algorithms to reconsider, and the Fonda images were placed back into the artwork (Nguyen 2018).

Numerous scholars have highlighted problems related to algorithms that use historic data such as the **Filter Bubble** (Pariser 2011) in which algorithms give you the same type of information based on previous history, which results in the user getting always similar information (lack of variety) and thus not stepping out of their bubble. This can lead users to always watch similar content, and push users into watching certain type of genres. However, the study of Netflix algorithms has revealed the complexities of these systems, in which they also use different strategies and relate content with recommendations that the user never thought of before (Plummer 2017).

Algorithms also contain **social biases**, similar to the question regarding the usefulness of the system, that can contain for example gender biases (Zarum 2018) (Spandana 2020), in the New York Times article and the explicit data, it was shown that although algorithms do not retrieve race specifically, they still target them based on previous search and deduce race, therefore showing mostly the same race. Questions appear about the **fairness of the system** and how to avoid this type of algorithmic biases. Many Netflix users were concerned with recommendations based on race (Ibqal 2018), which can result in discriminatory outcomes. As an example, a Twitter user noted that Netflix was generating artwork featuring black cast members[21], but for movies in which those black actors had minor roles.

[20] "Confusion and usefulness" are ethically relevant but not necessarily or directly ethical issues. They are ethically concerning or problematic, and affect the "necessity" of algorithms.
[21] Twitter link: https://twitter.com/BCMorrow/status/1036049714620325888?s=20

Even if Netflix does not retrieve data from race and gender, social biases can appear by trying to influence user's based on their data which contains biases.

3.7 References

Burrus, M. 2020. Expectation-Maximization (EM) Algorithm: Solving a Chicken and Egg Problem. Retrieved from: https://towardsdatascience.com/solving-a-chicken-and-egg-problem-expectation-maximization-em-c717547c3be2

DeLeon, Haley. (2016). The Ethical and Privacy Issues of Recommendation Engines on Media Platforms. Retrieved from: https://towardsdatascience.com/the-ethical-and-privacy-issues-of-recommendation-engines-on-media-platforms-9bea7bcb0abc

Epstein, M. 2019. Netflix's 'Black Mirror' Creates Same Privacy Problems it Warns Against. RealClear. Retrieved from: https://www.realclearpolicy.com/articles/2019/01/11/netflixs_black_mirror_creates_same_privacy_problems_it_warns_against__110983.html

Gorgoglione, M., Panniello, U., & Tuzhilin, A. (2019). Recommendation strategies in personalisation applications. Information & Management, 56(6), 103143.

Narayanan, A., & Shmatikov, V. (2008, May). Robust de-anonymization of large sparse datasets. In 2008 IEEE Symposium on Security and Privacy (sp 2008) (pp. 111-125). IEEE.

Nguyen, N. 2018. Netflix Wants To Change The Way You Chill. Buzzfeed. Retrieved from: https://www.buzzfeednews.com/article/nicolenguyen/netflix-recommendation-algorithm-explained-binge-watching

Pariser, E. (2011). The filter bubble: How the new personalised web is changing what we read and how we think. Penguin.

Plummer, L. (2017). This is how Netflix's top-secret recommendation system works. WIRED. Retrieved from: https://www.wired.co.uk/article/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like

Hallinan, B., & Striphas, T. (2016). Recommended for you: The Netflix Prize and the production of algorithmic culture. New media & society, 18(1), 117-137.

Iqbal, N. (2018). Film fans see red over Netflix 'targeted'posters for black viewers. The Guardian, 20.

Keller, J. (2018) How to clear your viewing history in Netflix. iMore. Retrieved from: https://www.imore.com/how-clear-your-viewing-history-netflix

Nicholson, C. (2019) A.I. Wiki: A Beginner's Guide to Important Topics in AI, Machine Learning, and Deep Learning. Pathmind. Retrieved from: https://pathmind.com/wiki/neural-network#:~:text=Neural%20networks%20are%20a%20set,labeling%20or%20clustering%20raw%20input.

Spandana, S. (2020). Why Am I Seeing This? How Video and E-Commerce Platforms Use Recommendation Systems to Shape User Experiences. Retrieved from: https://d1y8sb8igg2f8e.cloudfront.net/documents/Why_Am_I_Seeing_This_2020-03-25.pdf

Singel, R. 2017. Netflix Spilled Your Brokeback Mountain Secret, Lawsuit Claims. Wired. Retrieved from: https://www.wired.com/2009/12/netflix-privacy-lawsuit/

Tingley, B. 2019. Netflix User Data Begins to Raise Security and Privacy Concerns. Soda. Retrieved from: https://www.soda.com/news/netflix-user-data-security-and-privacy-concerns/

Vandervilt, T. (2013). The Science Behind the Netflix Algorithms That Decide What You'll Watch Next. WIRED. Retrieved from: https://www.wired.com/2013/08/qq-netflix-algorithm/

Wang, C., Zheng, Y., Jiang, J., & Ren, K. (2018). Toward privacy-preserving personalised recommendation services. Engineering, 4(1), 21-28.

Yu, A. 2019. How Netflix Uses AI, Data Science, and Machine Learning — From A Product Perspective. Medium. Retrieved from: https://becominghuman.ai/how-netflix-uses-ai-and-machine-learning-a087614630fe

Zarum, L. 2018. Some Viewers Think Netflix Is Targeting Them by Race. Here's What to Know. The New York Times. Retrieved from: https://www.nytimes.com/2018/10/23/arts/television/netflix-race-targeting-personalisation.html

4 Case study 3: Personalised political advertisements – The case of algorithms in Facebook

4.1 Objective

Recent scandals such as Cambridge Analytica[22] confront us with the issues of the use of technologies to steer people's political opinions. This case study focuses on algorithms used by Facebook to personalise political campaigns. The use of these algorithms will be studied to share light on the ethical and legal issues of the use of social networks for the personalisation of political advertisements.

4.2 Technology description

Political campaigns nowadays use social networks and their respective algorithms to find specific voters, their political beliefs and ideals, and target them with personalised message to influence their votes. Facebook (FB) is one of the most used social networks for political advertisement (Wadhwa 2015). This case study focuses on the algorithms and practices that are used to personalise political advertisements. A thorough study has been conducted on Facebook the past year (Ali et al. 2019) by a team of researchers that used real political ads and by being undercover, they payed FB to promote them. The ads were aimed to two categories, these are centred in US, by being supporters of Trump and Bernie Sanders, and the researchers used same parameters to deliver both ads, hypothesising that both ads will get the same audience (Trump and Sander's ads should go to both conservatives and liberals). This research will be used to investigate the ethical and legal consequences of the technology used on FB for political campaigns.

The authors mention two steps to deliver political ads: *ad creation* and *ad delivery*. First, the *ad creation* is when the advertiser provides FB with the content of the ad and specifies the *target audience*, an *objective* and a *budget*. Secondly, the *ad delivery* refers to the process by which the platform decides which ad will be displayed to which user, as they argue "before displaying an ad to a user, the platform will hold an ad auction to determine which ad, from among all ads that user is eligible to see" (Idem p.3). There are different ways for advertisers to target ads, such as the data from user's demographics and interests online, but also offline behaviour, as the authors point out "often acquired without user's explicit consent or knowledge" (Idem p.3). In this context of political propaganda, FB derives characteristics that indicate whether a user is interested in different political candidates as well as general behaviour such as "likely to engage with US political conservative content". The authors interestingly point out that the exact methodology by which FB infers certain characteristics from users is not disclosed to the public, but it is argued that it likely involves data processing from users and algorithms, e.g. pages or interaction with specific content. In addition, FB also customize audiences and allows advertisers to target users directly using *custom audiences* (Venkatadri et al 2018).

[22] For more information about Cambridge Analytica visit the article of The Guardian: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

The *objective* for the *ad creation* is important, the advertisers have to distinguish between some common goals such as "reach" (showing the ad to as many users as possible), "traffic" (showing the ad to those who will be most inclined to click) "App Installs" (showing the ad to the users most likely to download the advertiser's app) (Idem p.4). Within each objective, advertisers must also specify *optimization,* for example with the objective "reach" the optimisation are: "reach" (by default, showing ads to as many users as possible), "impressions" (showing ads as many times as possible)[23], or "Link Clicks" (getting as many people as possible to visit their campaign page). Finally, the *budget* also makes a difference. When creating an ad, the advertiser must tell FB their *ad budget*: "These budgets are either a daily or lifetime budget for the ad, allowing FB to spend their money… according to an algorithm that is nor publicly known" (Idem p.4).

After the *ad creation* phase is finished, it is submitted to FB for *ad review*. This process consists of a combination of automated and manual review to prevent abuse or violation of advertising policies[24] however the authors argue that the processes and mechanisms that the company uses to accept the ads are not clear nor precise, and the reasons why FB reject some ads are not explained to the advertisers. It is important to highlight that the *ad delivery* is not merely influenced by the bid price, but FB considers other personalised features[25] such as the performance of the ad and the platform's estimation of how relevant the ad is to the user, it includes feedback of how people interact with the ad, such as if they hide it (there is a possibility to hide advertisements on FB). Facebook's COO Sheryl Sandberg argued that this model of targeted ads, including the use of sensitive information, brings great benefits to small businesses and allows the social network to continue being free (Salinas 2018).

Furthermore, the *ad delivery* refers to the FB algorithms that decide which users see which ads (Matsakis 2019). FB also provides detailed statistics to advertisers about how their ad is being delivered, such as the number of times the ad was shown (impressions) and the number of unique users who saw the ad (reach), and the number of times users clicked on the ad (clicks) (Ali et al. 2019 p.5). These statistics contain information about gender, age, and location. However, the authors highlight that FB did not reveal information about people's political ideas, but still took them into account.

FB states in their rules for advertisers that if an ad is relevant to a person it will be considered more important than higher bids: "[W]e subsidize relevant ads in auctions, so

[23] To see the complete list of objectives, go to the article by Ali et al. 2019 page 4.

[24] Facebook Ad review process: https://www.facebook.com/business/help/204798856225114?id=649869995454285

[25] Facebook: About the delivery system: Ad auctions. https://www.facebook.com/business/help/430291176997542

more relevant ads often cost less and see more results. In other words, an ad that's relevant to a person could win an auction against ads with higher bids"[26]. In this line, the research has shown that FB algorithms make it harder and more expensive for campaigners to get its messages delivered to users who do not in principle, agree with them, even if the campaigners do not want to use information such as racial or gender data "Facebook delivers our ads with content from Democratic campaigns to over 65% users registered as Democrats, while delivering ads from Republican campaigns to under 40% users registered as Democrats, despite identical targeting parameters" (Ali et al. 2019, p.2). Thus, FB allows to design ads' audience but the algorithms decide which user will see the ad based on the user's interests (Edelman 2019).

4.3 Context

There are different social networks that run political ads such as Twitter, Instagram or Facebook. However, last year Twitter decided to change its policy and ban all political advertising from its platform (Conger 2019). This case study focuses on Facebook's algorithms to investigate how this technology personalise users' political ads and what are its ethical and legal issues.

4.3.1 Potential Use

The possibilities of FB personalised algorithms for political campaigns are very specific and centred in the organisation of political campaigns. It is exposed that the use of algorithms to find possible voters and influence them can only increase in the future of politics (Bartlett 2018). Although Google and Twitter have banned political targeting, FB is still remaining neutral and allowing this to happen (Roose 2020). New technologies such as AI and beacons[27] can be merged with personalised algorithms for political campaigns and make the personalisation even stronger.

4.4 Relationship to Workpackage

Facebook retrieves users' data to target specific political advertisements, this has proven to be a method to personalise advertisement through algorithms and data analytics. This WP focuses on personalisation technologies, and therefore FB algorithms are a personalised technology for political advertisements.

4.5 Relationship to ESR PhD topic

The PhD topic focuses on the digital surveillance of individuals. This case study offers the possibility to investigate the surveillance of individuals to steer political opinions. While it focuses on FB algorithms, this case also helps to investigate the possibilities for political parties to surveil citizens.

[26] Facebook: About the delivery system: Ad auctions: https://www.facebook.com/business/help/430291176997542?id=561906377587030

[27] See case-study on Beacons by Ana Fernandez Inguanzo

4.6 Review of ethical and legal issues

4.6.1 Methodology

For this study, a critical analysis was used to identify relevant literature on Facebook technology used to personalise political advertisements. Firstly, Google Search was used as a main source for non-academic literature such as newspapers, blogs and other relevant literature. Secondly, Google Scholar and Scopus was used for academic articles. Similarly, the legal research focused on the finding based on Google Search. The terms that were used in this search engines to conduct the research were divided into search words: “Facebook” “Facebook Algorithms”, “Facebook Political Advertisements”, “ethics” “moral” “issues”, “privacy” “law”. The findings were prepared on the basis of a critical selection of the most relevant articles and this analysis will provide a comprehensive summary of the ethical and legal issues related to Facebook personalised algorithms.

4.6.2 Findings

The description of the technology (section 1.1.2) has shown that FB plays a significant role in determining which users see which ads, based on its own decisions about which ads are most likely to be relevant to particular users. This technology has positive attributes, such as keep people informed about elections and to send relevant content to the users, thus not bothering with the wrong information to people that are not interested in. The nature of this information is very sensitive and to send wrong information might result in users being angry. FB helps to deal with personalised political campaigns. However, this case-study only deals with ethical and issues, while acknowledging the benefits of this technology.

FB is delivering ads to users that are more likely to be related to the campaign’s political views, which can create a “political polarization” also called an informational filter bubble (Ali et al. 2019, p.2) (Edelman 2019). This polarization brings issues such as **unfair political campaigns** (Isaac & Kang 2020) by means of personalised political ads, and highlights the importance of social networks in the political role in society: “Political advertising cuts to the heart of Facebook’s outsize role in society, and the company has found itself squeezed between liberal critics, who want it to do a better job of policing its various social media platforms, and conservatives, who say their views are being unfairly muzzled” (Idem). This raises questions such as how companies like FB decide how much political content they permit; e.g the Democratic ads are delivered to 65% of users registered as Democrats, while Republican ads to under 40% of users registered as Democrats. They have also shown that it is more expensive for political campaigns to have their content delivered to the users that FB considers not related with a particular political opinion, e.g they show how to conservative users, Sander’s ads (progressive party) were sent to fewer people than Trump’s ads (conservative) (Idem p.2). FB define targeted audiences, introducing demographic and political biases in the reached audiences, even beyond those intended by the advertiser. This limits people’s exposure to different political viewpoints and raise serious concerns about how FB and other platforms are, in fact, creating filter bubbles and making it more expensive for campaigners to change this:

Facebook is making decisions about which ads to show to which users based on its own priorities (presumably, user engagement with or value for the platform). But in the context of political advertising, Facebook's choice may have significant negative externalities on political discourse in society at large (Idem p.3).

On the contrary, there are opposing views arguing that showing users political ads that they do not support appears to increase their intolerance (Bail et al 2018). While the US Federal Election Commission argued that instead of banning political ads, the platform should limit or modify the capacity of "microtargeting", and ensure that a broader population can be informed (Weintraub 2019). Google also announced[28] that it will significantly reduce election ad targeting in order to promote visibility for all political ads. On the contrary, Zuckerberg has claimed that the freedom of speech is also important and that he "did not want to be in the position to police what politicians could and could not say to constituents. Facebook's users, he said, should be allowed to make those decisions for themselves" (Isaac & Kang 2020). FB argument refers to being a private company, and that they should not have to censor any politician (Sullivan 2019).

Lack of unified regulation is an important issue, in the absence of a regulation, FB and other companies are left to design their own privacy policies (Dave 2019) (Friedersfdorf 2019). Under pressure from authorities, FB in 2018 introduced new initiatives to oversight political ads, where the director of product management Rob Leathern affirmed that they are learning from every country and that one of their solutions are that "Facebook believes that holding the ads in a library for seven years is a key part of fighting interference" (Idem). While GDPR in Europe refers to the use of personal data and dictates norms to their use, political advertisements go further and relates to the democratic values of society. This is an "ad hoc approach" (Dave 2019) with different policies and transparency depending on the region, referring to local laws and governments and civil society groups.

Lack of transparency is also an ethical and legal issue, the GDPR in Europe refers to transparency as an important part of the regulation. However, there is a necessity to inform the population about targeted ads in political campaigns, while the average voter might not even understand the consequences of running political ads and targeted campaigns. In the previous section (1.1.2) it was shown that FB's algorithms are not publicly available, and thus the inferences about people are not specified, such as, what type of algorithms explain who is most likely to click? Users and neither advertisers have the possibility to review how this is being done. The author's results suggest that FB limits this possibility and has significant power over political discourse through ad delivery algorithms without **public accountability** or scrutiny (Ali et al. p.13). In addition, they rightly argue that "researchers, regulators, and campaigns lack access to algorithms and data required for a more thorough study of ad delivery skews and their likely impacts….although much has already been said

[28] Google Blog (2019). An Update on our political ads policy. https://www.blog.google/technology/ads/update-our-political-ads-policy/

about the inadequacy of current ad transparency tools provided by ad platforms[29] our work draws attention to the need to expand these efforts to account for ad delivery algorithms as well" (Idem). This has similar repercussions in accuracy.

Related to the unfair political campaigns, there is a **lack of algorithmic fairness**, the concern about how platforms show "relevant" ads to users may raise issues. It was exposed how FB influences the target audience, for example "ads targeting the same audience but with different content can be shown to over 95% women or less than 15% women, depending only on the content of the ad and not on the advertiser's targeting choices or competition from other advertisers" (Idem p.5). The authors rightly point out that in some ads, this might be desirable, however in the field of political ads this relevancy can raise serious issues such as equality. Similarly, **social biases** are also present when FB use their own political targeting features such as "likely engagement with conservative US political campaigns", in this case the authors argue that democratic ads are sent to 60% liberal users and only 25% of republican campaigns to the liberal users (Idem p.2). Even when companies choose to show their ads to inclusive audiences, FB can deliver it more to men than women (Matsakis 2019).

In this line, **control and manipulation** is also an issue to consider. Matz et al. (2017) conducted a large-scale experiment in which they showed that FB ads tailored to individual's psychological characteristics brings higher click-through and conversion rates compared to non-personalised ads and mismatching ads. It is not clear whether personalised political ads are any different from any other commercial ad such as unhealthy drinks (Friedersdorf 2019) but the ability of FB algorithms to match people's political ideas with similar content raise questions of control, similar to the unfairness of political campaigns. In relation to this, FB refused to fact-check their political ads (Ingram 2019), and this bring also moral issues of manipulation and control (Sullivan 2019), however this falls out of the scope of this case-study focus on personalisation technologies.

There are **implications for the future** in regard to personalised political advertisement. It has been shown how FB has enormous power over how people obtain political discourse. It is also likely that this will only continue to grow in the future. It has been made clear that a unified regulation of digital and political advertisement is needed and transparency requirements will enable a better understanding and research about ad targeting and the delivery of political ads. The public (users) and the campaign managers need more information (transparency) about the operations in ad delivery algorithms and their real-world effects. Ad platforms should increase transparency around political ads. Appropriately, the authors argue that "Ad platforms could also disable delivery optimization for political content, or at least allow advertisers to do so. They could also introduce more nuanced user-facing controls for political content delivery and expand public ad archives to make them more accessible and usable by everyone" (Ali et al. 2019, p.14).

[29] Facebook's Ad Archive API is Inadequate. The Mozilla Blog. Retrieved from: https://blog.mozilla.org/blog/2019/04/29/facebooks-ad-archive-api-is-inadequate/

4.7 References

Ali, M., Sapiezynski, P., Korolova, A., Mislove, A., & Rieke, A. (2019). Ad Delivery Algorithms: The Hidden Arbiters of Political Messaging. arXiv preprint arXiv:1912.04255.

Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., & Rieke, A. (2019). Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes. arXiv preprint arXiv:1904.02095.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. Science, 348(6239), 1130-1132.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. Proceedings of the National Academy of Sciences, 115(37), 9216-9221.

Bartlett, J., Smith, J., & Acton, R. (2018). The future of political campaigning. Demos.

Conger, K. 2019. *Twitter Will Ban All Political Ads, C.E.O. Jack Dorsey Says. The New York Times. Retrieved from: https://www.nytimes.com/2019/10/30/technology/twitter-political-ads-ban.html*

Edelman, Giglad. 2019. How Facebook's Political Ad System Is Designed to Polarize - Want to reach voters across the aisle online? That'll cost extra, a new study finds. WIRED. Retrieved from: https://www.wired.com/story/facebook-political-ad-system-designed-polarize/

Friedersdorf, Conor. (2019). Doubt Anyone Who's Confident That Facebook Should Ban Political Ads. Retrieved from: https://www.theatlantic.com/ideas/archive/2019/11/twitter-facebook-political-ads/601174/

Isaac Kang. (2020). Facebook Says It Won't Back Down From Allowing Lies in Political Ads. The New York Times. Retrieved from: https://www.nytimes.com/2020/01/09/technology/facebook-political-ads-lies.html

Ingram, Mathew. (2019). On Facebook, disinformation, and existential threats. The Media Today. Retrieved from: https://www.cjr.org/the_media_today/facebook-disinformation-antitrust.php

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. Proceedings of the national academy of sciences, 114(48), 12714-12719.

