Data Surveillance and Digital Humanism

An Educational Perspective Through the Medium of Documentary Filmmaking

Master Thesis

Arturo Fabbro

Aalborg University
Electronics and IT

Electronics and IT
Aalborg University
http://www.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Data Surveillance and Digital Humanism: An Educational Perspective Through the Medium of Documentary Filmmaking

Theme:
Digital surveillance; documentary filmmaking

Project Period:
Spring Semester 2023

Participant(s):
Arturo Fabbro

Supervisor(s):
Carlos Mauricio Castano Diaz

Copies: 1

Page Numbers: 63

Date of Completion:
May 24, 2023

Acknowledgments

The author would like to thank Elisavet Angouria-Tsorochidou for performing interrated reliability. A thanks is also due to Sandro Fabbro, Nadim Haidar and Eleftherios Kosmidis for their precious theoretical insights.

Abstract

This paper details the theoretical background and the production of an educational documentary film whose main purpose is to get its audience involved in a wider discussion on digital surveillance and technology.

Some contemporary theories on digital surveillance view the extraction of data from users as part of a mechanism of “behavior modification”, a potential threat whose consequences might have troubling implications for the meaning of human agency itself. At the same time, studies show that users might not care enough about their privacy as to take action, due to a feeling of "digital resignation". Getting a user involved into a discussion on his or her own privacy is further complicated by the fact that issues of surveillance and algorithmic intelligence seem almost “non-human” for us to understand. Therefore, the question might arise of of how can digital surveillance be represented and communicated.

Starting from the central role that the concept of "human nature" has in the discourse on surveillance, this paper proposes a theory of “digital humanism” as a strategic shift in focus towards the human user and the experiential nature of interacting with algorithms. To attempt at answering the question of how can surveillance be represented, an educational documentary film was produced according to the framework of digital humanism. The film aimed at encouraging viewers’ self-reflection by looking at the tight interplay between user behavior, algorithmic knowledge, and the folk theories that are developed to understand the world of algorithms. The film included a varied set of characters, their everyday life and beliefs on the topic of technology. Eventually, the film suggested the idea of “personal data obfuscation” as a tactic to deceive algorithms and retain one’s own privacy and autonomy.

Qualitative testing was carried out eight participants, with the aim of investigating whether the film was able to impact the viewer’s own algorithmic awareness, especially in relation to tactics of personal data obfuscation. Results showed that the film as an educational method succeeded in getting the audience engaged on the topic of digital surveillance, which was thought to be very relatable to the viewer’s own everyday life. However, not all communicative goals were reached, and the film was criticised for lacking technical information. Constructive criticism was also raised towards the concept of "personal data obfuscation". Future research can consider the educational framework described in the paper and further broaden it in conjunction with traditional methods for computer science literacy.

Contents

Chapter 1

Introduction

> As people whose data is being collected, what we know of the situation is problematic, and what we do not know is substantial.
>
> - Finn Brunton and Helen Nissenbaum [10]

The vast architecture for tracking, data mining, and behavior prediction put together by big tech corporations such as Google and Meta, has been the cause for increasing concern over the past years. The smartphones resting in our pockets are incredibly powerful tools capable of sensing a highly diverse set of environmental and behavioral data, whether we consent to it or not [45]. Social medias map our network of friends, capture our interests and beliefs, and measure our interaction. In fact, our overall behavior while surfing the internet is tracked and analysed. Internet anonymity is a utopia of the past. Geolocation services track our position at all times. Smart appliances measure our life at its most private, inside the home, while AI systems for facial recognition keep doing the job once we leave our house. More and more aspects of human life (emotions, desires, bodily responses etc.) are captured, turned into analysable data and fed to prediction algorithms. At a time when many services and industries are shifting towards a model of integration with digital systems (education, health care, architecture, transportation etc.), scholars have rightfully raised the question of the possible impact that a large system of surveillance might have on human agency and autonomy.