Matsakis, L. 2019. Facebook's Ad System Might Be Hard-Coded for Discrimination
The social network says it shows users the most "relevant" ads, but a new study suggests the system perpetuates bias. WIRED. Retrieved from: https://www.wired.com/story/facebooks-ad-system-discrimination/

Roose, Kevin. 2020. Buckle Up for Another Facebook Election. New York Times. Retrieved from: https://www.nytimes.com/2020/01/10/technology/facebook-election.html

Salinas, S. (2018). Sheryl Sandberg delivered a passionate, defiant defense of Facebook's business. CNN. Retrieved from: https://www.cnbc.com/2018/04/26/facebooks-sheryl-sandbergs-brilliant-defense-of-the-ad-business.html#:~:text=Sheryl%20Sandberg%20delivered%20a%20passionate%2C%20defiant%20defense%20of%20Facebook's%20business,-Published%20Thu%2C%20Apr&text=Facebook%20for%20weeks%20has%20had,to%20l

awmakers%2C%20investors%20and%20users.&text=But%20COO%20Sheryl%20Sandberg%20elevated,business%20demographic%20into%20the%20conversation.

Sullivan O', Donnie (2018). Facebook's refusal to fact-check Trump could be its defining 2020 decision. CNN. Retrieved from: https://edition.cnn.com/2019/10/10/tech/facebook-false-trump-ads-analysis/index.html

Venkatadri, G., Liu Y., Andreou A., Goga O., Loiseau P., Mislove A., (2018). Privacy Risks with Facebook's PII-based Targeting: Auditing a Data Broker's Advertising Interface. In IEEE Symposium on Security and Privacy

Wadhwa, T., 2015. How Facebook Is Shaping Who Will Win the Next Election. Huffpost. Retrieved from: https://www.huffpost.com/entry/how-facebook-is-shaping-w_b_6201652?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuc2N1LmVkdS9ldGhpY3Mtc3BvdGxpZ2h0L3NvY2lhbC1tZWRpYS1hbmQtZGVtb2NyYWN5L3RoZS1ldGhpY3Mtb2YtZ2l2aW5nLXBlb3BsZS1hLXZvaWNlLWFuZC1wb2xpdGljYWwtYWR2ZXJ0aXNpbmctb24tZmFjZWJvb2sv&guce_referrer_sig=AQAAAGuZFupIoVcySS3xW8s3uLSwsfln829UMj7CblY1vqf6FHpKK-wAQchdnfEBUB_m5V3XzE74iTNpYCfDia-ibpyJe5XGMg6V8zXpiQEcQy98vE4aj13JpEF2TOwj9bJlV0IXBaAi-4yXH5MdH0KbS7rdTeO9ms_AziNpYRDUm_Ha

Weintraub E. L. (2019). Don't abolish political ads on social media. Stop microtargeting. The Washington Post. Retrieved from: https://www.washingtonpost.com/opinions/ 2019/11/01/dont- abolish- political- ads- social- media- stop- microtargeting/.

5 Case study 4: Persuasive profiling

5.1 Objective

The aim of this case study is to investigate how persuasive profiling works and what its ethical and legal implications are. Persuasion profiling is a form of adaptive technology that seeks to systematically model which persuasive means are particularly effective for a unique individual. They are defined as “profiles that specify estimates of the effects of particular influence strategies on an individual” (Kaptein & Eckles, 2010).

For many marketers, accurate consumer segmentation is the ultimate ambition (Yeung, 2018). More accurate segmentation lead to higher relevance, which in turn lead to higher revenue and lower costs. The more a specific a marketing message is tailored to viewer tastes, preferences and desires, the greater the chance of success. In online environments with ubiquitous surveillance possibilities, the ability to collect data on user interests and behaviour is endless, enabling very precise profiling of individual users (Yeung, 2018). This technique is of course interesting for marketers and others who seek the maximise their persuasive impact on a large audience.

5.2 Technology description

Persuasive profiling is about creating user profiles that have predictive power. This involves the collection, processing of enormous amount of data. Subsequently, the profiles are used to target users to achieve any goal chosen by the designer. It is important to note that ‘behavioral targeting’ is often used as a more general concept, encapsulating all forms of adjustments to past user behavior. While there exists a broad field of literature on end adaptive personalisation and the broader term behavioral targeting, **persuasive profiling** has not received much explicit attention.

5.3 Context

In the beginning of persuasion profiling, most messages focussed on motivational orientation. Ads were framed either in terms of benefits or losses to the consumer (Hirsch *et al.*, 2012). However, persuasion profiling has become more advanced and increasingly focus on psychological modelling and digital psychometry (HBR, 2018). An important psychological model for modelling personality is the big five model. This model describes five dimensions of personality, namely extraversion, openness, neuroticism, conscientiousness and agreeableness (Gosling *et al.*, 2003). This model has been extensively tested and successfully used in many studies (Zhao & Seibert, 2006). It has also been largely replicated in various cultures (Rolland, 2002). However, the big five is met with scepticism and it has been reported that Facebook “segments users by many more dimensions than do models such as the Big Five” (Eckles; nature, 2018).

5.3.1 Potential use

With psychographic models in hand, marketeers can build links between personality traits and online behaviours. At the same time, “relating behavioural science to big data” constitutes a challenge for many engineers (Mills, 2019). One way to tackle this challenge is to infer psychological characteristics from Facebook likes, tweets, browsing histories and YouTube histories (HBR, 2018). Behavioral data that are collected of a person are called their ‘digital footprint’ (Matz *et al*., 2017). From the traces people leave digitally, it is possible to construct a relatively accurate image of their personalities. It has been found that based on 200 Facebook likes or more, algorithmic modelling can better predict the five personality traits of a person than their spouses (Youyou *et al*., 2014).

1.1.4 **Relationship to Workpackage**

Persuasion profiling is strongly connected to personalisation (the title of the workpackage) It is important to note that personalisation can occur on two levels. First, a product, service or information can be personalised, like in recommender systems. This form of personalisation is called either ‘choice personalisation’ (Mills, 2019), or ‘end adaptive’ personalisation (Kaptein & Eckles, 2010). Alternatively, a message can be personalised. This form of personalisation is called either ‘delivery personalisation’ (Mills) or ‘means adaptive’ personalisation (Kaptein & Eckles). Persuasion profiles are a method to model personal persuasion styles and techniques. They are thus a form of means adaptive personalisation.

5.4 Relationship to ESR PhD topic

Persuasion profiles can be seen as a way of personalising digital nudges (the topic of my PhD). The techniques selected to deliver personally through persuasive profiling are most often nudges, such as framing (Kaptein *et al*., 2015). Persuasion profiling thus facilitates personalised digital nudging and makes it more effective .

5.5 Review of ethical and legal issues

5.5.1 Methodology

For this case study a literature search is conducted in three academic databases, namely Google Scholar, Scopus and Hein Online. I first searched with the search string “persuas* profil*” in HeinOnline and the same search string combined with the additional set (‘ethic*’ OR ‘moral*’ OR ‘virtue*’ in the other two databases. When these searches did not result in sufficient results, I added the search strings “psycho* profil*”, “behavior* profil” and “behavior* target*”. This review is exploratory in the sense that it aims to provide a clear image of the ethical and legal issues of persuasive profiling but in no way aims to offer a complete overview of the entire literature on the topic.

5.5.2 Findings

Not surprisingly, psychological profiling of advertisements has received severe criticism and suspicion. Kaptein & Eckles note that persuasive technologies in general stand on “uneasy ethical ground” (Kaptein & Eckles, 2010, p 9). Especially of personalised persuasive technologies, many concerns have been brought forward (Mills, 2019). I will now discuss the three most important and often cited ethical and legal concerns of persuasive profiling.

Privacy

The first concern which is both ethical and legal is that persuasive profiling could lead to the infringement of privacy. Privacy is often conceived to regard only personal data (Mavriki & Karyda). However, it should be extended to include constructed data, as in (persuasive) profiling (idem). The construction of user profiles can have many adverse effects, such as: "losing the ability to shield intimate and personal details of their private lives", "embarrassment from the unexpected disclosure of details", "identity theft or other forms of financial fraud" and "the unexpected use of a consumer's profile to make adverse decisions about how to treat her" (Berger, 2011, p 18). Due to the specific nature of persuasive and psychological profiling, privacy conceptions that merely look at which data are disclosed are too limited. Psychological profiling requires us to look at "how the data are being used, that is, in which context and for which purpose, becomes crucial" (Matz *et al.*, 2020, p 118).

Manipulating and influencing choices harm a specific form of privacy, namely decisional privacy (Zarsky, 2019). Decisional privacy is defined as "the right against unwanted interference with our decisions and actions" (Lanzing, 2019). Decisional privacy is harmed by persuasive profiling since it is specifically designed to - and have proved to be effective at – influencing consumer behavior. The effectiveness of personalised psychological profiling is so strong that it can be perceived as a form of manipulation and a loss of autonomy for the consumer (Susser *et al.*, 2019).

Gräf (2017) argues from a neo-republican stance that profiling causes people to be "subject to an extended capacity of arbitrary control" (p 446). He argues that the gaining knowledge of an individual, as in persuasive profiling, is a form of domination as it constitutes "interference of power in relation to the individuals' autonomy and thus their freedom to build identity and self" (Gutwirth & de Hert, 2008).

Digital market manipulation

The third concern is that persuasive profiling (and psychological marketing methods more generally) leads to the distortion of digital markets, hampering welfare. Market manipulation is a situation in which consumer choices do no longer necessarily reflect the true preferences of consumers (Calo, 2012). When the persuasion of marketeers is so strong that it trumps consumer preferences, this distorts markets and leads to suboptimal choices (Cofone & Robertson, 2017). Market manipulation will necessarily become widely used because if some firms use psychological profiling and hereby effectively extract market value other firms are forced to follow suit, rendering market manipulation through psychological profiling an endogenous feature of the market (Hanson & Kysar, 1999). Therefore, psychological profiling can cause economic markets to function inefficiently and so reduce total welfare. Market manipulation through (persuasive) profiling is not only an ethical problem, but also a legal problem because it creates new challenges to consumer protection law (Helveston, 2016).

Harm and adverse effects

Finally, persuasive profiling can have harmful and adverse effects (Mittelstadt *et al.*, 2016). Persuasive profiling could overly target vulnerable people (O'neil, 2012) or systematically look for ‘psychological weak spots’ (wired, 2011). Thereby, it enforces a mechanism in which people enjoying high social status people are offered quality products, while vulnerable people are overly persuaded to purchase inferior products. Ultimately, profiling limits options available to all users (Steindel, 2011) and may lead to “unfair discrimination and stigmatization” (Hildebrandt & Koops, 2010).

Additionally, “plausibly connected” targeted ads may “lead consumers to adjust their self-perceptions to match the implied label” (Summers *et al*., 2016). This form of ‘social labelling’ implies that behavioral targeting can influence user’s view of their own identity (HBR, 2016). Thus, persuasive profiling could harm identity forming and self-determination. Perhaps the most notorious application of persuasive profiling is its use by Cambridge Analytica, the company that collected psychographic data (illegitimately) from Facebook and used them to influence voters in the 2008 US Presidential election (Wired, 2012). The Cambridge Analytica scandal showed that persuasive profiling can also be used to influence democratic elections, casting doubt over the validity of democratic processes and the rationality of voter decisions (Ward, 2018). The full ethical and legal implications of persuasive profiling in democratic elections have yet to be investigated.

Benefits

There are two main benefits of using persuasive profiling. Firstly, persuasive profiling contributes to a more personalised experience for users, as it delivers the “appropriate persuasive ends, fitting the user context and activities”. For example, some people like to see ads that use humor, while others prefer to see what authoritative figures recommend. Persuasive profiling can tailor to these preferences.

Secondly, personalised profiling promises one huge advantage for marketeers: increased profits. The effectiveness of personalised persuasive messages is well-documented (IBM, 2019). Hirsch *et al*. (2012) found that personalised messages were rated more positively for all five characteristics. Matz *et al*. conducted three filed experiments in which they found that “persuasive appeals that were matched to people’s extraversion or openness-to-experience level resulted in up to 40% more clicks and up to 50% more purchases than their mismatching or generic counterparts” (Matz *et al*., 2017).

5.6 References

Aalberts, R. J., Nill, A., & Poon, P. S. (2016). Online Behavioral Targeting: What Does the Law Say? *Journal of Current Issues & Research in Advertising*, *37*(2), 95–112.

Bennett, S. C. (2010). Regulating online behavioral advertising. *J. Marshall L. Rev.*, *44*, 899.

Berger, D. D. (2010). Balancing consumer privacy with behavioral targeting. *Santa Clara Computer & High Tech. LJ*, *27*, 3.

Calo, R. (z.d.). Digital Market Manipulation. *THE GEORGE WASHINGTON LAW REVIEW*, *82*, 58.

Cofone, I. N., & Robertson, A. Z. (2017). Consumer Privacy in a Behavioral World. *Hastings LJ*, *69*, 1471.

Gräf, E. (2017). When Automated Profiling Threatens Our Freedom: *European Data Protection Law Review*, *3*(4), 441–451. https://doi.org/10.21552/edpl/2017/4/6

Gutwirth, S., & De Hert, P. (2008). Regulating profiling in a democratic constitutional state. In *Profiling the European citizen* (pp. 271–302). Springer.

Helveston, N. (z.d.). *REGULATING DIGITAL MARKETS*. *13*, 63.

Hildebrandt, M., & Koops, B.-J. (2010). The challenges of ambient law and legal protection in the profiling era. *The Modern Law Review*, *73*(3), 428–460.

Hirsh, J. B., Kang, S. K., & Bodenhausen, G. V. (2012). Personalised persuasion: Tailoring persuasive appeals to recipients' personality traits. *Psychological science*, *23*(6), 578–581.

Kaptein, M., & Eckles, D. (2010). Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling. In T. Ploug, P. Hasle, & H. Oinas-Kukkonen (Red.), *Persuasive Technology* (Vol. 6137, pp. 82–93). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-13226-1_10

Kaptein, M., Markopoulos, P., de Ruyter, B., & Aarts, E. (2015). Personalising persuasive technologies: Explicit and implicit personalisation using persuasion profiles. *International Journal of Human-Computer Studies*, *77*, 38–51. https://doi.org/10.1016/j.ijhcs.2015.01.004

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences*, *114*(48), 12714–12719. https://doi.org/10.1073/pnas.1710966114

Matz, Sandra C, Appel, R. E., & Kosinski, M. (2020). Privacy in the age of psychological targeting. *Current Opinion in Psychology*, *31*, 116–121. https://doi.org/10.1016/j.copsyc.2019.08.010

Matz, Sandra C, & Netzer, O. (2017). Using Big Data as a window into consumers' psychology. *Current Opinion in Behavioral Sciences*, *18*, 7–12. https://doi.org/10.1016/j.cobeha.2017.05.009

Mavriki, P., & Karyda, M. (2019). Automated data-driven profiling: Threats for group privacy. *Information & Computer Security*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/ICS-04-2019-0048

Steindel, T. A. (2010). A Path Toward User Control of Online Profiling. *Mich. Telecomm. & Tech. L. Rev.*, *17*, 459.

Summers, C. A., Smith, R. W., & Reczek, R. W. (2016). An audience of one: Behaviorally targeted ads as implied social labels. *Journal of Consumer Research*, *43*(1), 156–178. Scopus. https://doi.org/10.1093/jcr/ucw012

Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. *Internet Policy Review*, *8*(2). Scopus. https://doi.org/10.14763/2019.2.1410

Ward, K. (2018). Social networks, the 2016 US presidential election, and Kantian ethics: Applying the categorical imperative to Cambridge Analytica's behavioral microtargeting. *Journal of Media Ethics*, *33*(3), 133–148. https://doi.org/10.1080/23736992.2018.1477047

Zarsky, T. Z. (2019). Privacy and Manipulation in the Digital Age. *Theoretical Inquiries in Law*, *20*(1), 157–188.

Online sources

https://cloud.ibm.com/docs/services/personality-insights?topic=personality-insights-references#arnoux2017
https://hbr.org/2016/04/targeted-ads-dont-just-make-you-more-likely-to-buy-they-can-change-how-you-think-about-yourself
https://hbr.org/2018/05/what-marketers-should-know-about-personality-based-marketing
https://www.nature.com/articles/d41586-018-03880-4
https://www.wired.com/2011/04/st-essay-persuasion-profiling/
https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/

6 Case study 5: Website morphing

6.1 Objective

The objective of this case study is to understand how website morphing works and what its ethical and legal challenges are. Website morphing changes User Interface (UI) elements for individual users, to optimise some metric (such as click through rate, time on site, or money spent) over all users (Hauser *et al*., 2009). Similar techniques have been called 'adaptive user interfacing' (Liu *et al.*, 2003) or 'GUI adaption' (Paskalev, 2009). Here, I will use the term 'website morphing'. "Website morphing customizes the look and feel of a website to each customer so that, over a series of customers, revenue or profit are maximized" (Hauser *et al*., 2009). So, website morphing is defined as changing the interface of a website or application according to each user's characteristics with the goal of improving the user's experience or some business goal.

In 1997, Steven Johnson argued interfaces have strong social and cultural impact. He thought a new form of criticism was necessary to capture the power of interfaces. Since then, a broad field of design ethics and ethical Human Computer Interaction (HCI) has emerged (see f.e. Brey, 2010; Verbeek, 2006; vd Hoven *et al*., 2015; Friedman & Kahn, 2003). With the advent of the personalised interface, new ethical and legal challenges also arise. It is the goal of this case study to identify and discuss these issues.

6.2 Technology description

The interface of a website has a strong impact on how a user perceives that website and consequently acts on it (Chang & Chen, 2008; Li & Yeh, 2010).Therefore, website designers need to optimise their website interfaces if they want to give users the optimal experience and maximise click-through-rates, engagement or revenue. However, since users are heterogeneous, different website interfaces may suit different users. To accommodate different user preferences, engineers seek to personalise website designs by adapting interfaces to different user tastes.

Website morphing can be differentiated from other forms of personalisation, such as the widespread personalisation of products and services as found in recommendation systems, personalised advertisements and personalised data processing (Salonen & Karlajuoto, 2016). Website morphing seeks to adapt the interface by changing the colours, pagination, order of content, images and menu options of a website. Website morphing is thus about the way in which information is presented and not about which information, services or products are offered. It is thus an instance of 'means adaptive' personalisation (Kaptein & Eckles, 2008).

Recently attempts have been made to personalise website design, content and interface to increase user trust and engagement (Urban *et al*., 2009). User judge a website's credibility and value within the first few seconds of visiting (Fogg & Tseng, 1999). Website morphing can help to improve the attributes of a website that determine user trust and engagement. To achieve this, interfaces can be adapted to cultural background (Gevorgyan &

Manucharova, 2009; Reinecke & Bernstein, 2013) or to 'cognitive styles' (Hauser, 2008) which are "a person's preferred way of gathering, processing, and evaluating information" (Hayes and Allinson 1998, p.850).

Website morphing is based on the idea that different users have different preferences of interface design, and that a chosen metric can be improved by displaying each users his or her preferred interface. If some characteristics of the user are known, the morphed version of the website can be immediately delivered by linking the IP address, Google / Facebook login or cookies of a user to some known characteristics of the user. This is for example possible when the persuasion profile (see CS1) of the user is known. Alternatively, websites morphs can be offered based on the location, year of birth or political affiliation of the user (Reinecke & Bernstein, 2013).

If on the other hand no (sufficient) characteristics of the user are known in advance, no morph can be offered at first. In this case, morphing is still possible, but relevant user characteristics have to be inferred from user behavior (Hauser *et al*., 2009). This form of event-driven morphing can be performed in real-time. Two strategies to implement real-time morphing were found in the literature.

First, Frias-Martinez *et al.* (2006) use cognitive style analysis (CSA) by Riding (1991) and test a regression and a classification approach to infer one of two user cognitive styles from their interaction with a library catalogue website. In this experiment, search choices are used to predict user cognitive style. The authors found regression analysis performs best. Second, Hauser *et al*. (2008) recommend a Bayesian approach. They argue the cognitive style of the user (impulsive vs deliberative, visual vs verbal and analytic vs holistic) can be inferred from clickstreams. Based on just a few clicks, they propose adapting the website by adding or deleting features "such as column headings, links, tools, persona, and dialogue boxes" (p 4).

6.3 Context

6.3.1 Potential use

There are many potential uses for website morphing. Most examples in the literature refer to companies tailoring their marketing message or showcased products to individual customers (Urban *et al.* (2009). However, government website could also use website morphing to tailor specific messages to specific groups, or even to give different people the same information but in different types of language that are appealing to them.

6.4 Relationship to Workpackage and ESR PhD topic

Website morphing is defined as the personalisation of web interfaces, so it is close to the focus of this workpackage. It is also strongly connected to the topic of my PhD, digital nudging. Many nudges aim to structure choices and make choices easy by adapting choices architectures to user cognition (Johnson *et al*., 2012). Website morphing can be seen as an instance of nudging, as it also tries to subtly direct user behavior in a chosen direction by adapting the architecture of a website to the user. Moreover, nudging and morphing both focus on the 'how' of information presentation rather than the 'what' (Urban *et al*., 2009).

6.5 Review of ethical and legal issues

6.5.1 Methodology

For this case study a literature search is conducted in three academic databases, namely Google Scholar, Scopus and Hein Online. I first searched with the search string "web* morph*" OR "adaptive UI" OR "adaptive GUI" OR "adaptive UX" OR "adaptive user interface" OR "personal* interface" in HeinOnline and the same search string combined with the additional set ('ethic*' OR 'moral*' OR 'virtue*' in the other two databases. In all databases, the 100 most relevant results were analysed (first ten pages).

6.5.2 Findings

No results were found of articles that explicitly focus on ethical or legal issues of website morphing. Two articles fleetingly mention risks of morphing. Varian (2009) mentions the risk of misclassification and stresses the importance of giving users the possibility to reset their profile, especially if morph profiles are shared across websites. Urban *et al*. (2009) very briefly mention that morphing techniques should respect privacy laws. It can be concluded that the ethical and legal risks of morphing have not received attention in the (academic) literature. As website morphing is a vastly different form of personalisation than product recommendation, future research should explore which ethical and legal risks website morphing introduces. The literature on the ethical and legal risks of personalisation and interfaces could function as a starting point for this endeavour.

"HCI and STS scholars have sought to describe the ethical and value-laden relationship between designers and design outcomes" (Gray & Chivukula, 2019). However, personalisation is completely absent from this discussion. Therefore, I will now identify three ethical and legal issues with interfaces, personalisation and nudging and briefly discuss how the issues in these three fields can be extended to our present discussion of website morphing.

From interface to website morphing

It has been argued that interfaces have a strong impact on how people experience the world: "[t]he interface has already changed the way we use our computers, and will continue to do so in the years to come. But it is also bound to change other realms of modern experience in more unlikely, unpredictable ways" (Johnson, 1997, p 25). Steven Johnson notes how in the digital world, there is much more information available than can be consumed. He argues "[the] interface is a way of mapping that strange new territory, a way for us to get our bearing in a bewildering environment". However, interfaces favor certain experiences over others, such as the personal over the social and images over text. Interfaces "will continue to change the way in which we imagine information, and in doing so they are bound to change us as well – for the better *and* for the worse" (Johnson, 1997, p 242).

Johnson's view of interfaces suggests it is important to pay close attention to their influence over us. However, when interfaces become personalised, this becomes much harder (Pariser, 2009). So, one of the important ethical issues of website morphing future research should pay attention to is the way in which different people's experiences are shaped by different interface morphs and the influence this has on society.

From personalisation to website morphing

Website morphing, like all personalisation technologies, needs user data to select the optimal morph for each user. This leads to the ethical and legal risk of invading privacy (Toch *et al*., 2012). Privacy risks loom large in all three stages of website morphing: data collection, user model creation and adaptation (morphing) since users may not want (1) to share their personal data; (2) be profiled and marked as belonging to a certain group; or (3) be served a personalised morph, tailored to their characteristics (idem). Legal issues may arise in European law as consumer protection laws and GDPR dictate that user should have control over their own personal data (Zarksy, 2019). These ethical and legal issues could easily be applied to the practice of website morphing.

From nudging to website morphing

Website morphing ultimately is about seamlessly influencing the choices a website user makes. To understand the ethical issues this creates, we can turn to the literature on nudging, which refers to the idea of 'choice architecture' which is defined as "the context in which people make" decisions (Thaler & Sunstein, 2008, p 3). Nudges are subtle interventions in that choice architecture to change people's behavior in predictable ways. Interfaces are an excellent example of a choice architecture. The buttons, colors and text in each interface influence how people perceive a website. Website morphing can thus be seen as an intervention in the choice architecture that predictably alters user behavior. Website morphing is thus liable to many of the critiques charged at nudging. It steers user behavior in an unnoticed and non-rational way, which can be said to be less than autonomous (Yeung, 2017). Some authors argue that changing people's behavior for selfish reasons is a form of manipulation (Blumenthal-Barby & Burroughs, 2012). And other argue that non-transparent influence is a form of manipulation (Hansen & Jespersen, 2012). These ethical and legal issues can serve as a starting point for future research into the ethics of website morphing.

Benefits

The issues discussed in the previous section show website morphing is promising for a variety of contexts. While most researchers focus on content personalisation, interface personalisation can offer enormous advantages to both users and companies. Urban *et al.* (2009) argue website morphing builds "empathy, trust and sales" as it makes users more comfortable on a website. In an exploratory study, Kwon & Kim (2012) found that interface personalisation was even more effective than content personalisation in increasing customer loyalty and trust on a news portal website.

Morphing will be especially interesting to e-commerce websites, which can use website morphing to increase their revenues. Hauser *et al*., (2009) found that they could increase revenue of British Telecom by 19,9% by offering users a morph after collecting 10 clicks. This would amount to an additional $80 million in sales. In a field study with Suruga bank, it was found that website morphing increased the amount of purchases by four percentage points (Hauser *et al*., 2010). Such figures will surely be of interest to commercial parties.

6.5.3 References

Chang, H. H., & Chen, S. W. (2008). The impact of customer interface quality, satisfaction and switching costs on e-loyalty: Internet experience as a moderator. *Computers in Human Behavior*, *24*(6), 2927–2944. https://doi.org/10.1016/j.chb.2008.04.014

Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems the CHI Is the Limit - CHI '99*, 80–87. https://doi.org/10.1145/302979.303001

Frias-Martinez, E., Chen, S. Y., & Liu, X. (2007). Automatic cognitive style identification of digital library users for personalisation. *Journal of the American Society for Information Science and Technology*, *58*(2), 237–251. https://doi.org/10.1002/asi.20477

Gevorgyan, G., & Manucharova, N. (2009). Does Culturally Adapted Online Communication Work? A Study of American and Chinese Internet Users' Attitudes and Preferences toward Culturally Customized Web Design Elements. *Journal of Computer-Mediated Communication*, *14*(2), 393–413. https://doi.org/10.1111/j.1083-6101.2009.01446.x

Hauser, J. R., Liberali, G., & Urban, G. L. (2014). Website morphing 2.0: Switching costs, partial exposure, random exit, and when to morph. *Management science*, *60*(6), 1594–1616.

Hauser, J. R., Urban, G. L., & Liberali, G. (2010). *When to Morph*. Cambridge, MA: MIT Sloan School of Management.

Hauser, J. R., Urban, G. L., & Liberali, G. (2011). *Website Morphing 2.0: Technical and Implementation Advances Com-bined with the First Field Experiment of Website Morphing*.

Hauser, J. R., Urban, G. L., Liberali, G., & Braun, M. (2009a). Rejoinder—Response to Comments on "Website Morphing". *Marketing Science*, *28*(2), 227–228.

Hauser, J. R., Urban, G. L., Liberali, G., & Braun, M. (2009b). Website Morphing. *Marketing Science*, *28*(2), 202–223. https://doi.org/10.1287/mksc.1080.0459

Johnson, S. (1997). *Interface culture: How new technology transforms the way we create and communicate*. Basic Books, Inc.

Johnson, E. J., Shu, S. B., Dellaert, B. G. C., Fox, C., Goldstein, D. G., Häubl, G., Larrick, R. P., Payne, J. W., Peters, E., Schkade, D., Wansink, B., & Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters*, *23*(2), 487–504. https://doi.org/10.1007/s11002-012-9186-1

Kwon, K., & Kim, C. (2012). How to design personalisation in a context of customer retention: Who personalises what and to what extent? *Electronic Commerce Research and Applications*, *11*(2), 101–116.

Li, Y.-M., & Yeh, Y.-S. (2010). Increasing trust in mobile commerce through design aesthetics. *Computers in Human Behavior*, *26*(4), 673–684. https://doi.org/10.1016/j.chb.2010.01.004

Liu, J., Wong, C. K., & Hui, K. K. (2003). An adaptive user interface based on personalised learning. *IEEE Intelligent Systems*, *18*(2), 52–57.

Paskalev, P. (2009). Rule based GUI modification and adaptation. *Proceedings of the International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing*, 1–7.

Reinecke, K., & Bernstein, A. (2008). Predicting user interface preferences of culturally ambiguous users. In *CHI'08 extended abstracts on Human factors in computing systems* (pp. 3261–3266).

Reinecke, K., & Bernstein, A. (2011). Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(2), 1–29.

Reinecke, K., & Bernstein, A. (2013). Knowing what a user likes: A design science approach to interfaces that automatically adapt to culture. *Mis Quarterly*, 427–453.

Salonen, V., & Karjaluoto, H. (2016). Web personalisation: The state of the art and future avenues for research and practice. *Telematics and Informatics*, *33*(4), 1088–1104.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Toch, E., Wang, Y., & Cranor, L. F. (2012). Personalisation and privacy: A survey of privacy risks and remedies in personalisation-based systems. *User Modeling and User-Adapted Interaction*, *22*(1–2), 203–220.

Urban, G. L., Hauser, J. R., Liberali, G., Braun, M., & Sultan, F. (2009). Morph the web to build empathy, trust and sales. *MIT Sloan Management Review*, *50*(4), 53.

Varian, H. (2009). Commentary—Discussion of "Website Morphing". *Marketing Science*, *28*(2), 224–224.

Violante, J. G. (2011). *Behavior-based personalisation: Strategies and Implications* [Thesis, Massachusetts Institute of Technology]. https://dspace.mit.edu/handle/1721.1/65820

Yeung, K. (2017). 'Hypernudge': Big Data as a mode of regulation by design. *Information, Communication & Society*, *20*(1), 118–136. https://doi.org/10.1080/1369118X.2016.1186713

Zarsky, T. Z. (2019). Privacy and Manipulation in the Digital Age. *Theoretical Inquiries in Law*, *20*(1), 157–188. https://doi.org/10.1515/til-2019-0006

7 Case study 6: Personalised e-learning

7.1 Objective

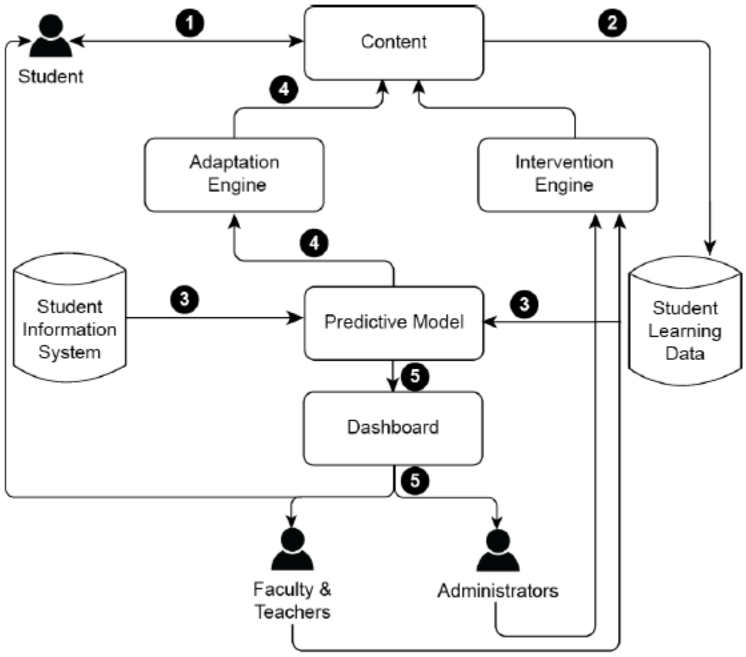
The objective of this case study is to discuss the ethical and legal issues with personalised e-learning technologies. Personalised e-learning systems are diverse so when describing their technology, we must be careful to include most applications without offering a generic description. We can focus on what most applications have in common, namely that “[i]n general, personalised-learning models seek to adapt the pace of learning and the instructional strategies being used to best fit each individual child's strengths, weaknesses, and interests" (Herold, 2017a; edweek). Typically, personalisation uses personal or behavioural data of users to modify any aspect of the learning system to optimize the learning experience (Bulger, 2016).

7.2 1.6.2 Technology description

Bulger (2016) distinguishes between responsive and adaptive systems. Responsive systems are more fixed as these systems can be customized or tweaked to change interface and menu’s (see CS2) or offer students the opportunity to choose their own learning path through predetermined material. Adaptive systems on the other hand are much more flexible. They “move beyond a predetermined decision tree and use machine learning to adapt to a student’s behaviour and competency” (Bulger, 2016, p 6). It is important to stress that behavioural data can be used to adapt for difficulty (Bulger, 2016), content (Bulger, 2016), learning style (Regan & Jesse, 2019) or learning orientation (Martinez, 1999; 2001).

A “prototypical” adaptive learning system works as follows (Bienkowski *et al.*, 2012): a student (1) interacts with an interface and course data, this (2) generates student data, these data are (3) combined with demographic data and serve as input for a predictive model that predicts the optimal score for a number of variables, these variables are (4) entered into the adaption engine which determines learning components are delivered to the student. The predictive model (5) also provides feedback about the learners to a dashboard for supervisors which allows them to keep track of learners and personalised adaptions. They can (6) manually intervene in the process by overturning the automated decisions of the adaption engine through an intervention engine. This process is depicted in figure below[30].

[30] Retrieved from Bienkowski *et al.*, 2012, p 18. Note that the sixth process has no number in the figure, but the process is displayed, nonetheless.



7.3 Context

Personalised e-learning is used in both traditional educational institutes and by new emerging actors, often tech firms and start-ups.

7.3.1 Potential use

Examples of traditional institutes that use personalised e-learning are found at all levels: K-12 US Primary schools (Dishon, 2017), secondary schools (Dishon, 2017) and institutes that offer higher education (Avella *et al*., 2016; Kay *et al*., 2012). According to Edweek over 97% of all US high schools were investing in personalised learning (Herold, 2017a; edweek). However, most known for innovations in e-learning are (commercial) learning platforms that offer Massive Open Online Courses (MOOC) such as Coursera, Udemy, Edx and Udacity (Rizzardini & Amado-Salvatierra, 2018, p 16). Large tech players Facebook and Microsoft are also showing interest as the Bill & Melinda Gates Foundation is investing in personalised learning (Regan & Jesse, 2019) and so is Mark Zuckerberg's charity (Herold, 2017b; edweek).

7.4 Relationship to Workpackage

Personalised e-learning is a prime example of personalisation by digital technologies and hence very suited as a case study in this workpackage.

7.5 Relationship to ESR PhD topic

The relationship between personalised e-learning and digital nudging – the subject of my PhD thesis – is less obvious but nevertheless important to address. Some authors investigate the use of nudges to improve education (Dimitrova *et al*., 2017; Palmer, 2020). However, this is highly controversial as we will see in the discussion on the ethical and legal issues beneath.

7.6 Review of ethical and legal issues

7.6.1 Methodology

For this case study a literature search is conducted in three academic databases, namely Google Scholar, Scopus and Hein Online. I first searched in HeinOnline using the following search string: "*e-learning*" OR "*edtech*" OR ("*personal**" AND "*learning*" AND "*analytics*"). In the other two databases, the same search string was entered combined with the additional string: "*ethic**" OR "*moral**". In all databases, the 100 most relevant results were analysed (first ten pages). As a next step, relevant articles from the bibliographies of the selected articles were selected. Finally, some articles were added ad hoc.

7.6.2 Findings

In this section we will discuss the ethical and legal issues with personalised e-learning found in the literature.

Learning analytics

While interesting, I will not discuss the ethical and legal issues about **privacy, surveillance and security** that personalised e-learning bring about. For a discussion of these topics see Lynch (2017), Lupon & Williamson (2017) and Slade & Prinsloo (2013) among others.

Personalisation

As education is personalised, various authors warn that care should be taken to **retain serendipity** (Ashman *et al*., 2014). Education is about confronting one's own views with that of others and learning to adjust one's worldview. Others emphasize that unexpected new information can be challenging (Slade Prinsloo, 2013; Hartman-Caverly, 2019). Another important ethical issue that arises due to personalisation is **managing trust** (Ashman *et al*., 2014). Learner engagement is eafected by their trust in the system (Siemens, 2013). Therefore, PLE's must be transparent, give learners the opportunity to participate in design choices and have good data practices (Kandratova *et al*., 2018). The issues of **transparency and control** are further emphasized as a separate ethical issue. "When data-intensive technology progresses to become increasingly complex and opaque, it is increasingly difficult to lay bare the values that are implicit in them" (Huis & Nagenborg, 2019, p 55). It is important to maintain "scrutability and user control of personalisation" to foster both trust and autonomy (Ashman *et al*., 2014; Slade & Prinsloo, 2013). However, a difficult legal issue arises as education is often aimed at minors who may not be able to give informed consent (Regan & Jesse, 2019). A final issue with personalisation is possible **discrimination.** Algorithms and data sets may contain biases (Alarcon *et al*., 2014). This may give rise to discrimination and unfair treatment of children of lower social and economic status (Drachsler & Greller, 2016; Regan & Jesse, 2019). Discrimination is an ethical issue but also has a legal dimension as discrimination violates human rights (EU charter of fundamental rights, article 21).

Measures

"Personalisation systems make assumptions and inferences about the user" (Ashman *et al*., 2014, p 13). However, these measures are not perfect, leading to a **lack of accuracy of inferring**. This constitutes the worry that algorithms could infer student characteristics wrongly and cause a "danger of false positives" (Bulger, 2016). The concern is that "students will be unnecessarily flagged for a cycle of remediation based on a faulty algorithm and a lack of understanding of how people learn" (Bulger, 2016, p18). Imagine a bright student who is erroneously classified as 'slow' or 'below average' and therefore is placed in a slower or easier learning track. She may lose interest in the program which is too easy for her and her performance could deteriorate even further, leading to a self-fulfilling prophecy.

But even if user models were perfect, an even more fundamental danger looms large because personalisation systems **cannot model success clearly** (Kondratova *et al*., 2018). To understand this, one must know that an adaptive system can function either based on human rules (categorize students and content and subsequently serve each type of student the appropriate type of content), or by optimising a metric (classify and serve in such a way that a variable x is maximised). For systems of the second kind, Bulger asks the paramount question: "what do personalised learning systems optimize for?" as "it is difficult to measure for or optimize for success when success isn't clearly defined" (Bulger, 2016, p 18-19). Not everything that is valuable about education can be captured in data since "[s]tudent success is a complex and multidimensional phenomenon" (Slade & Prinsloo, 2013, p 13). Therefore, it can be questioned if any (set of) metric(s) is suited to guide educational choices.

Whichever metric is chosen, **too much focus on metrics** can be a vice because optimizing metrics could become a goal in itself (O'neil, 2012). An adverse consequence of overly emphasizing metrics is the risk of viewing students as a passive mass that must be manipulated to optimise metrics (Slade & Prinsloo, 2013). This entails the worry about students being pushed by university managers who may see personalisation as a tool to quickly and efficiently output degrees (Ashman *et al*., 2014).

One metric often measured is **student engagement**. Student engagement has two important characteristics in personalised e-learning. First, it "is presumed to be malleable, responsive to contextual features, and amenable to environmental change" (Fredericks et al., 2004). Second, it is often measured as number of resources accessed and amount of time spent (Bulger, 2016). This idea of malleability combined with the ability to perfectly track time spent in digital environments made e-learning applications "a site for behavioural intervention and nudging" (Knox *et al*., 2020; Bradbury *et al*., 2013). Examples of nudges and attention engineering are devising personalised nudges to increase video watching (Dimitrova *et al*., 2017) and sending out personalised emails to increase course participation (Palmar, 2020).

Measuring and optimizing student engagement is problematic because time spent is a poor proxy for engagement (Berliner, 1990). In addition to being an imperfect metric, engagement is strongly related to **issues regarding social (attention) engineering and**

nudging (Jones, 2017). Nudging may be problematic because it hinders reflection which is vital for learning (Regan & Jesse, 2019). Regan and Jesse argue that nudges in personalised systems may “entail more direction than suggestion” (p 13). In addition, attention engineering in education is critiqued as antithetical to students’ intellectual freedom and development as self-sufficient learners and independent thinkers” (Hartman-Caverly, 2019, p 24).

This is agreed upon by Knox *et al*. who fear behavioural science causes learners to be perceived as irrational subjects who need to be steered to preferable outcomes. Here, personalised learning is critiqued for its connection to **behaviourism** (Kohn, 2015). His critique is that personalised e-learning breaks down education in separate learnable chunks, which are then learned through behaviourist conditioning at the expense of constructivist or cognitivist approaches to education (da Silva *et al*., 2012). Behaviourism can be especially harmful in the educational context as it harms autonomy and empowerment because it steers all learners to predefined goals, instead of allowing them to discover their own values (Knox *et al*., 2020).

Learning theories

While personalised e-learning is popular as we have seen, it is criticized for a **lack of evidence** for its effectiveness and inherent difficulty to come by this evidence (Ashman *et al*., 2014; Bulger, 2016). It has been argued the “hype outweighs the research” (Herold, 2017a; edweek). In addition, it is argued that personalisation **neglects social dimension of teaching** (Bulger, 2016; Kohn 2015) and leads to **skill degradation** (Ashman *et al*., 2014; Kohn).