While there is great variety between the opinions of scholars on the extent and danger of digital surveillance, the will to maintain a questioning attitude towards this system and its ideological foundations is what has driven both research and activism. A similar consensus on the need for caution against digital technology

is not always found among the general public. A common yet perplexing attitude towards technology is what is termed the “privacy paradox” [51][33]: when questioned about it, most users show concern in relation to the amount of private information they share online, and they express the wish of disclosing information only according to their desired degree of privacy. In reality, however, their actual behavior when using digital technology does not reflect their intentions, and they end up sharing more than they had wished for.

Trepte et al. [51] argue that greater privacy literacy among users is what is required in order to fill this gap. Draper and Turow [19] do not share the same idea. They believe instead that an explanation for user inaction is not to be located within a lack of knowledge on the topic; it is rather a very natural response to a feeling of *digital resignation*, which stems from the belief that no alternative to the existing system is possible and that surveillance is inescapable; from a feeling of lack of power over the matter; and from an increased isolation between users over the question of being surveilled. For the authors, digital resignation is actively encouraged by corporate processes that benefit from the neutralisation of criticism and dissent, as it has been historically seen for what concerns the tobacco and mining industries. Thus, in their opinion, change should come from the empowerment of users (see Discussion).

This paper acknowledges both perspectives, and it starts from the consideration that both knowledge of surveillance and the empowerment of users are needed as drives for positive social change in the context of technology usage. Therefore, it sets as its aim that of informing on the topic of privacy and digital surveillance, and, at the same time, raising self-awareness and critical reflection algorithmic interaction. There are, however, many underlying questions as to how can these issues be communicated.

This paper outlines a theory of "digital humanism" as a way to focus on and represent issues of digital surveillance through a method that has the empowerment of people as its goal. To do so, it focuses on the production of a documentary film, titled "Data Surveillance", as an application of the theory of digital humanism.

Qualitative research was performed with an audience of eight people. Test participants were interviewed in order to study how they reacted to the movie, and whether it proved successful as an educational strategy, both for what concerns the communication of the topic of digital surveillance, as well as for its potential to raise self-awareness among its viewers.

1.1 Paper overview

The **Literature review** presents theories that informed the production as well as the communicative goal of the educational documentary film titled "Data Surveillance". Firstly, the question of what it means to portray digital technology is explored, together with the difficulties this poses. Then, the following section provides a brief overview of existing theories on privacy and digital surveillance. The aim of this section is to investigate the extent to which digital surveillance poses a serious threat; where does this threat lie *concretely*; and which criticisms have been moved against popular models of surveillance. The third section presents theories on how users experience algorithms, and the extent to which they become aware of them. Furthermore, the section takes on the topic of "folk theories" that users develop to understand how technology functions. The following section discusses possible counteractions strategies against digital surveillance. Firstly, it starts from considering the agency of people in the *co-constitution* of users and algorithms. Then, a theory of "personal data obfuscation" is outlined as an "everyday" tactic against tracking and surveillance. As such, personal data obfuscation might be proven a useful tool to counteract the influence that "algorithmically-determined identities" have over our own. The last section collects a few theories on the study of digital technology from a "materialist" perspective.

The **Methods** first introduces a theory of "digital humanism" as a shift in focus for the understanding of digital surveillance that has human nature at its centre. Then, the following section details how the previously discussed theories where applied to the production of the documentary film. The last section of the chapter takes on the research question of whether a film approach to the theory of digital humanism is able to educate and empower its audience. Specifically, the question concerns whether the focus given in the film to algorithmic awareness and folk theories is able to raise the spectator's own awareness of his or her relation to algorithms, especially for what concerns the act of obfuscating personal data. The section then goes into detail on how qualitative testing was performed on an audience of test participants.

The **Results** chapter discusses the findings that emerged from the analysis of the interviews with test subjects.

Finally, building from these results, the **Discussion** section presents an evaluation of the extent to which three initial research hypotheses were confirmed as correct or disproved by test data. The chapter introduces new literature to justify the findings, and suggests possible ways in which this research might be improved in the future.