Moreover, the argument is made that personalisation **lacks underlying pedagogies** (Bartolomé *et al*., 2018) and uses no coherent learning strategy or overall philosophy. Building new educational tools “requires an engagement with pedagogy, politics and ideology which has so far been more conspicuous by its absence than its salience” (Griffiths, 2020, p 52). This critique constitutes the idea that personalised e-learning systems require student to gain knowledge but cannot address more fundamental aspects of education such as critical reflection, citizenship and social skills.

Benefits

Personalisation is widely used and quickly adopted because it promises great improvements in education (Bulger, 2016). According to Ashman *et al*. (2014) personalised e-learning promises three main benefits. The first is ‘engagement’, it makes students more motivated (Ashman *et* al,. 2014; Samah *et al*., 2011). The second is ‘economy’, it makes learning more cost effective. The third is ‘outcome’, it leads to better learner achievements (Ashman *et* al,. 2014; Pogorskiy, 2015). Another hopeful outcome is that personalised e-learning has a certain egalitarian and democratic appeal because it makes knowledge available for previously excluded groups (Yu, 2002).

7.7 References

Amo, D., Fonseca, D., Alier, M., García-Peñalvo, F. J., Casañ, M. J., & Alsina, M. (2019). Personal Data Broker: A Solution to Assure Data Privacy in EdTech. *International Conference on Human-Computer Interaction*, 3–14.

Anwar, M. M., Greer, J., & Brooks, C. A. (2006). Privacy enhanced personalisation in e-learning. *Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services*, 1–4.

Ashman, H., Brailsford, T., Cristea, A. I., Sheng, Q. Z., Stewart, C., Toms, E. G., & Wade, V. (2014). The ethical and social implications of personalisation technologies for e-learning. *Information & Management*, *51*(6), 819–832.

Attwell, G. (2007). Personal Learning Environments-the future of eLearning. *Elearning papers*, *2*(1), 1–8.

Avella, J. T., Kebritchi, M., Nunn, S. G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Journal of Asynchronous Learning Network*, *20*(2). Scopus.

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *JEDM| Journal of Educational Data Mining*, *1*(1), 3–17.

Bartolomé, A., Castañeda, L., & Adell, J. (2018). Personalisation in educational technology: The absence of underlying pedagogies. *International Journal of Educational Technology in Higher Education*, *15*(1), 14.

Bellini, C., De Santis, A., Sannicandro, K., & Minerva, T. (2019). Data management in learning analytics: Terms and perspectives. *Journal of E-Learning and Knowledge Society*, *15*(3), 133–144. Scopus. https://doi.org/10.20368/1971-8829/1135021

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *Proceedings of conference on advanced technology for education*, 1–64.

Bulger, M. (2016). Personalised learning: The conversations we're not having. *Data and Society*, *22*.

da Silva, N. S. A., da Costa, G. J. M., Prior, M., & Rogerson, S. (2012). The evolution of e-learning management systems: An ethical approach. In *Virtual Learning Environments: Concepts, Methodologies, Tools and Applications* (pp. 67–79). IGI Global.

Dimitrova, V., Lau, L., Piotrkowicz, A., Weerasinghe, A., & Mitrovic, A. (2017). Using learning analytics to devise interactive personalised nudges for active video watching. *UMAP 2017 - Proceedings of the 25th Conference on User Modeling, Adaptation and Personalisation*, 22–31. Scopus. https://doi.org/10.1145/3079628.3079683

Dishon, G. (2017). New data, old tensions: Big data, personalised learning, and the challenges of progressive education. *Theory and Research in Education*, *15*(3), 272–289. https://doi.org/10.1177/1477878517735233

Drachsler, H., & Greller, W. (2016). *Privacy and analytics—It's a DELICATE issue a checklist for trusted learning analytics*. *25-29-April-2016*, 89–98. Scopus. https://doi.org/10.1145/2883851.2883893

Fiedler, S. H., & Väljataga, T. (2011). Personal learning environments: Concept or technology? *International Journal of Virtual and Personal Learning Environments (IJVPLE)*, *2*(4), 1–11.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59–109.

Griffiths, D. (2020). The Ethical Issues of Learning Analytics in Their Historical Context. *Lecture Notes in Educational Technology*, 39–55. Scopus. https://doi.org/10.1007/978-981-15-4276-3_3

Hall, R., & Stahl, B. (2015). Against commodification: The university, cognitive capitalism and emergent technologies. In *Marx and the Political Economy of the Media* (pp. 65–97). Brill.

Hartman-Caverly, S. (2019). Human Nature Is Not a Machine: On Liberty, Attention Engineering, and Learning Analytics. *Library Trends*, *68*(1), 24–53. https://doi.org/10.1353/lib.2019.0029

Huis, I., & Nagenborg, M. (2019). It's getting personal: The ethical and educational implications of personalised learning technology. *Journal of Philosophy in Schools*, *6*(1).

Jones, K. M. L. (2017). Learning Analytics and Its Paternalistic Influences. *Learning and Collaboration Technologies: Technology in Education, Lct 2017, Pt Ii*, *10296*, 281–292. https://doi.org/10.1007/978-3-319-58515-4_22

Kay, D., Korn, N., & Oppenheim, C. (2012). Legal, risk and ethical aspects of analytics in higher education. *Analytics series*.

Knox, J., Williamson, B., & Bayne, S. (2020). Machine behaviourism: Future visions of 'learnification' and 'datafication' across humans and digital technologies. *Learning, Media and Technology*, *45*(1), 31–45. Scopus. https://doi.org/10.1080/17439884.2019.1623251

Kondratova et al_2018_Supporting Trust and Engagement in Personalised Learning.pdf. (z.d.).

Kondratova, I., Molyneaux, H., & Fournier, H. (2018). Supporting Trust and Engagement in Personalised Learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10925 LNCS*, 44–59. Scopus. https://doi.org/10.1007/978-3-319-91152-6_4

Lupton, D., & Williamson, B. (2017). The datafied child: The dataveillance of children and implications for their rights. *New Media & Society*, *19*(5), 780–794.

Lynch, C. F. (2017). Who prophets from big data in education? New insights and new challenges. *Theory and Research in Education*, *15*(3), 249–271. Scopus. https://doi.org/10.1177/1477878517738448

Martinez, M. (2001). Key design considerations for personalised learning on the web. *Journal of Educational Technology & Society*, *4*(1), 26–40.

Maseleno, A., Sabani, N., Huda, M., Ahmad, R., Jasmi, K. A., & Basiron, B. (2018). Demystifying learning analytics in personalised learning. *International Journal of Engineering & Technology*, *7*(3), 1124–1129.

Omotoyinbo. (2016). EDUCATIONAL TECHNOLOGY AND VALUE NEUTRALITY. *Socialinių Mokslų Studijos*, *8*(2), 163–179.

O'neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.

Palmer, M. A. (2020). Improving student success by being automatically personal. *Computers in Education Journal*, *11*(1). Scopus.

Pogorskiy. (2015). *Using personalisation to improve the effectiveness of global educational projects*. https://journals-sagepub-com.dcu.idm.oclc.org/doi/10.1177/2042753014558378

Polonetsky, J., & Tene, O. (2014). Who Is Reading Whom Now: Privacy in Education from Books to MOOCs. *Vanderbilt Journal of Entertainment and Technology Law*, *17*(4), 927–990.

Prinsloo, P., & Slade, S. (2017). *An elephant in the learning analytics room—The obligation to act*. 46–55. Scopus. https://doi.org/10.1145/3027385.3027406

Pykett, J. (2009). Personalisation and De-Schooling: Uncommon Trajectories in Contemporary Education Policy. *Critical Social Policy*, *29*(3), 374–397.

Regan, P. M., & Jesse, J. (2019). Ethical challenges of edtech, big data and personalised learning: Twenty-first century student sorting and tracking. *Ethics and Information Technology*, *21*(3), 167–179.

Rizzardini, R. H., & Amado-Salvatierra, H. R. (2018). Exploring new ways to increase engagement in full-path MOOC programs. *International Conference on Learning and Collaboration Technologies*, 16–25.

Samah, N. A., Yahaya, N., & Ali, M. B. (2011). Individual differences in online personalised learning environment. *Educational Research and Reviews*, *6*(7), 516–521.

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*. https://doi.org/10.1177/0002764213498851

Siemens, G., & Baker, R. S. d. (2012). Learning analytics and educational data mining: Towards communication and collaboration. *Proceedings of the 2nd international conference on learning analytics and knowledge*, 252–254.

Slade, S., Prinsloo, P., & Khalil, M. (2019). *Learning analytics at the intersections of student trust, disclosure and benefit*. 235–244. Scopus. https://doi.org/10.1145/3303772.3303796

Slade, Sharon, & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510–1529.

Wang, Y. (2016). Big Opportunities and Big Concerns of Big Data in Education. *TechTrends*, *60*(4), 381–384. Scopus. https://doi.org/10.1007/s11528-016-0072-1

Yu, P. K. (2002). Bridging the Digital Divide: Equality in the Information Age Bridging the Digital Divide: Equality in the Information Age: Introduction. *Cardozo Arts & Entertainment Law Journal*, *20*(1), 1–52.

Online sources

Herold, 2017a: https://www.edweek.org/ew/articles/2017/11/08/the-cases-against-personalised-learning.html

Herold, 2017b: https://www.edweek.org/ew/articles/2017/06/29/chan-zuckerberg-to-push-ambitious-new-vision-for.html

Kohn, 2015: https://www.alfiekohn.org/blogs/personalised/

https://ethicaledtech.info/wiki/Meta:Welcome_to_Ethical_EdTech

8 Case study 7: Personal Voice Assistants

8.1 Objective

The objective of this case study is to describe what personal voice assistants are, in what sense they are personalised technologies, and their ethical and legal implications. The focus lies on ethical issues to do with the personalised aspects of voice assistants, but attention will also be given to legal issues and more general ethical issues regarding personal voice assistants.

8.2 Technology description

Personal voice assistants – also referred to as digital assistants, AI virtual assistants, intelligent assistants, and so on – are software systems that can provide services to an individual user on the basis of a conversational user interface. These software systems are most commonly embedded in one's personal devices (such as a smartphone, personal computer, smart watch or smart speaker) and can be connected to other devices (such as lights, televisions or even shoes) (Barkho, 2019; Chung et al., 2017). The conversational user interface that characterizes personal voice assistants entails that the assistant can recognize and understand users' spoken commands and in turn responds to users in spoken word. This is possible due to natural language processing (NLP), a subfield of artificial intelligence (AI).

The most known examples of personal voice assistants are Amazon Alexa, Google Assistant, Apple's Siri, and Microsoft Cortana, but other companies have or are developing similar systems too (Chung et al., 2017). The services a personal voice assistant can provide depend on the nature of the company that developed the particular system as well as the devices the software system can be connected to. For example, Amazon's Alexa can interact with users, play music, make to do lists, set alarms, stream podcasts, play audiobooks, provide information (weather, traffic, news, etc.), order products on Amazon, or order food from delivery services that collaborate with Amazon.

Voice assistants are *personal* in that are specifically designed to resemble a human personal assistant and to feel personal to a user, in the sense that the assistant is always there for *you*, always ready to help *you*. The design of personal voice assistance most commonly exists in a female voice and name, sophisticated use of language, and an extrovert, caring and servile personality.

8.3 Context

Today, personal voice assistants are mainly used in the context of one's private life and private home. In other words, the services that personal voice assistants provide mainly assist individuals in private aspects of their lives. Personal voice assistants could also be used in a professional context (Finnegan, 2020; Prist, 2018) and for customer service online, via phone call, or even in a physical store, restaurant, or other type of facility. In this case study, however, personal voice assistants are studied only in the private context described above. Both present services of personal voice assistants and potential future uses are taken into consideration.

8.3.1 Potential use

There are some functions that are not yet embedded in existing personal voice assistant, but are currently in the making or likely to be developed in the future. Currently in the making are for example personal voice assistants that can recognize and respond to human emotions (Day, 2019; Johnson, 2019) or adapt their services to different users of a single device (Amazon Alexa, n.d.). A potential future service of personal voice assistants could be social companionship, meaning that one would not just turn to his/her personal assistant for practical tasks but also for social interaction or emotional connection. Another possible development is that personal voice assistants become connected to wearables, making it possible to get to know a user and his/her routines, health conditions, and emotions even better and adjust services accordingly.

8.4 Relationship to Work Package

This case study is part of a work package on the ethics of personalisation. Voice assistants are a *personalised* technology because they can offer personalised services (e.g. recommending certain songs, ordering products, filtering news), by analysing and remembering user's preferences (e.g. music taste, previous purchases, interest in certain news items) and also by recognizing a user's voice (e.g. in cases where multiple residents use a single smart speaker).

8.5 Relationship to PhD topic (ESR13)

For my PhD project I investigate how video analytics applications can impact individuals' autonomy. Personal voice assistants are not an application of video analytics, hence I cannot directly use this case study in my dissertation. But the case is nevertheless interesting in light of my research, because voice assistants can affect human autonomy in ways similar to many video analytics applications.

8.6 Review of ethical and legal issues

8.6.1 Methodology

This case study is based on academic literature as well as media and grey literature. The academic literature was found by means of a critical search in Google Scholar, Scopus and HeinOnline as well as by a backwards snowballing technique. The media and grey literature were found using Google Search. For both searches a mixture of search terms was used, namely: 'personal* voice assistant', 'digital assistant', or 'AI assistant' in combination with 'ethic*', 'moral*', or 'legal*'. The selected search results could be from any discipline, as long as they are less than 10 years old (i.e. published after 2010) and discussed ethical or legal issues/concerns regarding current or future uses of personal voice assistants (as I described them above). This literature search was not strictly systematic, but should nevertheless provide a relevant overview of the ethical and legal issues regarding personal voice assistants. Below I discuss the found ethical and legal issues, after which I conclude with a few words on possible beneficial aspects of the technology.

8.6.2 Findings

The selected literature covers a wide variety of ethical and legal issues regarding personal voice assistants. A first, and obvious, set of concerns has to do with **privacy and data-governance**, which are both ethical and legal issues. Voice assistants threaten privacy because at times they record audio data without the user actively using the system or being aware of the fact that the system is recording (Chung et al., 2017; Green, 2018; Temkin & Dorsett, 2019). In other words, voice assistants could eavesdrop on users' private conversations without them realizing it. Boughman et al. rightfully ask "should one expect privacy in the communications he engages in around a voice-controlled digital assistant?" (Boughman et al., 2017, p. 1). To make things worse, voice assistants can be present in one's entire house and in some cases even carried around on smart phones and people are likely to act more careless around these discrete, almost invisible devices. Moreover, privacy is at risk because of the sensitive information that speech could reveal (Cox, 2019). Voice data reveals information that mere text could not (e.g. one's accent can reveal where one is originally from and the tone of one's voice can unveil one's emotional state). In that sense, voice data might be "more private" than other types of personal data (Boughman et al., 2017). A further issue is the fact that the data and information received by voice assistants can be used by third parties, for example to send personalised advertisements. Green (2018) points out that voice assistants make it easier for third parties to obtain information about individuals and Davis (2016) moots that personal voice assistants might be giving companies too much access to people's lives.

Despite the various worries raised by specialists, ethicists, legal experts and journalists, studies have shown that, with exception of the always-on mode, voice assistants are not raising much privacy concerns among users (Fruchter et al., 2018; Manikonda et al., 2018). Similarly, Backe (2018) and Brandom (2019) note that users *appear* to be okay with trading off their privacy for convenience. Schwartz (2019), on the other hand, says that the majority of smart speaker users "don't want their voice assistant's personalisation ability to improve" because they do not want to share more personal data.

Trust is another issue that is often raised in the literature. It is important that users can trust their voice assistants because of the personal role they fulfil and place they take in one's private life. Therefore, assistants should be plenty robust against hacking or misuse, which is tricky because of the 'always-on mode . Users should also be able to trust that companies use personal data in ways that does not violate users' privacy.

Furthermore, personal voice assistants can threaten the **autonomy** of users. First of all, the use of voice assistants can lead to de-skilling. Outsourcing certain tasks to AI assistants might cause cognitive degeneration – e.g. not knowing how to manually turn on other devices, or how to search for and filter information. According to Danaher (2018), this de-skilling is problematic only when the outsourced task has intrinsic value. The use of personal voice assistants might also lead to social de-skilling, either by outsourcing human-human communication to personal assistants (Danaher, 2018) or replacing human interaction with personalised interaction with one's voice assistant (Mattin, 2019). Personal voice assistants can also threaten the autonomy of users by manipulating choices and choice architectures, for example when they make music recommendations or order products (Danaher, 2018; Davis, 2016).

Another issue has to do with the common user interface of voice assistant systems, which is **discriminating women and minorities** (Adams & Loideáin, 2019; Bogost, 2018, Kleber, 2018; Prescot, n.d.). The combination of a female voice and name with a subordinate role and personality reinforces sexist stereotypes, such as the idea that women are always at your service and it is okay to boss them around (Kleber, 2018). Furthermore, the accent and language use of voice assistants are perceived to be 'white' and well-educated – thus lacking diversity (Prescot, n.d.). The discriminatory design of voice assistants is not just an ethical but also a legal concern, because such discrimination can be perceived as a human rights violation (Adams & Loideáin, 2019).

Many of the mentioned ethical concerns become more pressing when thinking about the 'generation voice', i.e. the generation that will grow up being used to personal voice assistants (Turkington, 2019). This generation might have a different understanding of privacy, never have to learn certain skills, and adopt stereotypes displayed by personal voice assistants. Growing up with a particular 'brand' of voice assistant (say, Alexa), might also cause them to be life-long customers of the respective company (Amazon). A personal voice assistant that is personalised on the basis of years of data will make it unattractive to switch from one to another. Voice assistants create a new way of consuming and "there exists a risk that the voice assistant becomes the gatekeeper by default that may lead to an abuse of market power" (Rabassa, 2019). This **market power** then forms another legal concern regarding personal voice assistants.

Personal voice assistants raise new ethical and legal questions about the rights that should be ascribed to the systems itself. Boughman et al. (2017) ask whether Amazon's Alexa has, could or should have any rights. Similar questions play in the field of robotics and these become more complex as AI systems improves, i.e., when voice assistants become more intelligent.

Finally, the technology does not only raise ethical or legal problems but also comes with a range of **benefits**, which can be deemed 'ethical aspects' of voice assistants. Mentioned earlier are the possible negative consequences that personal voice assistants can have on individuals' autonomy. However, voice assistants can positively affect autonomy too. The convenience that these devices offer can save individuals time, allow them to perform multiple activities at the same time or even to do things they otherwise could not have. The systems also have the potential of detecting cases of domestic violence and audio data can be used as evidence in court. Voice assistants could even act as moral exemplars or teachers, although they are currently not designed to.

8.7 References

Amazon Alexa (n.d.). Add Skill to your Personalisation. https://developer.amazon.com/en-US/docs/alexa/custom-skills/add-personalisation-to-your-skill.html (Accessed May 14 2020).

Adams, R. & Loideáin, N.N. (2019). Addressing indirect discrimination and gender stereotypes in AI virtual personal assistants: The role of international human rights law. *Cambridge International Law Journal, 8*(2), 241-257.

Backe, N. (2018, November 26). Is Amazon Alexa invading privacy? *Medium*. https://medium.com/@nils.backe/is-amazon-alexa-invading-privacy-analysis-of-an-ethical-dilemma-f7e064ab6dba

Barkho, G. (2019, March 9). Nike's Futuristic New Sneakers allow Siri to Control your Shoe Laces. *Observer*. https://observer.com/2019/09/nike-siri-enabled-sneakers-control-shoe-laces/

Bogost, I. (2018, January 24). Sorry, Alexa is not a feminist. *The Atlantic*. https://www.theatlantic.com/technology/archive/2018/01/sorry-alexa-is-not-a-feminist/551291/

Boughman, E., Beth, S., Sella-Villa, D. & Silvestro M. (2017). "Alexa, Do You Have Rights?": Legal Issues Posed by Voice-Controlled Devices and the Data They Create. Business Law Today. 1-5.

Brandom, R. (2019, September). To use Alexa, you have to trust Amazon: You're trading privacy for convenience – and the deal keeps getting worse. *The Verge*. https://www.theverge.com/2019/9/26/20885512/amazon-alexa-voice-assistant-privacy-features-trust

Chung, H., Jorga, M., Voas, J. & Lee, S. (2017). Alexa, Can I Trust You? *Computer, 50*(9), 100-104. doi: 10.1109/MC.2017.3571053

Cox, T. (2019, May). The ethics of smart devices that analyze how we speak. *Harvard Business Review*. https://hbr.org/2019/05/the-ethics-of-smart-devices-that-analyze-how-we-speak

Danaher, J. (2018). Toward an Ethics of AI Assistants: an Initial Framework. *Philosophy and Technology, 31*, 629-653. https://doi.org/10.1007/s13347-018-0317-3

Davis, B. (2016, November 7). The problem with voice user interfaces like Amazon Alexa. *eConsultancy*. https://econsultancy.com/the-problem-with-voice-user-interfaces-like-amazon-alexa/.

Day, M. (2019, May 23). Amazon is Working on a Device that can Read Human Emotions. *Bloomberg*. https://www.bloomberg.com/news/articles/2019-05-23/amazon-is-working-on-a-wearable-device-that-reads-human-emotions

Finnegan, M. (2020, January 2). 2020: The year the office finds its voice? *Computerworld.* https://www.computerworld.com/article/3509470/2020-the-year-the-office-finds-its-voice.html

Fruchter, N., Liccardi, I.: Consumer attitudes towards privacy and security in home assistants. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, pp. LBW0501–LBW0506.* ACM, New York (2018)

Green, D. (2018). Big Brother is listening to you: digital eavesdropping in the advertising industry. *Duke Law & Technology Review. 16*(1). 352-392.

Johnson, K. (2019, July 8). Amazon Alexa may soon know if you're happy or sad. *VentureBeat.* https://venturebeat.com/2019/07/08/amazons-alexa-may-soon-know-if-youre-happy-or-sad/

Kleber, S. (2018, May 23). Why our voice assistants need ethics. https://magenta.as/whats-missing-from-siri-and-alexa-ethics-25abdd6b4e7f

Manikonda, L., Deotale, A., & Kambhampati, S. (2018). What's up with Privacy? User Preferences and Privacy Concerns in Intelligent Personal Assistants. *2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES '18)*, 229-235. https://doi.org/10.1145/3278721.3278773

Mattin, D. (2019, July 10). Voice technology could soon be your new best friend. Here's why. *WeForum*. https://www.weforum.org/agenda/2019/07/voice-technology-personalisation/

Prescot, A. (n.d.). Why Alexa is Racist and Sexist. https://digitransglasgow.github.io/ArtificiallyIntelligent/contributions/12_Alexa.html

Prist, A. (2019, December 18). How Voice Assistants Transform the Enterprise. *Medium*. https://medium.com/voiceui/how-voice-assistants-transform-the-enterprise-f2281154049e

Rabassa, V. (2019). Connected objects, voice assistant, digital platform and data: a new way of consuming, an increasing market power for the tech giants? *European Commission conference on shaping competition policy in the era of digitalization.*

Schwartz, E.H. (2019, December 6). Smart Speaker Owners are Uncertain if Personalisation is Worth Sharing Data: Survey. *Voicebot.ai.* https://voicebot.ai/2019/12/06/smart-speaker-owners-are-uncertain-if-personalisation-is-worth-sharing-data-survey/

Temkin, B.R., & Dorsett, B. (2019). Lawyers' Digital Assistants Raise Ethics, Privacy Concerns. https://works.bepress.com/barry_temkin/62/.

Turkington, E. (2019, October 13) What happens when 'generation voice' grows up? *Observer*. https://observer.com/2019/10/voice-assistants-generation-alpha-big-tech-brands-implications/

9 Case study 8: Personalised News

9.1 Objective

The objective of this case study is to describe what personalised news is and to give an overview of ethical issues related to personalised news. These ethical issues are related to individuals (i.e. the news consumers), society at large, and the profession of journalism. Additionally some important matters for legislation are highlighted.

9.2 Technology description

I define personalised news as the selecting of news articles and sources as well as the strategic ordering and placing of articles by a certain media platform, in response to the perceived identity and preferences of a user. The first aspect of this definition – i.e. selecting news and strategically displaying it – is what makes news personalised. The content of articles itself is not necessarily different for different readers, but the articles that are shown or recommended to them and the order in which these articles are shown are tailored to individuals. This tailoring to individuals is done on the basis of personal data, from which information about one's preferences, identity, or location are retrieved. Certain news is then shown to certain individuals because it is expected to be relevant to them (e.g. local events) or in line with their interests.

Personalisation can be implicit or explicit (Haim et al., 2017). Implicit personalisation is the kind of personalisation I here described – i.e. personalisation done by media platforms on the basis of acquired data about individuals. Explicit personalisation is then the personalisation of news on the basis of preferences or interests that users explicitly communicated to news providers, for example via the settings of their account. Instead of distinguishing between implicit and explicit personalisation, some refer to the latter as customisation and the former as personalisation (Edson et al., 2015). Although implicit personalisation refers to what we have defined as personalisation in the introduction, I will here discuss the ethical and legal implications of both implicit and explicit personalisation.

9.3 Context

Everyone who consumes news online will encounter personalisation in some way. This can be on social media, where one can come across articles that friends share or like, or articles from news sources they follow themselves. This encounter with news is personalised because news feeds on for example Twitter, Instagram or Facebook, filter and select information on the basis of an individual's data or profile. Individuals can also get personalised news on news websites. News websites can be websites from a single news source (e.g. The New York Times), where the content that is displayed first or recommended to individuals is selected on the basis of data about that individual (implicit personalisation) or interests that he or she explicitly communicated (e.g. using personalisation settings). News websites can also be platforms that show and suggest articles from different sources (e.g. Google News).

9.3.1 Potential uses

Personalisation of news and information can become even more common than it is today when the technology is improved and/or other types of personal data are used as input. For example, data about one's emotions obtained from smart home appliances with embedded camera's or microphones could be used to personalise content in accordance with one's mood – a very sad or stressed person might want to start the day with some good news, instead of being confronted with suffering elsewhere in the world.

Another possible development is the personalisation of content. Until now, news or information is mainly personalised in that they are shown to individuals in a different order. However, the titles used or style of writing might be changed to make it more appealing to certain audiences, the complexity might be adjusted to a reader's level of comprehension, and shocking details could be left out for younger or vulnerable audiences.

9.4 Relationship to work package

Personalised news, or personalised information, is one of the most obvious examples of personalisation and can therefore not be missing in a work package that seeks to discuss personalisation technologies.

9.5 Relationship to ESR PhD topic

Personalised news is not directly related to my PhD topic, which is the impact of video analytics on individual autonomy. But studying this technology is nevertheless relevant, as ethical concerns can be similar to those related to video analytics and connected to autonomy.

9.6 Review of ethical and legal issues

9.6.1 Methodology

This case study is based on academic literature found by a critical search in Google Scholar, Scopus and HeinOnline. In addition, I used the popular science book *The Filter Bubble: What the Internet is Hiding from You* by Eli Pariser, which is much cited in the found academic literature on personalised news or information. I used a mixture of search terms: 'personal*' AND 'news' or 'information' AND 'ethic*' or 'legal'. I selected search results that appeared to be relevant from the first 10 pages shows. In selecting articles I did not take discipline, journal, or year of publishing into account.

9.6.2 Findings

Personalised news raises a variety of ethical concerns, about the impact on individuals, on society and on journalism. I have divided these concerns in different categories and discuss them below. Although there is a section focusing on legal issues, many of the ethical issues discussed in other sections have a legal dimension too. In the final section I point out some benefits or 'ethical aspects', rather than 'ethical issues', that are related to the technology.

Filter bubble

In his book *The Filter Bubble: What the Internet is Hiding from You*, Pariser discusses the idea that personalisation of information will create a 'filter bubble' in which users only come across news articles, video's, posts, and so on, that confirm their interests and worldviews. Similarly, some speak of 'echo chambers' (Sunstein, 2002) or an 'information cocoon' (Sunstein, 2018). A filter bubble is problematic because it can prevent people from being confronted with different political views, discovering new interests, or fully knowing what is happening in the world. According to Sunstein, a democratic society requires that citizens 1) "exposed to materials that they would not have chosen in advance", 2) "have a wide range of common experiences" and 3) are able to "distinguish between truth and falsehood" (Sunstein, 2018, p. 85).

However, Pariser's filter bubble might be not more than just a theory. Haim et al. (2017) conclude that the negative effects are not found in practice, they write that

Overall, our findings suggest that the filter-bubble phenomenon may be overestimated in the case of algorithmic personalisation within Google News. (…) while personalisation effects were visible (which provides support for the applicability of our method), the results did not blind out essential shares of information (which the filter-bubble hypothesis would suggest).

More studies would be needed to show to what extent filter bubbles, echo chambers and information cocoons are really the case and whether they really have the expected effects on individuals and society. The existence of a filter bubble and its related consequences will also differ between platforms that offer personalised news feeds, as the data and algorithms they use for personalisation can vary significantly.

Transparency

The filtering of information, in the case of implicit personalisation, is an opaque process. How news articles are filtered is not visible to users. In other words, users do not know why they are shown certain articles, from certain sources, in a certain order. Various scholars have argued that it should be more transparent to users how information is filtered before it is presented to them (Bozdag & Timmermans, 2011; Diakopoulos & Koliska, 2016; Koene et al., 2015; Pariser, 2012). Personalised news filters have a major impact on individuals, because they define what users render to be important or real (Pariser, 2012). This influence put individuals in a vulnerable position. In addition to that, individuals cannot give informed consent to the personalisation of news if they do not know how the news they consume is filtered (Koene et al, 2015). To strengthen the position of users, they should be given more information or even control over the filtering processes (Bozdag & Timmermans, 2011). However, as Diakopoulis and Koliska (2016) point out, there is little business incentive to disclose information about personalisation algorithms. Also, giving users insight into the algorithms used to personalise news might not do much good, as many of them will not be able to understand this information or be overwhelmed by the excess of information. Therefore, Diakopoulis and Koliska suggest that as set standards to guide algorithmic filtering should be introduced.

Autonomy & Privacy

Autonomy entails governing one's own desires and making independent choices to pursue those desires. The lack of control over and knowledge of algorithmic filtering in personalised news can harm people's autonomy, because it limits people's ability to make informed choices or to make choices at all (Bozdag & Timmermans, 2011). Having an algorithm decide what information is of interest to one also keeps people from governing their own desires. Furthermore, autonomy might be harmed because data analytics makes it possible to detect users' vulnerabilities, which would enable another to manipulate users (Ignatidou, 2019; Koene et al., 2015). Having and gathering such vulnerable data from individuals is not only compromising autonomy, but also privacy. The creation of user profiles too can be understood as a breach of users' privacy (Koene et al., 2015).

Identity

Several concerns have been raised about the effect of personalised news on individuals' identity. As Reijers et al. (2016, p. 130) note "personalisation invariably integrates underlying assumptions about what a person is". Today, in personalised news the underlying assumption about personhood, or identity, is that it is static (Ignatidou, 2019). Pariser (2012) refers to this as the "one-identity problem". He argues that, instead of acknowledging differences between a 'work-self' and 'home-self', or between short term and long term desires, personalisation filters wrongfully assume that one's identity is static and continuous over time. What is more, the recommender systems do not take one's context into account (Bozdag & Timmermans, 2011).
Not only is identity not a static thing, media can shape identity. As personalised news plays an important role in what news an individual consumes, it plays an important role in the shaping of identity. Reijers et al. (2016) have analysed this process as a "narrative shaping", they argue that personalised news configures a narrative understanding of the self and the world. This shaping, however, is a mutual process. That is, in personalised news, an individual's identity shapes the news content while the news in turn shapes that individual's identity (Pariser, 2012).

The effect of personalised news on the identity of individuals can be seen as an ethical problem because it harms the autonomy of individuals, but also because it is a powerful tool that can be taken advantage of (e.g. as a propaganda tool).

Fairness

Algorithmic filtering of news articles and sources can enclose human biases or forms of discrimination. For example, it can show news based on one's predicted IQ level. Which means that, if the algorithm thinks you are 'stupid', you get less news from well established newspapers and more gossip about celebrities (Pariser, 2012). This implicit discrimination can be more subtle than, for example, discrimination by algorithms used for predictive policing or filtering resume's (Ignatidou, 2019). Not just individuals can be harmed by such discrimination, it also threaten the fairness and equality of political discourse (see 'Democracy and societal issues' below) and enforces a kind of classicism in society.

However, personalised news can also have positive implications with regards to fairness. Traditionally, a newspaper's editor would decide which articles are shown first – whether it is on a website or in an actual paper. This means that, news about local matters or minorities in a society would never make the front page. In other words, it seems to ignore their problems and their role in a society. Personalised news can promote inclusivity, it can show individuals the news that is most relevant to them, that acknowledges their place in society and their concerns, and make news more meaningful to all members of society (Ignatidou, 2019).

Democracy and societal issues

As mentioned, personalised news can promote inclusivity and make news more relevant to an individuals' identity and preferences. The downside of this is that the potential of serendipity or confrontation with the Other are side-lined. Personalised news can be a threat to democracy because it only confirms people's political views, rather than confront them with other perspectives, that can challenge their views or even change them. In addition to that, the news people are shown might intentionally manipulate their political views (e.g. Cambridge Analytica's involvement in the 2016 US elections and the Brexit campaign). News personalisation algorithms have agenda setting power (Helberger et al., 2016).

Furthermore, if people are only shown news that involves members of their direct community, this gives a distorted view of the members of society. Certain groups in society might become ignorant about the presence and struggles of other groups in society, which implies that their voice – or call for help – cannot be heard. Related to this, is the so-called 'friendly world problem' (Pariser, 2012). If content that individuals are more likely to like is prioritized, people might no longer see the negative news – e.g. about poverty, war or injustice elsewhere in the country or in the world.

Changing nature of journalism

Traditionally, media could protect that people were confronted with the troubles going on in different parts of society or the world, by simply putting that news on a front page. Media also played an important role in maintaining transparency of political discourse (Mittelstadt, 2016) and holding power to account (Ignatidou, 2019). The emergence of personalised news means that media move away from mass communication (Thurman & Schifferes, 2012). It makes it more difficult for traditional media sources to promote a fair political discourse, to encourage serendipity or respond to public needs. As traditional newspapers or media do not control what articles are shown by other news sites (e.g. Google News), they loose editorial autonomy. Personalised news and the emergence of different news sites that offer news from various sources, has changed the dynamic between journalists and commercial interests of media (Ignatidou, 2019).

Legal issues

Personalised news can raise various legal issues, some of which are closely related to the aforementioned ethical issues. A first legal issue has to do with privacy and data protection.

Legislation should protect that there are limits to what data can be used to personalise news and whether third parties can have access to such data. Secondly, legislators should question when and to what extent personalised information is acceptable. As Cavender (2017) points out, personalisation can threaten fundamental rights such as the freedom of expression, which in turn threatens democracy (see discussion above). According to Cavender, corporations should not be free to decide what information is shown to individuals and what not, instead adequate legislation should be put in place by governments to protect individuals' free and uninhibited access to information. Similarly, Helberger et al. (2016) argue that there should be clear legislation to prevent the use of personalised news to manipulate political views during political campaigns. Also, regulation should be in place to ensure that personalisation serves the interest of individuals and not solely of advertisers – which is the case today according to Feuz et al. (2011).

Benefits

Of course personalised news is not without benefits. Personalisation makes news feeds more relevant for users, which might cause some individuals to consume at least some news articles where they otherwise would not at all. Another aspect of this relevance is that it can show individuals local news items, that are likely to be relevant to him or her, but otherwise would not make a front page. Personalisation also makes news consumption less time consuming, it is an efficient way to get access to the items you are likely to be looking for otherwise. Furthermore, if a downside of personalisation is that individuals lack broad information about what is happening in the world, then an upside is that they can gain a deeper insight into issues that particularly interest them and are more likely to come across likeminded people.

9.7 References

Bozdag, E. & Timmermans, J. (2011). Values in the filter bubble Ethics of Personalisation Algorithms in Cloud Computing. 1st International Workshop on Values in Design – Building Bridges between RE, HCI and Ethics, Lisbon (Portugal) September 6, 2011.

Bozdag, E. (2013). Bias in algorithmic filtering and personalisation. *Ethics and Information Technology*, 15. 209-227. DOI 10.1007/s10676-013-9321-6

Cavender, B. (2017). The Personalisation Puzzle. *Washington University Jurisprudence Review, 10*(97). 97-121.

Diakopoulos, N. & Koliska, M. (2016). Algorithmic Transparency in the News Media. *Digital Journalism.* DOI: 10.1080/21670811.2016.1208053

Edson C., Tandoc Jr. & Ryan J. T. (2015). The Ethics of Web Analytics. *Digital Journalism, 3*(2). 243-258. DOI: 10.1080/21670811.2014.909122

Eskens, S., Helberger, N. & Moeller, J. (2017). Challenged by news personalisation: five perspectives on the right to receive information. *Journal of Media Law, 9*(2). 259-284. DOI: 10.1080/17577632.2017.1387353

Feuz, M., Fuller, M., & Stalder, F. (2011). Personal Web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation. *First Monday*, *16*(2). https://doi.org/10.5210/fm.v16i2.3344

Haim, M., Graefe, A., & Brosius, H. (2017). Burst of the Filter Bubble? Effects of personalisation on the diversity of Google News. *Digital Journalism*. DOI: 10.1080/21670811.2017.1338145.

Helberger, N., Irion, K., Möller, J., Trilling, D., de Vreese, C.H. (2016). Shrinking core? Exploring the differential agenda setting power of traditional and personalised news. *Info, 18* (6). 26-41. doi:http://dx.doi.org/10.1108/info-05-2016-0020

Ignatidou, S. (December 2019). *AI-driven Personalisation in Digital Media: Political and Societal Implications*. Chatham House Research Paper.

Koene, A., Perez, E., Carter, C.J., Statache, R., Adolphs, S., O'Malley, C. ... (2015). Ethics of personalised information filtering*. Proceedings of Internet science: second international conference*, INSCI 2015, Brussels (Belgium), May 27-29, 2015. Springer International Publishing. 123-132.

Lewis, S.H. & Westlund, O. (2014) Big Data and Journalism: Epistemology, expertise, economics, and ethics. *Digital Journalism*. 1-20. http://dx.doi.org/10.1080/21670811.2014.976418

Mittelstadt, B. (2016). Auditing for Transparency in Content Personalisation Systems. *International Journal of Communication*, 10. 4991-5002.

Pariser, E. (2012). *The Filter Bubble: What the Internet is Hiding from You*. Penguin Books.

Portilla, I. (2018). Privacy concerns about information sharing as trade-off for personalised news. *El profesional de la información, 27*(1). 19-26. https://doi.org/10.3145/epi.2018.ene.02

Reijers, W., Gordijn, B., & O'Sullivan, D. (2016). Narrative Ethics of Personalisation Technologies. In Kreps, D., Fletcher G., & Griffiths, M. (Eds.), *Technology and Intimacy: Choice or Coercion,* proceedings of 12th IFIP TC 9 International Conference on Human Choice and Computers, Salford (UK), September 7-9, 2016. Springer International Publishing. 130–140. http://doi.org/10.1007/978-3-319-44805-3_11

Sunstein, C. (2002). The law of group polarization. *The Journal of Political Philosophy, 10*(2). 175–195.

Sunstein, C. (2018). Is Social Media Good or Bad for Democracy? *Sur - International Journal on Human Rights, 15*(27). 83-89.

Taylor, D.G., Davis, D.F. & Jillapalli, R. (2009). Privacy concern and online personalisation: the moderating effects of information control and compensation. *Electronic Commerce Research*, 9. 203–223. DOI 10.1007/s10660-009-9036-2

Thurman, N. and Schifferes, S. (2012). The Future of Personalisation at News Websites: Lessons from a Longitudinal Study. *Journalism Studies, 13*(5-6). doi: 10.1080/1461670X.2012.664341

Treiblmaier, H., Madlberger, M., Knotzer, N. & Pollach, I. (2004). Evaluating Personalisation and Customization from an Ethical Point of View: An Empirical Study. *Proceedings of the 37th Hawaii International Conference on System Sciences*, January 5-8, 2004.

10 Case study 9: The use of facial recognition for personalisation

10.1 Objective

The objective of the present case study is to disclose the possible uses of facial recognition for personalisation purposes and what ethical and legal implications such uses of facial recognition can have. To this end, I will look not only at literature about uses of facial recognition for personalisation, but also about literature about the ethics of facial recognition technologies in general. However, the focus of this literature review is to find what ethical and legal issues are specifically related to the use of facial recognition for the personalisation of products or services.

10.2 Technology description

'Facial recognition' refers to the use of computer vision techniques to analyse individuals' facial features and make inferences about their identities, emotions, or personal traits like age, gender or ethnicity (Mery, 2019; Sample, 2019). In some cases, the term 'facial recognition' is used to refer solely to the identification of individuals. In the present context, however, I will employ the term in a broader sense, including all traits that could be recognized from one's facial features.

In short, facial recognition works as follows: deep learning algorithms are used to detect faces on videos or images and map out facial features, these facial features are then analysed by comparing them against a large database of images of faces (Mery, 2019; Sample, 2019). To recognise one's identity, the given image needs to be compared with a database of faces that contains that individual's image. But even when one's identity cannot be determined, personal traits or emotions are sought to be recognised from someone's facial features.[31]

Facial recognition, or facial analysis, can facilitate the personalisation of products or services. First of all, services can be personalised by the recognition of emotions, facial expressions, or traits like age or gender. Emotion recognition can be used to personalise services from social robots (Khosla et al., 2015), for automatic pain recognition of patients (Thiam et al., 2017), to help autistic children understand their own emotions in a personalised way (Gay et al., 2013)[32], or to personalise advertisements by tailoring them to the mood of a consumer (Chatterjee, 2019; Jess, n.d.). Recognising facial expressions also makes it possible to detect personal highlights when one watches videos, which enables the personalisation of future video summaries and recommendations (Joho et al., 2011), or to personalise one's gaming experience (Blom et al., 2014). Moreover, the analysis of consumers' facial expressions when they see an add or look at products in a store can

[31] It should be noted that there is no consensus on the idea that emotions can be recognized from facial expressions, which means that there is no solid scientific foundation for the use of emotion recognition for personalisation or other purposes.