Chapter 2

Literature review

2.1 Representing complexity

An issue that arises when considering how to represent the complexity of digital technology is the form that big tech corporation's business model should take. Following Zuboff [54] (see Surveillance capitalism), understanding what drives big tech's profit is of the greatest importance in order to be conscious of the threat that digital surveillance poses. What is the source of their revenue? Why do they design their products the way they do? And then, most importantly, how can this process be shown? Similar questions concerning the "nature of the *image* of capitalism" have been taken on before in the history of cinema, most famously by Soviet filmmaker Sergei Eisenstein. Eisenstein's question on the representation of capital can be summarised in this way: "How is it possible to see the capitalist mode of production [when] most of the time we only see its effects but not its hidden causes?" [5]. The question arose to him somewhere around 1927, when he decided to embark on a film rendition of Karl Marx's *Das Kapital*, a project he eventually never carried to completion due to the change in the political landscape of the USSR during the 30s. His goal was to start a didactic project, focusing not so much on the conceptual passages of *Das Kapital*, but rather on Marx's dialectical method itself. The approach he devised, termed *cinedialectic*, would make the spectator "feel-and-think" what he or she sees on the screen. To achieve this goal, the director would show "thousands of tiny details" [5] of the lives of people and workers that would serve as direct examples of how capitalism operates.

In order to provide a complete picture of the system of digital surveillance, however, a second issue needs to be taken on, that is how to represent complex algorithms for data analysis. Firstly, by attempting to unpack this problem, an even greater epistemological obstacle might arise, that is how to approach and understand a digital "intelligence" or mode of reasoning that does not follow the

laws of the human mind. This is the type of intelligence possessed by deep learning models that D'Abbraccio and Facchetti [20] call an "alien" one, "or at least one that cannot be linked back to human intelligence"[1]. Quoting an article by James Bridle [8], they posit that "the more powerful our tools become, the less we are able to comprehend them". A similar position is taken on by digital media scholar Lev Manovich when he discusses the possibilities that AI generated art might hold for the future [39]. For Manovich, the capabilities of AI when creating art are not simply an imitation of human work, but they have the potential to go beyond our very understanding of what art fundamentally is, generating something "that we humans are not able to create because of the limitations of our bodies, brains, and other constraints', something that "we would deeply love once we see it".

Secondly, on a more practical level, further inquiry into how complex machine learning algorithms operate is hindered by their proprietary nature, so much that the impossibility to study or understand why an AI models has produced a certain output poses serious questions to authorities and policy makers. The field of "eXplainable AI" (XAI) seeks to influence a shift of paradigm towards a more responsible development of AI tools, one that is built around concepts of "fairness, model explainability and accountability" [1].

2.2 Models of privacy and digital surveillance

2.2.1 Contextual integrity and big data privacy

Accounts of the meaning of privacy differ greatly, going from *reductionist* views that trace back claims of privacy to other values (liberty, security, democracy etc.), to theories that see privacy as valuable in itself [30]. For what concerns digital surveillance, the need to reformulate theories on privacy has been driven by their ability, or lack thereof, to keep yielding satisfactory conclusions in the face of new technological developments [30] [44]. Nissenbaum's [44] formulation of privacy as *contextual integrity* is particularly useful to understand how the meaning of privacy changes from a "human" context to a digital one, that is from sharing personal information between a human actor and another, to a transactions with and through digital medias. According to her account, "there are no arenas of life *not* governed by *norms of information flow*", that is to say that individuals negotiate their privacy in all contexts of public and private life. These negotiations are regulated by *norms of appropriateness*, which determine which information is deemed allowable or even expected *within a certain context*; and by *flow* or distribution of information-movement, that is the transfer of information from one party

[1] My translation from Italian.

to another. Nissenbaum's normative framework of contextual integrity draws attention to when and how norms of appropriateness and information flow have been breached. This formulation is strategic as it exposes a privacy violation in a situation in which it would otherwise go unnoticed, that is the change from one context (such as that of human to human interaction) to another (such as a digital network). A useful example might be that of the gathering of purchasing data from consumers:

> In the past, it was integral to the transaction between a merchant and a customer that the merchant would get to know what a customer purchased. [...] Although the online bookseller Amazon.com maintains and analyzes customer records electronically, using this information as a basis for marketing to those same customers seems not to be a significant departure from entrenched norms of appropriateness and flow. By contrast, the grocer who bombards shoppers with questions about other lifestyle choices-e.g., where they vacationed, what movies they recently viewed, what books they read, where their children attend school or college, and so on-does breach norms of appropriateness.