[32] The example referred to here is the app *CaptureMyEmotion*, that allows users to take selfies and uses emotion recognition to label their emotions and teach them about their own emotions.

provide valuable feedback that can be used to improve the efficiency of advertisements and product display (Jess, n.d.). Information about gender, age, or other traits that can help to personalise advertisements that one is shown in public spaces or it can be combined with other data, like geolocation, to increase the efficiency of targeted advertisements (Jess, n.d.).

A second way in which facial recognition can enable the personalisation of products or services is by identifying a known user or customer. For example, when several household members use the same device, this device can adjust its settings to each user upon recognising them (McDonogh, 2012). Another example would be the personalisation of customer service in stores or in the hospitality industry (McDonogh, 2012; Munyaradzi et al., 2014). When one is a regular customer of a certain bar, hotel or supermarket, he or she can be recognized upon entrance and services can be tailored to that customer on the basis of his or her customer profile.

The here mentioned examples of the use of facial recognition for personalisation are examples that can be found in the literature today. What is clear from the found examples, is that facial recognition is a very useful tool for personalisation. As facial recognition is still an emerging technology, its uses for personalisation purpose can increase in the coming years. The future of facial recognition and other biometric technologies might not be in security, as is often thought, but in personalisation (McDonogh, 2012).

10.3 Context

Facial recognition can be used in other ways, where it is less clear whether the service or product can be deemed 'personalised'. One example of this is the use of facial recognition to provide access to apps, devices, or spaces. Facial recognition can be used via the camera on one's phone or on a doorbell, to grant access only to those who are recognised and authorised by the system. Facial recognition as a means of access has some advantages in comparison to keys or PINs, the latter are less secure because they can "get lost, can be found or stolen by others, and can also be given to third parties" (Nonak & Hahn, n.d.). However, other biometrics, such as fingerprints, can be used for this as well and might be less intrusive.

Facial recognition can also facilitate and change the shopping process and experience in several ways. For example, facial recognition can enable self-checkout shopping, which entails that a camera recognizes a costumer and keeps track of the products he takes, to withdraw the costs of those products from the customer's pre-registered bank account. Facial recognition can also be used for augmented reality (AR) applications, that allow customers to try on clothes or make-up – either at home via an app or in the store via a smart mirror. A known example of this is the Visual Artist app of beauty brand Sephora (Chatterjee, 2019; Jess, n.d.), that enables customers to try on make-up products at home using their smartphone's front camera. Luxury brand Rebecca Minkoff uses facial recognition to offer automated customer services in stores, which is said to better suit the tastes of millennial audiences, who frequently avoid personal communication with store assistants (Jess, n.d.). Automated customer advice would also prevent social friction in the shopping experience, i.e. feeling judged by staff members (Arthur, 2017).

These are examples of the use of facial recognition to offer personal services, such as granting access or giving shopping advice, but they are not necessarily *personalised* products or services. Therefore, only the examples mentioned in paragraph 1.1.2 will be taken into consideration in the review of ethical and legal issues below.

10.3.1 Potential uses

The example uses of facial recognition for personalisation that are mentioned in 1.1.2 are all examples that are mentioned in the literature, but not yet widely embedded in society. In a way, therefore, these are all potential uses of facial recognition for personalisation.

10.4 Relationship to work package

In this work package we discuss the ethics of personalisation. Facial recognition offers tremendous opportunities for enabling the personalisation of products or services. It is therefore of relevance to this work package to disclose what ethical and legal issues the use of facial recognition might raise, in particular what issues arise from the use of facial recognition for the purpose of personalisation.

10.5 Relationship to ESR PhD topic

For my PhD, I look into the impact of video analytics or computer vision applications on human autonomy. Facial recognition is one such application, hence the present case study is closely related to my PhD research.

10.6 Review of ethical and legal issues

10.6.1 Methodology

The literature used for the present review of ethical and legal issues, as well as the technology description provided in 1.1.2 and 1.1.3, was found by means of a critical search in Google Scholar, Scopus and Hein Online. A few of the cited articles were found by a backwards snowballing technique. At first, I used a mixture of search terms: ‘personalisation’ or ‘personalised’ AND ‘facial recognition’ AND ‘ethic*’ or ‘moral’. However, the relevant results were very limited. Therefore searched instead for 1) ‘personalisation’ or ‘personalised’ AND ‘facial recognition’, and 2) ‘facial recognition’ AND ‘ethic*’ or ‘moral’. In the case of Google Scholar, the search results were over 10 pages, so I limited the selection of search results to the first 10 pages.

10.6.2 Findings

The critical search did not enable me to find literature that specifically deals with the ethical and/or legal implications of the use of facial recognition for the personalisation of products or services. The search also showed that there are a number of articles discussing examples of personalised products or services that implement facial recognition. However, the use of facial recognition for personalisation nevertheless seems to be an under-researched topic. A reason for this can be that applications of facial recognition that can be understood as personalised technologies, are not explicitly given that label. Another reason might be that the use of facial recognition for security ends is given more attention in research than its use for personalisation purposes. However, as the brief technology description in 1.1.2 shows, facial recognition can be a very useful tool for the personalisation of goods and services.

Because no literature on the ethical or legal implications of using facial recognition for personalisation was found, there are no ethical or legal issues to be reviewed in this section. There is, however, a much richer literature on the ethics of facial recognition in general, which can help to determine what ethical issues might be related to the use of facial recognition for personalisation purposes. Therefore, I will here give a brief overview of the found ethical and legal issues regarding facial recognition in general.
In an early article on the ethics of using facial recognition in public spaces, Brey (2004) argues that the main risks are error, function creep and privacy. Error of facial recognition, like false positives or negatives (Introna, 2005), are especially problematic when facial recognition is used for security purposes. For example, facial recognition algorithms have been showed to work significantly worse on people with darker skin tones, which means that black people are more likely to be misidentified as criminals (Garvie & Frankle, 2016; Jess, n.d.). Racial bias in facial recognition algorithms can also be problematic when facial recognition is used for personalisation purposes. If products or services would work less well for people with darker skin tones, this is a kind of (unfair) social exclusion.

Function creep and privacy are ethical as well as legal issues. When facial recognition algorithms are widely available and cameras are present in all spheres of one’s life, the technology can potentially be used for other purposes than initially planned or communicated. Given the sensitive information that computer vision can reveal, the widespread use of camera’s and facial recognition technologies can threaten people’s privacy. Because of that, it should be critically assessed whether personalised services can be offered by other, less intrusive means (Jess, n.d.; Wiewíorowski, 2020).

Personalisation by means of facial recognition can, of course, have positive consequences too. The well-being of elderly can increase by personalised care robots (Khosla et al., 2015), personalised gaming makes the gaming experience available for a larger social group (Blom et al., 2014), it can be of great value to the social development of autistics children, and self-checkout shopping and automated customer service can increase people’s autonomy, in that it allows them to gain full control over the shopping process (Jess, n.d.). A positive aspect of the use of facial recognition in general is the security potential it offers on an individual level. Facial recognition can replace use of keys, cards or codes, solving the problem of lost or stolen keys, passwords, and so on. Furthermore, the use of facial recognition to unlock phones or homes can also be more convenient or efficient to an individual, especially when this person has a disability that makes it more challenging to use traditional entry methods like keys. Finally, facial recognition as a means of entries or identification is more hygienic that the use of keys or having to touch others’ ID-cards. Which, in light of the COVID-19 pandemic might be an important consideration in choosing to adopt facial recognition technologies.

10.7 References

Arthur, R. (2017) "The Automated Future of the Fashion Store: Where Self-Checkouts and Human Touch Collide", [online] Available at: https://www.forbes.com/sites/rachelarthur/2017/02/01/the-automated-future-of-the-fashion-store-where-self-checkouts-and-human-touch-collide/#5fdd28985d6c [Accessed on 25 June 2020].

Blom, P.M., Bakkes, S., Tien Tan, C., Whiteson, S., Roijers, D., Valenti, R., & Gevers, T. (2014). Towards Personalised Gaming via Facial Expression Recognition. *Proceedings of the Tenth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE 2014)*. 30-36.

Brey, P. (2004). Ethical aspects of facial recognition systems in public spaces. *Journal of Information, Communication and Ethics in Society, 2*(2). 97-109. https://doi.org/10.1108/14779960480000246

Broad, E. (October 4, 2017). *Who gets held accountable when a facial recognition algorithm fails?* Retrieved from http://ellenbroad.com/facial-recognition-who-gets-held-accountable/

Chatterjee, S. (2019). Big Data Analytics in e-Commerce: Understanding Personalisation. *2nd International Workshop on Advances in Social Sciences (IWASS 2019)*. 201-208. DOI: 10.25236/iwass.2019.029

Connolly, E. (March 15, 2019). *Facial Recognition Technology Raises Ethical, Legal, Privacy Concerns in Health Care*. Retrieved from https://www.psychiatryadvisor.com/home/practice-management/facial-recognition-technology-raises-ethical-legal-privacy-concerns-in-health-care/

Crawford, K. (2019). Regulate facial-recognition technology. *Nature, 572*. 565.

Garvie, C. & Frankle, J. (April 7, 2016). Facial-Recognition Software Might Have a Racial Bias Problem. *The Atlantic.* Retrieved from https://apexart.org/images/breiner/articles/FacialRecognitionSoftwareMight.pdf

Gay, V., Leijdekkers, P., Agcanas, J., Wong, F., & Wu, Q. (2013). CaptureMyEmotion: Helping Autistic Children Understand their Emotions Using Facial Expression Recognition and Mobile Technologies. *BLED 2013 Proceedings*. 10. Retrieved from http://aisel.aisnet.org/bled2013/10

Introna, L.D. (2005). Disclosive ethics and information technology: disclosing facial recognition systems. Ethics and Information Technology, 7. 75-85. DOI 10.1007/s10676-005-4583-2

Jess, C. (n.d.). *Facial Recognition Technology: New Marketing Opportunities and Challenges*. Retrieved from https://15writers.com/facial-recognition-technology-new-marketing-opportunities-and-challenges/

Joho, H., Staiano, J., Sebe, N. & Jose, J.M. (2011). Looking at the viewer: analysing facial activity to detect personal highlights of multimedia contents. *Multimedia Tools and Applications*, (51). 505-523. DOI 10.1007/s11042-010-0632-x

Khosla, R., Nguyen, K., Chu, M.T. & Tan, Y.A. (2015). Robot Enabled Service Personalisation Based on Emotion Feedback. *MoMM2016*. 28-30. DOI: 10.1145/3007120.3007167

Kriebitz, A. & Lütge, C. (2020). Artificial Intelligence and Human Rights: a Business Ethical Assessment. *Business and Human Rights Journal, 5*. 84-104. doi:10.1017/bhj.2019.28

McDonogh, E. (2012). Biometrics for the mass market – are we ready? *Biometric Technology Today*. 9-11.

Mery, D. (2020). Face Analysis: State of the Art and Ethical Challenges. Dabrowski J., Rahman A., Paul M. (eds) Image and Video Technology. PSIVT 2019. Springer. 14-29. https://doi.org/10.1007/978-3-030-39770-8_2

Munyaradzi, M., Prudence, M., & Tarirayi, M. (2014). Use of Facial Recognition for Data Personalisation in Customer Relationship Management (CRM): Case of Great Zimbabwe Hotel. *International Journal of Engineering Trends and Technology, 14*(1).

Nouak, A. & Hahn, V. (n.d.). Access control and surveillance. https://link.springer.com/content/pdf/10.1007%2F978-3-540-88546-7_98.pdf

Sample, I. (July 29, 2019). What is facial recognition – and how sinister is it? *The Guardian*. Retrieved from https://www.theguardian.com/technology/2019/jul/29/what-is-facial-recognition-and-how-sinister-is-it.

Thiam, T., Kessler, V., & Schwenker, F. (2017) Hierarchical Combination of Video Features for Personalised Pain Level Recognition. *ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.* Bruges (Belgium), 26-28 April 2017, i6doc.com publ.

Wickins, J. (2007). The ethics of biometrics: the risk of social exclusion from the widespread use of electronic identification. *Science and Engineering Ethics, 13.* 45-54. DOI 10.1007/s11948-007-9003-z

Wiewiórowski, W. (February 21, 2020). *AI and Facial Recognition: Challenges and Opportunities*. European Data Protection Supervisor Blog. Retrieved from https://edps.europa.eu/press-publications/press-news/blog/ai-and-facial-recognition-challenges-and-opportunities_en.

11 Case study 10: Personalised Medicine – Continuous Glucose Monitors

11.1 Objective

This case study will present and address selected ethical and legal aspects of the use of continuous glucose monitoring sensors (CGM). These will include issues and opportunities connected to personalisation done by CGM and those generally associated with CGM as an example of a personalisation technology.

11.2 Technology description

Continuous glucose monitors are personal medical devices worn by type 1 diabetic patients in their everyday life. They take constant readings of blood glucose levels without requiring any involvement of the patient. Although they can be used for purely informational purposes similarly to other self-tracking devices, they are commonly coupled with insulin pumps which administer glucose when the blood glucose level of the user reaches a specific threshold. This makes it possible to formulate personalised treatments for diabetic patients. Whereas in traditional diabetes management patients are asked to take insulin at predetermined times or at the onset of certain symptoms, with the help of CGM insulin injections can be synchronised with live blood glucose readings and consequently administered when they are most needed, possibly increasing the efficiency of treatment. Additionally, when coupled with insulin pumps, the use of CGM can make diabetes management less invasive and more convenient – patients no longer have to strictly monitor parameters and stick the needles themselves as the process is automated and the insulin is administered by a wearable device that can be concealed under clothes and does not inflict pain. Moreover, CGM can also be used together with smartphone and other apps in order to provide patients with recommendations relating to their level of activity, diet and general well-being – based on their blood glucose level, a patient could see suggestions to eat or abstain from eating something sweet, engage in more exercise on that day or even contact a doctor. Combined with broader health apps and devices, CGM can be used as part of more general personalised healthcare solutions. Commonly used CGM include devices manufactured by Dexcom, FreeStyle Libre, and GlucoTrack, although this varies in different parts of the world.

11.3 Context

The use of CGM stretches beyond the clinical contexts as patients are able to wear these devices in their everyday life and make use of data and recommendations provided by CGM on their own. However, since the employment of these sensors is part of type 1 diabetes treatment, actors relevant for this study include diabetic patients and their families, but also healthcare professionals and private and public institutions involved in providing healthcare, such as insurance companies and government agencies. Ethical and legal issues affect all of these actors to a significant extent.

11.3.1 Potential use

Personalised healthcare solutions are often discussed as the future of healthcare (Dunn et al., 2018) and CGM are not an exception – proponents and developers of these devices promise that they would facilitate or even possibly automate the management of diabetes, while simultaneously limiting costs and reducing the workload of healthcare professionals – some discussion goes as far as to mention the future possibility of turning CGM into an “artificial pancreas” (Farrington, 2018; Quintal et al., 2019). Patient empowerment is also one of the promises, as CGM are supposed to give users greater insight into their condition, bringing both psychological and practical benefits to their lives. Last but not least, the data collected through CGM is sometimes lauded for its potential to change diabetes research and possibly lead to new ways of addressing the illness.

11.4 Relationship to Workpackage

CGM are a personalisation technology as the data they collect serves as a basis for the personalised health and lifestyle recommendations provided by algorithms and doctors, as well as for personalised treatment, i.e. the administering of insulin through the insulin pump coupled with the CGM.

11.5 Relationship to ESR PhD topic

The subject of my PhD research are the ethical issues surrounding self-tracking practices and technologies. This case study analyses CGM, which are an example of self-tracking technology, as well as medical and care practices involving these sensors.

11.6 Review of ethical and legal issues

11.6.1 Methodology

The foundation for this case study is a literature search conducted in Google Scholar, Scopus and Hein Online. In order to extract the most relevant articles and provide a comprehensive overview of ethical and legal aspects of CGM, I used a mixture of search terms from two sets combined into search strings. The first set included the terms ‘personal* medic*’, ‘continuous glucose monitor*’ ‘self-track* AND medic*’, and the second included the terms ‘ethic*’, ‘moral*’, ‘virtue*’ and ‘legal*’, and all have been adjusted depending on the search database. Some of the discussed articles deal with personalised medicine in general, but their findings are applied here to CGM, where relevant. Moreover, some papers identified general ethical and legal issues and opportunities connected to CGM and served as a basis for inferring aspects connected to personalisation in particular. A standard Google search helped with identifying press articles discussing CGM.

11.6.2 Findings

Although CGM have an undeniable positive impact on the well-being of diabetic patients (Dunn et al., 2018; Zapatka, 2019), some ethical concerns surrounding their use can be identified (Quintal et al., 2019). While management of diabetes has always involved frequent testing of blood glucose level, CGM conduct such tests virtually all the time, making this kind of continuous testing the new standard. As such, it could be argued that the use of CGM and other similar personal medical devices is not merely an option for diabetic patients, but might be seen as **the only responsible choice** (Crawford, 2004, 2006; Lupton, 2012, 2013a). Traditional methods of monitoring are inconvenient, more time consuming and not as reliable (Becker, 2017; Zimberoff, 2015) and patients opting not to use CGM could be seen as failing to follow medical advice, similarly as if they opted not to follow the prescribed diet or maintain the recommended level of activity. Furthermore, the greater accuracy of CGM could mean that patients willing to treat their health "seriously" would be **expected or even forced to monitor a greater number of potentially relevant factors** and to do so in a greater number of circumstances (Lupton, 2016). Consequently, while it is true that CGM are helpful in the management of diabetes, they could also be seen as **putting pressure on patients** by requiring them to take more responsibility over their health, and **inducing greater or even excessive care for health**, potentially overmedicalizing the life of patients or contributing to health-related anxiety (Lupton, 2013b).

Alternatively, overreliance on CGM could lead to a **false sense of security for patients** or even prompt them to ignore traditional ways of treatment and management of diabetes such as regular exercise and proper diet (Farrington, 2018; Smith & Vonthethoff, 2017), a development especially dangerous in case of device malfunctions (O'Connor, 2019). The great complexity of these devices could also mean that patients will not be fully aware of the CGM's purpose and will not be able to give **informed consent** to their use (Klugman et al., 2018). Consequently, the use of CGM and their possible negative impact on health perception is an area ripe with ethical issues and some scholars have suggested that it could potentially benefit from increased regulation (Carroll, 2014).

The increased use of CGM also **impacts the role of medical professionals**. The possibility to remotely monitor the condition of patients might mean that medical professionals will spend more time investigating data in front of computers rather than engaging in direct contact with their patients (Gabriels & Moerenhout, 2018), especially if CGM are used by health providers as a cost-cutting measure (Coughlan, 2006). Additionally, the automated management of disease enabled by data-driven medicine might mean that the **patients' needs are not fully addressed** and that the **autonomy** of individuals involved in the treatment might be at risk (Rich & Miah, 2017). Sharon argues that even if new kinds of **relationships and solidarity** would develop among patients and practitioners as a result of the use of personal medical devices, it is not certain that this would be to the benefit of everybody involved (Sharon, 2017).

Moreover, since CGM could offer users personalised recommendations and personalised treatment independently of the expertise of the physicians, the **norms and algorithms used by the developers of the app merit great scrutiny**: are the solutions offered by CGM really personalised or do they follow pre-existing models, only slightly adapting those to the needs of individual patients? In traditional management of diabetes, recommendations and treatment would be adjusted during regular check-ups – this is not necessarily the case with CGM and other health self-tracking devices which often **apply the same model to all users**, regardless of their differences (Kent, 2018). These concerns clearly show that personalised medicine is at risk of becoming **excessively impersonal**, **enforcing generalised norms** and prompting the users to engage in **self-discipline** in order to fit them (Ajana, 2017). Moreover, the relatively high individual cost of CGM (Zimberoff, 2017) could also increase **health inequalities** (Fox, 2017). Similarly, Piras and Miele note that the money and effort required to self-track for health could mean that patients needing such health interventions the most, might not actually be able to use them (Piras & Miele, 2017). Patients unable to afford mainstream solutions or not satisfied with the current offers have also attempted to create their own devices, which, due to the lack of regulatory approval and oversight, could expose them to great dangers (Kim, 2019).

Finally, there are also concerns related to **data management, privacy, law and policy**. The increased complexity of “appified” medicine could make it more difficult to create robust laws and policies addressing technologies like CGM (Terry, 2018). Scelsi (2015) argues that the increased use of internet-connected devices in healthcare might create new **threats to privacy**, while Evans (2016) discusses how consumer interests and the public good potentially arising from personal health data will need to be carefully balanced in order to meet the challenges related to data management. This is particularly relevant if we consider that the ownership of data by private companies could effectively **limit access to data** (Sharon, 2016). In turn, Rowe (2018) outlines concerns related to the sharing of data produced by personal medical devices, arguing that it is **not clear, who owns the data**, who can access it and who will be made aware of potential device malfunctions. She also notes that insurance companies could use this data to inform decisions that would not be beneficial to patients. Although the authors discussed legal, privacy, policy and data management issues connected to personal medical devices in general, their suggestions and concerns are also relevant to CGM in particular.

Despite these issues, however, it has to be noted that CGM come with several advantages which need to be taken into account when analysing these technologies. Most prominently, some authors observed that their use has a potential to improve patient wellbeing and help develop healthy habits, while improving public health outcomes and reducing costs (Ajana, 2017; Dunn et al., 2018; Zapatka, 2019; Zimberoff, 2017). Others noted the potential personalised medical technologies have to improve diabetes research and even give patients more agency in clinical contexts while helping them make sense of their condition (Dunn et al., 2018; Evans, 2016; Farrington, 2018; Gabriels & Moerenhout, 2018).

11.7 References

Ajana, B. (2017). Digital health and the biopolitics of the Quantified Self. *Digital Health*, *3*, 1–18.

Becker, R. (2017, May 25). *Apple’s needleless blood sugar tracker has an uphill battle in front of it*. The Verge. https://www.theverge.com/2017/5/25/15685148/apple-watch-glucose-tracker-blood-sugar-monitoring-diabetes

Carroll, N. R. (2014). Mobile Medical App Regulation: Preventing a Pandemic of Mobilechondriacs. *St. Louis University Journal of Health Law & Policy*, *7*(2), 415–448.

Coughlan, R. (2006). The Socio-Politics of Technology and Innovation: Problematizing the 'Caring' in Healthcare? *Social Theory & Health*, *4*(4), 334–352. https://doi.org/10.1057/palgrave.sth.8700078

Crawford, R. (2004). Risk Ritual and the Management of Control and Anxiety in Medical Culture. *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*, *8*(4), 505–528. https://doi.org/10.1177/1363459304045701

Crawford, R. (2006). Health as a meaningful social practice. *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*, *10*(4), 401–420. https://doi.org/10.1177/1363459306067310

Dunn, J., Runge, R., & Snyder, M. (2018). Wearables and the medical revolution. *Personalised Medicine*, *15*(5), 429–448. https://doi.org/10.2217/pme-2018-0044

Evans, B. J. (2016). Barbarians at the Gate: Consumer-Driven Health Data Commons and the Transformation of Citizen Science. *American Journal of Law & Medicine*, *42*(4), 651–685. https://doi.org/10.1177/0098858817700245

Farrington, C. (2018). Data as Transformational: Constrained and Liberated Bodies in an 'Artificial Pancreas' Study. In R. Lynch & C. Farrington (Eds.), *Quantified Lives and Vital Data: Exploring Health and Technology through Personal Medical Devices* (pp. 127–154). Palgrave Macmillan. https://doi.org/10.1007/978-1-137-09593-0

Fox, N. J. (2017). Personal health technologies, micropolitics and resistance: A new materialist analysis. *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*, *21*(2), 136–153. https://doi.org/10.1177/1363459315590248

Gabriels, K., & Moerenhout, T. (2018). Exploring Entertainment Medicine and Professionalization of Self-Care: Interview Study Among Doctors on the Potential Effects of Digital Self-Tracking. *Journal of Medical Internet Research*, *20*(1), e10. https://doi.org/10.2196/jmir.8040

Kent, R. (2018). Social Media and Self-Tracking: Representing the 'Health Self'. In B. Ajana (Ed.), *Self-Tracking: Empirical and Philosophical Investigations* (pp. 61–76). Palgrave Macmillan.

Kim, M. (2019, December 16). *DIY diabetes tech gains popularity with patients and parents fed up with clunky mainstream medical devices*. Washington Post. https://www.washingtonpost.com/health/fed-up-with-clunky-diabetes-machines-do-it-yourselfers-re-engineer-devices-and-create-apps-and-software/2019/12/13/3f7c4e20-16c4-11ea-9110-3b34ce1d92b1_story.html

Klugman, C. M., Dunn, L. B., Schwartz, J., & Cohen, I. G. (2018). The Ethics of Smart Pills and Self-Acting Devices: Autonomy, Truth-Telling, and Trust at the Dawn of Digital Medicine. *The American Journal of Bioethics*, *18*(9), 38–47. https://doi.org/10.1080/15265161.2018.1498933

Lupton, D. (2012). M-health and health promotion: The digital cyborg and surveillance society. *Social Theory & Health*, *10*(3), 229–244. https://doi.org/10.1057/sth.2012.6

Lupton, D. (2013a). The digitally engaged patient: Self-monitoring and self-care in the digital health era. *Social Theory & Health*, *11*(3), 256–270. https://doi.org/10.1057/sth.2013.10

Lupton, D. (2013b). Quantifying the body: Monitoring and measuring health in the age of mHealth technologies. *Critical Public Health*, *23*(4), 393–403. https://doi.org/10.1080/09581596.2013.794931

Lupton, D. (2016). The diverse domains of quantified selves: Self-tracking modes and dataveillance. *Economy and Society*, *45*(1), 101–122. https://doi.org/10.1080/03085147.2016.1143726

O'Connor, A. (2019, December 2). In Weekend Outage, Diabetes Monitors Fail to Send Crucial Alerts. *The New York Times*. https://www.nytimes.com/2019/12/02/well/live/Dexcom-G6-diabetes-monitor-outage.html

Piras, E. M., & Miele, F. (2017). Clinical self-tracking and monitoring technologies: Negotiations in the ICT-mediated patient–provider relationship. *Health Sociology Review*, *26*(1), 38–53. https://doi.org/10.1080/14461242.2016.1212316

Quintal, A., Messier, V., Rabasa-Lhoret, R., & Racine, E. (2019). A critical review and analysis of ethical issues associated with the artificial pancreas. *Diabetes & Metabolism*, *45*(1), 1–10. https://doi.org/10.1016/j.diabet.2018.04.003

Rich, E., & Miah, A. (2017). Mobile, wearable and ingestible health technologies: Towards a critical research agenda. *Health Sociology Review*, *26*(1), 84–97. https://doi.org/10.1080/14461242.2016.1211486

Rowe, E. A. (2018). Sharing Data. *Iowa Law Review*, *104*(1), 287–324.

Scelsi, C. (2015). Care and Feeding of Privacy Policies and Keeping the Big Data Monster at Bay: Legal Concerns in the Age of the Internet of Things. *Nova Law Review*, *39*(3), 391–436.

Sharon, T. (2016). The Googlization of health research: From disruptive innovation to disruptive ethics. *Personalised Medicine*, *13*(6), 563–574. https://doi.org/10.2217/pme-2016-0057

Sharon, T. (2017). Self-Tracking for Health and the Quantified Self: Re-Articulating Autonomy, Solidarity, and Authenticity in an Age of Personalised Healthcare. *Philosophy & Technology*, *30*(1), 93–121. https://doi.org/10.1007/s13347-016-0215-5

Smith, G. J. D., & Vonthethoff, B. (2017). Health by numbers? Exploring the practice and experience of datafied health. *Health Sociology Review*, *26*(1), 6–21. https://doi.org/10.1080/14461242.2016.1196600

Terry, N. P. (2018). Appification, AI, and Healthcare's New Iron Triangle. *Journal of Health Care Law and Policy*, *20*(2), 117–182.

Zapatka, C. (2019, November 19). *The chaotic industry behind the insulin I need to live*. The Verge. https://www.theverge.com/2019/11/19/20966695/insulin-industry-diabetic-type-1-drug-price-cost-manufacturing-access

Zimberoff, L. (2015, May 27). Glucose-Sensing Contacts and More Brilliant Diabetes Tech. *Wired*. https://www.wired.com/2015/05/diabetes-monitoring-tech/

Zimberoff, L. (2017, October 17). A Diabetes Monitor That Spares the Fingers. *The New York Times*. https://www.nytimes.com/2017/10/17/well/live/a-diabetes-monitor-that-spares-the-fingers.html

12 Case study 11: Self-Tracking for Fitness

12.1 Objective

This case study will present and address selected ethical and legal aspects of the use of self-tracking devices in fitness-oriented activities. While the focus will be on the ethical and legal dimension of personalisation connected with the tracking of physical activity, other factors will also be outlined.

12.2 Technology description

Fitness-tracking is a very popular practice facilitated by the introduction of specialised wearable devices such as fitness bands and smartwaches, as well as the inclusion of dedicated sensors in smartphones. People engaging in fitness-tracking use their devices to collect data relating to their physical activity, such as their daily step count, heart rate, distance travelled and others. The tracking can either be dependent on user input of relevant data, or on a passive collection of relevant metrics by the device. In practice, it is often a combination of both, with the user selecting categories, goals and context for tracking, and the device registering datapoints. Popular fitness apps and devices, for example Fitbit, Strava, Nike Fuel, Apple Watch and Apple Health, allow users to keep detailed record of their past and present fitness levels, and to share and compare their data with others, most often in numerical or visual forms (graphs, activity maps etc.). Most importantly for this case study, fitness-tracking involves also personalised lifestyle, diet and exercise recommendations which range from minuscule nudges encouraging users to walk more to fully-fledged training regimes constructed on the basis of personal data. Typically, a user might see recommendations on when to exercise, what activities to include in their training, how to improve their abilities or endurance, and what food or drinks to consume in order to make the most of their regimen. Fitness self-tracking can be said to be goal-oriented (e.g. towards weight loss) and dependent on actionable insights from the apps and devices. Consequently, the personalised recommendations issued by the devices are at the centre of tracking physical activity – even if some users use their fitness-tracking apps and devices purely for statistical purposes, the algorithms at work provide relevant insights, graphs and predictions automatically as part of their functioning.

12.3 Context

Fitness-tracking is a predominantly individual endeavour, so most ethical and legal issues that can be identified impact the primary users the most. However, the communal dimension of fitness-tracking is evident in ways the users are encouraged to share their data and compare themselves with others – consequently, some concerns arise with regards to fitness-tracking’s impact on privacy, interpersonal relations and the idea of a community promoted by these technologies. Some institutional actors, such as insurance companies and public bodies have also raised various levels of interest with the data collected by users of fitness apps. Last but not least, the way the companies behind these technologies are able to use and access data is also relevant.

12.3.1 Potential use

The development of tracking sensors promises greater accuracy, limited costs and improved ease of use of fitness-tracking sensors. Moreover, as algorithms used for fitness-tracking purposes grow in complexity, the recommendations and predictions are expected to become more relevant and useful to all kinds of users. Combined, these factors could bring great improvements in public health and facilitate the management of habits for diverse users. Moreover, data collected through fitness-tracking could improve our knowledge connected to the human body and physical exercise.

12.4 Relationship to Workpackage

The data collected by fitness-tracking devices can be used to provide personalised recommendations relating to individuals' lifestyle, diet and exercise regimens. Moreover, the data collected through tracking can be used to create a personalised profile of the user that may later serve as a basis for personalised advertisements or other persuasive activities.

12.5 Relationship to ESR PhD topic

The subject of my PhD research are the ethical issues surrounding self-tracking practices and technologies. Fitness-tracking is one of the best example of such practices and technologies.

12.6 Review of ethical and legal issues

12.6.1 Methodology

The foundation for this case study is a literature search conducted in Google Scholar, Scopus and Hein Online as well as online media sources. In order to find the discussion of the most relevant ethical and legal aspects of self-tracking for fitness, I used search terms from two sets that were combined into search strings. The first set included the terms 'fitness track*', 'activity track*' 'self-track* AND fitness', and the second included the terms 'ethic*', 'moral*', 'virtue*' and 'legal*'. All have been adjusted to fit the engine used by specific databases. Additional articles were identified through backwards snowballing done on the search results. Some authors researched self-tracking technologies in general, but their findings are discussed here in the context of fitness tracking.

12.6.2 Findings

Although many fitness apps and devices are supposed to provide personalised recommendations and services, some researchers note that this is not always the case. These recommendations and services are based on pre-established norms that should be attained by the users (Reijneveld, 2017) and these **norms often do not reflect the diversity of the users** and endorse the body image and standards that are most suited to young, white males (Lupton, 2018; Ruckenstein & Pantzar, 2015). However, the suggestions and analyses constructed on the basis of the employed models are often presented as "standard", "typical" or "recommended", which could give the discussed apps and devices great **norm-prescribing effects**, especially if users believe the promise of personalisation and expect to see information that is tailored to them, not based on an abstract model. Consequently, the underrepresentation of bodily norms specific to minority and vulnerable groups can lead to a **distorted view of the "standard body"**. This could further marginalise those who are already discriminated against in the sphere of physical activity. For example, an overweight person might find it extremely difficult to reach the levels of activity endorsed by the developers, which in turn might lead them to feeling alienated and giving up on physical activity in the end. Additionally, data produced through fitness tracking has great **impact on how trackers perceive their bodies and environment** (Kristensen & Ruckenstein, 2018; Lupton et al., 2018). The persuasive force of the recommendations issued by fitness apps and devices could also convince some users to engage in far-reaching **practices of self-regulation** or to excessively depend on these devices in the management of their daily life (Schüll, 2016), which could be seen as a **threat to their autonomy and health** when taken to the extreme.

Some have criticised fitness tracking for promoting **a narcissistic worldview** and reducing empathy by showing only the perspective of relatively wealthy, fit people. Others have noted that fitness tracking **prompts users to engage only in activity that is labelled as productive** and thus much more desirable than leisure (Till, 2014), which consequently promotes the narrative of the most valuable personal contributions in life being connected to the increasing of individual productivity and one's capacities as a labouring subject (Fotopoulou & O'Riordan, 2017). Gabriels and Coeckelbergh (2019) noted that the networked character of fitness tracking encourages people to monitor the activity of others by looking at data shared by self-tracking apps and devices on social media and other channels, sometimes even without the **data subjects' knowledge or consent**. This could lead to concerns about **voyeurism and privacy violations** occurring whenever others are granted access to the data users would prefer to remain undisclosed.

Moreover, fitness-tracking technologies have been demonstrated to be significantly **biased against women**. Since the male body is considered the standard by developers of the apps and devices, women are often issued less useful recommendations and the data collection is less reliable, while at the same time, they are expected to put more effort into conforming to the norms (Lupton, 2018; Sanders, 2017). The sensors embedded in smartphones often work well only if the devices are stored in pockets, while the sizes of many wearable fitness-tracking devices make it more difficult for women to wear them (Criado Perez, 2020, pp. 159–160). Fitness band readings are also reported to be less accurate in connection with some activities that are traditionally more likely to be done by women[33], such as pushing

[33] Of course, men also push prams, but parenting responsibilities are still more commonly assigned to women and the inaccuracy of fitness tracker in this instance affects them disproportionately more. Mowing the lawn could be a good example of an activity traditionally reserved for men that would not be accurately recorded by a fitness tracker, but it has to be noted that even the most dedicated gardener would not spend as much time pushing the lawnmower as mothers typically spend pushing prams.

prams. Some women users have reported that a significant portion of their daily activity has not been recorded by their devices (Lupton & Maslen, 2018).

Privacy issues are also a great problem in the case of fitness tracking (Lanzing, 2016, 2019; Reijneveld, 2017). The data on users' physical activity can be used to study their behaviour patterns, real-time location and health. There is a great risk involved with this **data getting accessed by third parties**, for example, insurance companies who could use it to discriminate against their customers by increasing insurance prices for people not following a strict training regimen. Similarly, online advertisers are able to target this data to offer highly persuasive ads to vulnerable individuals. A big challenge to protecting privacy of fitness-trackers is that even regulatory bodies are unable to say where the data ultimately ends up and how it is used (Christovich, 2016). Some scholars have also pointed out that the abundance of self-tracked data is highly useful for governments that are already engaging in massive **surveillance** of the population (Lupton, 2016; Sanders, 2017). However, while there exist concerns connected to third parties' potential access to self-tracked fitness data (Karanasiou & Kang, 2016), these concerns should be addressed in a way that does not infringe on the benefits self-trackers gain through their practices (Schüll, 2019). Moreover, the discussed tracking is not always consensual – most contemporary smartphones have relevant sensors installed in them and there are often **no ways to opt out** of step counting or location sharing. While these can be used to provide fitness feedback, they can be also used by bad-faith actors to **monitor individuals of interest**, threatening in some cases not only privacy but even national security (Thompson & Warzel, 2019). Additionally, in the context of neoliberal governance, it is the **individuals that are often burdened with the responsibility** to safely manage their collected personal data (Smith, 2016).

Some concerns have also been raised in regards to the **commodification of daily activity** that is connected to fitness tracking (Dewart McEwen, 2018; Fox, 2017; Ruckenstein & Pantzar, 2015; Till, 2018). Since data collected through fitness bands and smartphones can be sold to third parties (Liebelson, 2014), fitness tracking can be seen as creating commodities on the basis of leisure activities, which were previously not part of profit-making (other than driving consumption). However, the users are not compensated for the value they help create by monitoring their activity which can be considered a form of **exploitation or unjust distribution of profit**.

Finally, it has to be noted that the use of self-collected data regarding physical activity has become a challenge in the legal field. For example, in a recent court case in Canada, a woman used Fitbit data to demonstrate that her level of activity has decreased after her injury for which her insurance company did not want to pay out compensation (Gibbs, 2014), while another woman was charged with a misdemeanour when her Fitbit contradicted her testimony (Moon, 2015). Although these are individual instances, they raise questions about the **possibility of self-tracking data being used in legal proceedings** as well as criminal investigations. It has to be noted, however, that the **disputable accuracy** of self-tracked data, as well as the **inherent biases** promoted by such devices, make it difficult to assess whether such data should indeed be considered admissible evidence in legal proceedings (Rutkin, 2015).

However, some authors note that self-tracking for fitness can bring real benefits to people practicing it (Fotopoulou & O'Riordan, 2017; Kristensen & Ruckenstein, 2018; Lanzing, 2016; Schüll, 2016, 2019). Fitness-tracking can be understood as a practice of self-care which improves well-being, facilitates behaviour change and increases knowledge about the body and public health. Moreover, it can also bring about a sense of community and improve interpersonal relations as users are able to compare their achievements with others and in some cases understand the needs of others better.

12.7 References

Christovich, M. M. (2016). Why Should We Care What Fitbit Shares?: A Proposed Statutory Solution to Protect Sensitive Personal Fitness Information. *Hastings Communication and Entertainment Law Journal*, *38*(1), 91–116.

Criado Perez, C. (2020). *Invisible Women*. Vintage.

Dewart McEwen, K. (2018). Self-Tracking Practices and Digital (Re)productive Labour. *Philosophy & Technology*, *31*(2), 235–251. https://doi.org/10.1007/s13347-017-0282-2

Fotopoulou, A., & O'Riordan, K. (2017). Training to self-care: Fitness tracking, biopedagogy and the healthy consumer. *Health Sociology Review*, *26*(1), 54–68. https://doi.org/10.1080/14461242.2016.1184582

Fox, N. J. (2017). Personal health technologies, micropolitics and resistance: A new materialist analysis. *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*, *21*(2), 136–153. https://doi.org/10.1177/1363459315590248

Gabriels, K., & Coeckelbergh, M. (2019). 'Technologies of the self and other': How self-tracking technologies also shape the other. *Journal of Inormation, Communication and Ethics in Society*, *17*(2), 119–127.

Gibbs, S. (2014, November 18). Court sets legal precedent with evidence from Fitbit health tracker. *The Guardian*. http://www.theguardian.com/technology/2014/nov/18/court-accepts-data-fitbit-health-tracker

Karanasiou, A. P., & Kang, S. (2016). My Quantified Self, my FitBit and I. *Digital Culture & Society*, *2*(1). https://doi.org/10.14361/dcs-2016-0109

Kristensen, D. B., & Ruckenstein, M. (2018). Co-evolving with self-tracking technologies. *New Media & Society*, *20*(10), 3624–3640.

Lanzing, M. (2016). The transparent self. *Ethics and Information Technology*, *18*(1), 9–16. https://doi.org/10.1007/s10676-016-9396-y

Lanzing, M. (2019). "Strongly Recommended" Revisiting Decisional Privacy to Judge Hypernudging in Self-Tracking Technologies. *Philosophy & Technology*, *32*(3), 549–568. https://doi.org/10.1007/s13347-018-0316-4

Liebelson, D. (2014, January 31). Are Fitbit, Nike, and Garmin planning to sell your personal fitness data? *Mother Jones*. https://www.motherjones.com/politics/2014/01/are-fitbit-nike-and-garmin-selling-your-personal-fitness-data/

Lupton, D. (2016). The diverse domains of quantified selves: Self-tracking modes and dataveillance. *Economy and Society*, *45*(1), 101–122. https://doi.org/10.1080/03085147.2016.1143726

Lupton, D. (2018). 'I Just Want It to Be Done, Done, Done!' Food Tracking Apps, Affects, and Agential Capacities. *Multimodal Technologies and Interaction*, *2*(2), 29. https://doi.org/10.3390/mti2020029

Lupton, D., & Maslen, S. (2018). The more-than-human sensorium: Sensory engagements with digital self-tracking technologies. *The Senses and Society*, *13*(2), 190–202. https://doi.org/10.1080/17458927.2018.1480177

Lupton, D., Pink, S., Labond, C. H., & Sumartojo, S. (2018). Personal Data Contexts, Data Sense, and Self-Tracking Cycling. *International Journal of Communication*, *12*, 647–665.