For an e-commerce company such as Amazon, it would be considered acceptable to collect data regarding the transactions a customer makes, as this type of data collection has always been deemed appropriate within the specific context of monetary transactions. Such is not the case, however, for other types of personal information, which Amazon does collect in great quantity.

Another useful formulation of privacy in a digital context is what Mai [38] terms "big data privacy". Alongside other models for privacy (the "access model", in which privacy is about the ability to limit or restrict others from acquiring information about oneself; and the "control model" in which privacy is understood as the ability to have control over one's own personal information [37]), Mai posits the need for a new "datafication model" of privacy that takes into account not just the way data is collected, but also how it is processed. The example he provides is that of a father who went to a Target store to complain about his daughter receiving coupons for maternity clothing. This happened before a pregnancy test confirmed that his daughter was indeed pregnant. Despite the fact that the daughter had consented to giving out personal data to Target about her purchases, whether the further processing of this information, which yielded the result that the woman was pregnant, constituted an actual breach of her privacy is, as Mai puts it, a "complicated question", as it does not fall under neither the "access model" of privacy, nor the "control model". Thus, the processing of personal data through data mining algorithms requires the definition of privacy to be further expanded.

2.2.2 Surveillance capitalism

The wide-ranging implications of datafication have been explored in context not limited to that of privacy. In *The Cost of Connection* [15], authors Couldry and Mejias relate the global process of data extraction, an information flow from "human life in all its forms to infrastructures for collection and processing", to the extractive nature of colonialism. For them, this process of extraction is "degrading" to life, it "abstracts" life and appropriates it through a conversion into value.

A similar emphasis on extraction is given by surveillance scholar Shoshana Zuboff. Although she shares similar goals to those of Couldry and Mejias, namely to make sense of the deep changes that happened in the past few decades which, according to all authors, have no precedent in human history, Zuboff's theory of "surveillance capitalism" constitutes a qualitative shift from focusing solely on the predictive properties of data processing, to its potential for "behavior modification". In her popular study from 2019 [54], Zuboff formulates the logic, intents and historical origin of a new global architecture founded on data extraction and surveillance. According to her research, users are exploited by big tech corporations for the extraction of "behavioural surplus" through their interactions on the digital world (*extraction imperative*). This surplus of data is then utilized to make predictions related to the physical world, to users' daily life, to their bodies and their selves (*prediction imperative*). The more accurate predictions become, the more they will be able to foresee users' intentions before they are formulated. As such, a "perfect" prediction is going to act upon a user's selfhood as an actual modification of his or her behaviour, desire, and will. Zuboff's theory, backed by numerous statements by big tech executives, who seem to confirm the ultimate goal of predictive algorithms as that of behavior modification, has wide-ranging implications. Behavior modification is formulated as a serious threat to human nature, to human autonomy and sovereignty, to the democratic order and liberal State, and to our very future.

In her theory, Zuboff is less concerned with coming up with a conceptualisation of privacy that might be useful in the context of surveillance capitalism, as in that scenario privacy itself would cease to be a meaningful concept. Not only is the meaning of privacy rendered unclear in this new scenario, but the future heralded by behavior modification poses a whole series of novel epistemological problems. Even our comprehension of what surveillance capitalism fundamentally is is problematic, as it is "unimaginable outside the *inscrutable* high velocity circuits of Google's digital universe"[2]:

> [...] the regime's most poignant harms, now and later, have been difficult to grasp or theorize, blurred by extreme velocity and camouflaged by expensive and illegible machine operations, secretive corporate prac-

[2]My emphasis.

> tices, masterful rhetorical misdirection, and purposeful cultural misappropriation. [55]

As such, Zuboff's theory is traversed by the issue of *understanding* which powers and which processes are at play in the system of surveillance. To overcome this problem, she might argue, we need new words and new images.

Different scholars have criticised Zuboff's work from a variety of angles, albeit sharing the gravity of her concerns. Evangelista [24] points to how Zuboff's work, by focusing almost exclusively on the operations of Silicon Valley corporations, effectively universalises the paradigm of surveillance capitalism without having first considered how it might play out in countries of the global South. Marxist critics have questioned Zuboff's formulation of surveillance capitalism as a "mutation" from the previous paradigm of industrial capitalism (instead of a continuation of it), as well as her reliance on the need to return to a "good" form of capitalism [34]. Furthermore, different authors [34] [3] question Zuboff's lack of solutions or alternatives to the paradigm she proposes. Lastly, the theory that is possibly the most poignant in undermining the premises of Zuboff's own is the one that Tim Hwang proposes in his book *Subprime Attention Crisis* [31]. Hwang casts serious doubts on the robustness of the industry of predictive online advertising and microtargeting, in a way that renders Zuboff's ideas on behavior modification markedly far off. Drawing a comparison with the 2008 financial crisis, Hwang argues that the industry of targeted advertising relies heavily on speculation, and that it ignores actual data showing the high costs and low profits. In part, this is due to an overestimation of the algorithms' ability to properly profile a user, while in reality their predictions might carry gross mistakes. This is a considerable resizing of their apparent pervasive powers. Furthermore, and most importantly, the greatest amount of digital advertising is hindered by digital ads fraud, low quality content, clickbaiting and click farms, bots [3] and even simply ad blockers, so much that, according to a 2017 study [6], 56% of the amount of money spent on digital ads the previous year was lost to fraudulent activity or unviewable content.

2.3 Algorithmic knowledge and folk theories

2.3.1 Experiencing algorithms

Learning how algorithms operate is made complex by their "black box" nature [47]: it is not possible to look "inside" an algorithm, both because it is protected as intellectual property, and also due to the increased complexity of whole clusters of

[3] According to a 2022 report, 42% of all internet traffic is constituted by bots, and 28% of all traffic is constituted by "bad" bots mimicking human behavior [2].

"algorithmic systems" [32]. However, as Cotter and Reisdorf [14] point out, algorithms remain knowable to some extent. Although different strategies have been suggested, such as that of reverse engineering [18] [32], Cotter and Reisdorf's study focuses on the idea of algorithms for news feed curation as *experiential technologies* [4]. They divide the sources that users have available to gain insight into the operational logic of algorithms between "exogenous" and "endogenous" sources. Exogenous sources of knowledge can be texts, media reports, and educational materials on algorithms, computer science and coding. However, as actual information on algorithms is often lacking, they argue that most learning occurs endogenously, that is through direct experience with algorithms. As users interact with a platform's algorithm, they *intuitively* form beliefs about how algorithms work. Quoting from Eslami et al. [23]:

> Just as the Facebook News Feed algorithm is likely trained by the act of clicking "like," so is the Facebook user *trained* by the algorithm's dissemination of some posts and not others. [5]

Building from the idea of algorithms as experiential technologies, Swart [49] identifies "explicit" behaviors through which users interact with feed curation algorithms (i.e. by personalising the feed parameters) and "implicit" behaviors (e.g. liking and sharing preferred posts, as explained in the above quotation). In her study, Swart looks at the habits that teenagers have when interacting with algorithms, and instead of focusing on what they actually know about algorithms (e.g. how algorithms mine certain types of data from users to create a customized feed), she considers the *perceived knowledge* and *tactics* that young people employ in their daily usage of digital technologies from a cognitive, affective and behavioral perspective. Thus formulated, this framework is helpful in tackling three issues when attempting at studying algorithmic knowledge [49] [28]:

a) although digital literacy is often assessed deductively, such an approach cannot be applied to the study of algorithmic knowledge, as their black-boxed nature makes it impossible to benchmark a user's skills. Thus, as no a priori measure of what an algorithmically literate user might look like cannot be known, a bottom-up approach is needed.

b) The opaqueness of algorithms makes them invisible to users. That is, according to the principle of "seamless design" [27], an algorithm that works properly is one whose curatorial work is not noticed by a user. Only when an algorithm produces unexpected or uncanny results is the awareness of a user triggered.