Moon, M. (2015, June 28). Fitbit tracking data comes up in another court case. *Engadget*. https://www.engadget.com/2015-06-28-fitbit-data-used-by-police.html

Reijneveld, M. D. (2017). Quantified Self, Freedom, and the GDPR. *SCRIPT-Ed*, *14*(2), 285–325. https://doi.org/10.2966/scrip.140217.285

Ruckenstein, M., & Pantzar, M. (2015). Datafied Life: Techno-Anthropology as a Site for Exploration and Experimentation. *Techné: Research in Philosophy and Technology*, *19*(2), 191–210. https://doi.org/10.5840/techne20159935

Rutkin, A. (2015). It's a Fitbit, Your Honour. *New Scientist*, *225*(3002), 17.

Sanders, R. (2017). Self-tracking in the Digital Era: Biopower, Patriarchy, and the New Biometric Body Projects. *Body & Society*, *23*(1), 36–63.

Schüll, N. D. (2016). Data for life: Wearable technology and the design of self-care. *BioSocieties*, *11*(3), 317–333. https://doi.org/10.1057/biosoc.2015.47

Schüll, N. D. (2019). The Data-Based Self: Self-Quantification and the Data-Driven (Good) Life. *Social Research*, *86*(4), 909–930.

Smith, G. J. D. (2016). Surveillance, Data and Embodiment: On the Work of Being Watched. *Body & Society*, *22*(2), 108–139. https://doi.org/10.1177/1357034X15623622

Thompson, S. A., & Warzel, C. (2019, December 19). Twelve Million Phones, One Dataset, Zero Privacy. *The New York Times*. https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html

Till, C. (2014). Exercise as Labour: Quantified Self and the Transformation of Exercise into Labour. *Societies*, *4*(3), 446–462. https://doi.org/10.3390/soc4030446

Till, C. (2018). Self-Tracking as the Mobilisation of the Social for Capital Accumulation. In B. Ajana (Ed.), *Self-Tracking: Empirical and Philosophical Investigations* (pp. 77–91). Palgrave Macmillan.

13 Case study 12: Fertility Tracking

13.1 Objective

This case study is devoted to ethical and legal aspects of the tracking of fertility with the help of digital technologies. The main focus of this study are the concerns and opportunities connected to personalisation embedded in popular fertility tracking applications.

13.2 Technology description

Fertility tracking is the practice of using software and (to a lesser extent) hardware to collect personal data related to reproduction and sexual activity for the purposes of facilitating or preventing pregnancy and, in case of women, monitoring the menstrual cycle. Fertility tracking is largely dependent on manual input of data points by users, but some devices offer passive collection of some relevant information (e.g. smart thermometers checking basal body temperature[34]). Through the collection of metrics concerning basal body temperature, frequency of sexual activity, type of menstrual flow or cervical mucus, mood patterns and others, fertility tracking apps provide personalised lifestyle recommendations (aimed at either increasing or decreasing the likelihood of getting pregnant) as well as personalised predictions regarding the timetable of the menstrual cycle and likely symptoms on a given day in the cycle. Typical information provided by these apps could include suggestions when the chance of conceiving is the highest, what volume of blood to expect during menstruation, what diet or activity patterns to follow to avoid symptoms like bloating or cramps, and even relationship advice aimed at helping the couples make the most of their attempts to conceive (e.g. how to convince the partner to have sex on a given day or how to leave out the stresses of everyday life out of the bedroom).

While most apps and devices used in fertility tracking are aimed at women, men are also sometimes included in the practice – either by receiving access to the readings from their partners' apps or by contributing some metrics themselves (e.g. testicular temperature or frequency of intercourse). Personalised recommendations and predictions aimed at men concern either suggestions on how they should adapt their sexual behaviour to best fit their partners' cycle and goals (e.g. facilitating or preventing pregnancy) or recommendations on how to maintain a healthy sperm count and sperm motility (including recommendations connected to lifestyle, diet and even best type of underwear). Popular fertility tracking apps include those developed by Glow, Clue, Ovia, Kindara and Flo.

13.3 Context

Fertility tracking is a practice aimed at either individual, predominantly female users, or couples willing to work together in order to facilitate or prevent pregnancy. However, gender roles, tabooisation of sex and social stigma attached to menstruation all result in the fact that it is impossible to neatly place fertility tracking on the individual-communal divide. Parties interested in a particular person's reproductive health and desires might include not

[34] See https://iproven.com/products/bbt-113 for one of many examples of such a product.

only that person and their partner, but also friends, relatives, employers, insurance companies, state institutions, researchers and even perfect strangers, as evidenced by the desire of many political groups to regulate women's reproductive rights. Consequently, while the tracking of fertility might be something undertaken by individual users or couples, sometimes in cooperation with a gynaecologist, the ethical and legal concerns surrounding the practice impact a significant number of varied stakeholders.

13.3.1 Potential use

In the foreseeable future, fertility tracking apps can grow in scope and accuracy to provide trustworthy predictions and recommendations for their users, possibly increasing access to sexual health and facilitating the planning or avoiding of pregnancy for a diverse group of users. Additionally, considering that fertility is an under-researched phenomenon, the collection of reproductive data could have great benefits for the improvement of our knowledge relating to fertility and sexuality in general.

13.4 Relationship to Workpackage

As noted above, personalised recommendations and predictions are a key element of fertility tracking. Moreover, personal data sourced from the users of fertility tracking apps is often used by advertisers, insurance companies and other actors in order to offer these users personalised advertisements, insurance plans and other services.

13.5 Relationship to ESR PhD topic

The subject of my PhD research are the ethical issues surrounding self-tracking practices and technologies. While my thesis focuses predominantly on quantified and passive collection of personal information, fertility tracking deals to a large extent with non-quantifiable, self-registered data (e.g. the assessment of the type of cervical mucus on a given day). It is, however, a paradigmatic example of self-tracking.

13.6 Review of ethical and legal issues

13.6.1 Methodology

The foundation for this case study is a literature search conducted in Google Scholar, Scopus and Hein Online as well as Google Search (in order to cover online media sources). In order to find the discussion of the most relevant ethical and legal aspects of fertility tracking, I used search terms from two sets that were combined into search strings. The first set included the terms 'fertility track*', 'menstrua* track*', 'cycle track*' 'period track*' and 'reproducti* track', and the second included the terms 'ethic*', 'moral*', 'virtue*' and 'legal*'. All have been adjusted to fit the engine used by specific databases. Some additional articles were identified through snowballing done on the search results.

13.6.2 Findings

Interestingly, relatively little research has been done on ethical and legal implications of fertility tracking, especially when compared with the kinds of self-tracking described in my other two case studies. This can be perhaps attributed to the fact that fertility, and especially menstruation, remains an underresearched phenomenon in general. Moreover, a significant number of concerns relating to fertility tracking has been identified by social scientists and journalists. This is undoubtedly a field that would merit more scrutiny from ethicists and philosophers, especially as it is relatively recent, diverse and rapidly evolving. Many concerns raised by authors discussed below have been the subject of scrutiny by developers, activists and policymakers, while many others do not apply to the same extent to all apps offering fertility-tracking. Consequently, some of the issues mentioned below, have been already addressed or are currently being addressed by at least some of the developers.

Although fertility tracking apps are often explicitly marketed as offering personalised services, it can be noted there is often surprisingly **little personalisation involved in the collection of data and the creation of recommendations and predictions** (Delano, 2015; Epstein et al., 2017; Eveleth, 2014; Fetters, 2018; Hall, 2017; Karlsson, 2019; Kressbach, 2019; Kroløkke, 2020; J. Levy & Romo-Avilés, 2019; Lupton, 2015). For instance, non-binary and queer users have reported that it is often impossible to input data and get predictions that are not designed with straight women in mind. Delano (2015) noted that her fertility tracker repeatedly prompted her to use contraception during sex, even though her partner, a woman, could not have gotten her pregnant. There was, however, no way to indicate that in the application, similarly to how it was impossible to report an irregular or very short cycle length.

This was something observed also by other authors, with Levy and Romo-Avilés (2019) noting that this could **stigmatise women experiencing periods labelled as 'pathological'** (even if their cycles are perfectly normal and healthy) and lead them to blame their bodies and lifestyles for perceived shortcomings. As described by Karlsson (2019), fertility tracking apps, together with other mechanism of societal pressure connected with fertility, often **shame women that do not want to or are unable to fit the image of a standard woman**. The promise of personalisation offered by these applications might make their **norm-prescribing effects** even stronger, since recommendations and predictions can be framed and perceived as aimed at a particular user and created on the basis of their personal data, not in reference to a predetermined and inadequately justified norm.

Fertility tracking is also widely seen as **reproducing gender roles and stereotypes**, with interfaces of their popular examples communicating information through sexist tropes (e.g. with the app giving its sexual advice while pretending to be a gossiping girlfriend) and using stereotypically feminine colours and themes in their presentation (Epstein et al., 2017; Eveleth, 2014; Hall, 2017; Hendl et al., 2019; Kressbach, 2019). More importantly, however, the content of the recommendations and predictions has also been criticised for **gender insensitivity and sexism**: by default, the apps treat women as seeking to get pregnant (Epstein et al., 2017), whereas men are met with the image of a sexually potent macho, who often manages multiple sexual partners and equates good sexual performance with physical fitness and effort (Kroløkke, 2020).

Additionally, even if **men are** sometimes able to take part in their partner's fertility tracking, they are often **largely excluded from the planning of pregnancy**, seeing their role reduced to merely supplying the sperm (Fetters, 2018). Developers of fertility tracking apps do not allow the users to personalise the content and presentation of the offered recommendations and predictions, and, consequently, they **do not reflect the diversity of users' bodies, gender roles, sexual orientations and tracking goals** (Costa Figueiredo et al., 2018; Epstein et al., 2017).

The information supplied by fertility tracking apps is also problematic. **The models and methods** which serve as the basis for personalised recommendations and predictions **are often outdated or scientifically suspect** and their accuracy is questionable, even if they are made on the basis of a large amount of data (Danaher et al., 2018; Epstein et al., 2017; Karlsson, 2019; Lee, 2019; Starling et al., 2018; Wiegel, 2016). Many fertility tracking apps base their output on variations on the fertility awareness method (FAM), which has been criticised for its low effectiveness and unreliability as well as lack of an uniform standard for how it should be practiced and developed. Moreover, as this method has been practiced in some forms ever since antiquity (and was updated in various ways and to various extents over the course of the 20th century), it is questionable whether FAM-based fertility tracking can offer its users any new insights and whether its recommendations substantially differ from others formulated on the basis of FAM, such as, for example, reproductive advice offered by the Catholic Church or advice passed on by previous generations of women (Wiegel, 2016).

This is particularly concerning as relatively **minor mistakes in predictions** (e.g. menstruation starting one day earlier than predicted by the app) **can have significant impact on the lives of the users**, especially when fertility tracking is used as a contraceptive method. For example, ovulation predicted to occur even a few days later than in reality could lead to an unwanted pregnancy, while menstruation starting earlier than predicted could leave a user without needed sanitary products, especially if they had great confidence in the app's output.

Moreover, the apps do not disclose how they arrived at specific recommendations (Novotny & Hutchinson, 2019) and this **lack of transparency**, as well as the overwhelming amount of information supplied by them can further **exacerbate negative emotions**, such as stress and anxiety, that are already often associated with fertility and menstruation (Costa Figueiredo et al., 2017, 2018). Finally, by presenting information as personalised predictions made on the basis of menstruation-related data, these apps can falsely suggest that some unrelated symptoms, like aches or mood swings, are connected to or caused by menstruation (Kressbach, 2019).

Some issues regarded to **access to data, privacy and consent** have also been noted (Danaher et al., 2018; Hall, 2017; Karlsson, 2019; Kresge et al., 2019; Kressbach, 2019; J. Levy & Romo-Avilés, 2019; K. E. C. Levy, 2015; Novotny & Hutchinson, 2019; Privacy International, 2019). Many of the **popular fertility tracking apps share users' data by**

default to various actors. These can include partners, advertisers, medical professionals and pharmacists, researchers, insurance companies and even employers, making it difficult to determine who exactly benefits from the sharing of fertility-related data. As noted by Levy (2015), fertility-related data is always sensitive, so it is doubtful whether anonymisation and aggregation are enough to protect the users from possible violations of privacy. Many of the apps are based in countries that have worse privacy protection regimes than the European Union, while some have been demonstrated to be non-compliant with GDPR and **shared user data with advertisers and Facebook without users' knowledge** (Privacy International, 2019). This is, of course, constantly changing, as even some developers called out by Privacy International quickly changed their data management practices. However, GDPR compliance is not something that can be reasonably evaluated and enforced by individual users, so **poor data management** is a serious threat to the privacy of those engaging in fertility tracking. Even where consent is obtained, privacy policies and data protection standards are often opaque, to the point where it is unlikely that users are able to give ***informed*** **consent** (Novotny & Hutchinson, 2019). This is particularly relevant, as pregnant women's data is extremely valuable to advertisers since after getting pregnant women's shopping habits drastically change and long-lasting consumer habits are formed – having children requires the parents to purchase large quantities of new kinds of products, like diapers, over an extended period of time (Kresge et al., 2019; Petronzio, 2014). The use of pregnancy-related data from fertility tracking apps in advertising could lead to **consumer manipulation**, but also give companies **unfair edge over their competitors** that do not have access or decide not to use this kind of information about their potential clients.

There are also risks associated with fertility-related data being **used by companies to discriminate against their employees** – some fertility tracking apps share anonymised and aggregated data with corporate clients in order to help them with planning for their employees' pregnancies (Hall, 2017). One tech company, Activision Blizzard, notably paid their female employees $1 per day in gift cards if they shared fertility data voluntarily (Harwell, 2019). This raises concerns regarding possible violations as well as the **power differences** between female employees and their employers.

As noted by some scholars, the sharing of data with partners is also problematic (Danaher et al., 2018; K. E. C. Levy, 2015; Privacy International, 2019), with many apps making it either too easy to gain access to intimate information or too difficult to remove another user's access privileges. This raises concerns about possible **breaches of trust among partners** (Danaher et al., 2018), but also about the possibility of **intimate surveillance** undertaken by abusive partners and stalkers (K. E. C. Levy, 2015).

Finally, the way fertility apps **frame fertility as an individual problem**, to be solved by personalised recommendations and predictions, is also an issue (Eveleth, 2014; Hall, 2017; Karlsson, 2019; Lupton, 2015). This creates a risk of even greater burden connected to fertility being placed on women and further shifting the responsibility for reproductive health from the state to individuals. For example, the increased use of fertility tracking apps might be seen as a substitute for previously employed systemic solutions aimed at improving reproductive health, like funding IVF treatments. However, despite the

individualisation of responsibility evident in the paradigm promoted by these apps, women's bodies are still subjected to surveillance and normative expectations connected to reproduction.

While most articles focused on identifying issues associated with digital fertility tracking, some have also discussed the practice's advantages (Danaher et al., 2018; Karlsson, 2019; Kressbach, 2019; Wiegel, 2016). Some proponents of fertility tracking note that it can have positive impact on reproductive health and knowledge about human reproduction, while making the planning or avoiding of pregnancy easier (Kressbach, 2019; Wiegel, 2016). Others noted that it could help normalise menstruation and reduce stigma associated with it (Karlsson, 2019). Danaher et al. focus on advantages that tracking can bring to relationships by increasing trust or facilitating non-standard relations of reciprocity, while providing partners more information about each other (2018).

13.7 References

Costa Figueiredo, M., Caldeira, C., Eikey, E. V., Mazmanian, M., & Chen, Y. (2018). Engaging with Health Data: The Interplay Between Self-Tracking Activities and Emotions in Fertility Struggles. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), 1–20. https://doi.org/10.1145/3274309

Costa Figueiredo, M., Caldeira, C., Reynolds, T. L., Victory, S., Zheng, K., & Chen, Y. (2017). Self-Tracking for Fertility Care: Collaborative Support for a Highly Personalised Problem. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 1–21. https://doi.org/10.1145/3134671

Danaher, J., Nyholm, S., & Earp, B. D. (2018). The Quantified Relationship. *The American Journal of Bioethics*, *18*(2), 3–19. https://doi.org/10.1080/15265161.2017.1409823

Delano, M. (2015, December 8). *I tried tracking my period and it was even worse than I could have imagined*. Medium. https://medium.com/@maggied/i-tried-tracking-my-period-and-it-was-even-worse-than-i-could-have-imagined-bb46f869f45

Epstein, D. A., Lee, N. B., Kang, J. H., Agapie, E., Schroeder, J., Pina, L. R., Fogarty, J., Kientz, J. A., & Munson, S. (2017). Examining Menstrual Tracking to Inform the Design of Personal Informatics Tools. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6876–6888. https://doi.org/10.1145/3025453.3025635

Eveleth, R. (2014, December 15). How Self-Tracking Apps Exclude Women. *The Atlantic*. https://www.theatlantic.com/technology/archive/2014/12/how-self-tracking-apps-exclude-women/383673/

Fetters, A. (2018, August 16). How Fertility Apps Exclude Fathers. *The Atlantic*. https://www.theatlantic.com/family/archive/2018/08/how-fertility-apps-exclude-fathers/567721/

Hall, M. (2017, July 25). *The Strange Sexism of Period Apps*. Vice. https://www.vice.com/en_us/article/qvp5yd/the-strange-sexism-of-period-apps

Harwell, D. (2019, April 10). Is your pregnancy app sharing your intimate data with your boss? *Washington Post*. https://www.washingtonpost.com/technology/2019/04/10/tracking-your-pregnancy-an-app-may-be-more-public-than-you-think/?utm_term=.09df9947f49c

Hendl, T., Jansky, B., & Wild, V. (2019). From Design to Data Handling. Why mHealth Needs a Feminist Perspective. In J. Loh & M. Coeckelbergh (Eds.), *Feminist Philosophy of Technology* (Vol. 2, pp. 77–104). J.B. Metzler. https://doi.org/10.1007/978-3-476-04967-4

Karlsson, A. (2019). A Room of One's Own? *Nordicom Review*, *40*(Special Issue 1), 111–123. https://doi.org/10.2478/nor-2019-0017

Kresge, N., Khrennikov, I., & Ramli, D. (2019, January 24). Period-Tracking Apps Are Monetizing Women's Extremely Personal Data. *Bloomberg.Com*. https://www.bloomberg.com/news/articles/2019-01-24/how-period-tracking-apps-are-monetizing-women-s-extremely-personal-data

Kressbach, M. (2019). Period Hacks: Menstruating in the Big Data Paradigm. *Television & New Media*, *00*(0), 1–21.

Kroløkke, C. (2020). Big sperm. The making of the (new) male repro-consumer. *NORMA*, 1–17. https://doi.org/10.1080/18902138.2020.1720335

Lee, N. (2019, July 24). *Are period and fertility tracking apps effective?* Engadget. https://www.engadget.com/2019-07-24-are-period-and-fertility-tracking-apps-effective.html

Levy, J., & Romo-Avilés, N. (2019). "A good little tool to get to know yourself a bit better": A qualitative study on users' experiences of app-supported menstrual tracking in Europe. *BMC Public Health*, *19*(1). https://doi.org/10.1186/s12889-019-7549-8

Levy, K. E. C. (2015). Intimate Surveillance. *Idaho Law Review*, *51*, 679–693.

Lupton, D. (2015). Quantified sex: A critical analysis of sexual and reproductive self-tracking using apps. *Culture, Health & Sexuality*, *17*(4), 440–453. https://doi.org/10.1080/13691058.2014.920528

Novotny, M., & Hutchinson, L. (2019). Data Our Bodies Tell: Towards Critical Feminist Action in Fertility and Period Tracking Applications. *Technical Communication Quarterly*, *28*(4), 332–360. https://doi.org/10.1080/10572252.2019.1607907

Petronzio, M. (2014, April 26). *How One Woman Hid Her Pregnancy From Big Data*. Mashable. https://mashable.com/2014/04/26/big-data-pregnancy/

Privacy International. (2019, September 9). *No Body's Business But Mine: How Menstruation Apps Are Sharing Your Data*. http://privacyinternational.org/long-read/3196/no-bodys-business-mine-how-menstruations-apps-are-sharing-your-data

Starling, M. S., Kandel, Z., Haile, L., & Simmons, R. G. (2018). User profile and preferences in fertility apps for preventing pregnancy: An exploratory pilot study. *MHealth*, *4*(6). https://doi.org/10.21037/mhealth.2018.06.02

Wiegel, M. (2016, March 23). *'Fitbit for your period': The rise of fertility tracking*. The Guardian. http://www.theguardian.com/technology/2016/mar/23/fitbit-for-your-period-the-rise-of-fertility-tracking

End of Deliverable D2.1.