[4] The rest of the discussion in this section will also focus primarily on algorithms that regulate feed curation on social media platforms.

[5] My emphasis.

c) Even using a term such as "algorithm" can be problematic, as not all users are aware of such a terminology. Despite this lack of a proper vocabulary, a person might very well be able to reflect upon how algorithms for feed selection operate.

Given the importance of the act of experiencing for the attainment of algorithmic knowledge, a question might arise of whether users are actually aware of it, or if the algorithmic knowledge that is acquired through experiencing is only a form of intuitive and unselfconscious learned interaction with the algorithm. A 2015 study [22] found that about 63% of users was unaware of the presence of algorithmic feed curation on Facebook, despite having developed strategies to interact with it. This should not surprise since, as Hamilton et al. [27] argue when tracing the history of user interaction design, the goal of seamlessness in the design of algorithms is that they should feel "effortless" or invisible to the human user. Designers should first consider patterns of human interaction, and build their codes from a process of "designing with" users, through a collaborative deduction of when and how awareness to algorithms arises.

2.3.2 Folk theories

In those instances in which users become aware of the curatorial work performed by algorithms on their social media feed, Eslami et al. [23] posit that people develop "folk theories" about why such changes occurred. Furthermore, for what concerns the group of "unaware users", they find that, once they have been made aware of the presence of algorithmic curation, they develop similar folk theories to justify the algorithm's decisions to those theories that the "aware group" had since the beginning of the experiment. In this sense, we might say that folk theories operate as the common sense theoretical grounding to the intuitive knowledge that users gain through experiencing algorithms. Folk theories may not reflect the actual logic followed by software engineers when developing the system, but they can impact user behaviour in positive ways, and even shape the evolution of the system [23]. In the same study, Eslami et al. identify a number of popular folk theories, such as the one, for example, according to which the amounts of likes and comments a content has received can be a measure of the likelihood of that content to appear on a user's feed. Other popular theories might concern the size of the actual audience a social media post can reach [4], or whether advertising companies surreptitiously gather audio data from a smartphone's microphone [49].

2.4 Strategies for counteraction: personal data obfuscation

2.4.1 User agency

A formulation of algorithmic knowledge as the one stated in the above section does not necessarily take into account the agency of a user. That is, it sees the process of gaining experience of algorithms as one-directional, going from a platform to the person interacting with it. This perspective is consistent with, and feeds into the idea of *digital resignation* (see Introduction), wherein a user who lacks any power over his or her situation is brought to inaction.

Pangrazio and Selwyn [46], however, take a different approach to the study of teenagers' understanding of social media data, stating that they "[took] care not to be pejorative about young people's agency" and that they decided not to push for a normative ideal of personal data practices. This is quite a departure from a common perspective on the topic of child technology usage and safety, which understands young people as having an "addiction" [40], thus equating teenagers' agency against algorithms to that of substance abusers. An example of how a user (in this case, also a young person) might impose his or her agency over the algorithm is given by Eslami et al. [23]:

> Interestingly, there were participants who used [a folk theory stating that the more interactions you have with a type of content, the more it is going to show up] in reverse, trying to counteract their own previous interaction to avoid the effects they theorized: "When I 'like' something, I usually hide it from my News Feed because I like it but I don't necessarily want to know all about it all the time". While they wanted to send a signal to their friend via the "like" feature, they did not want their News Feed to change.

A theory of algorithmic knowledge that takes into account a more reciprocal relationship between user and algorithm might therefore start from Kitchin's consideration [32] of algorithms as being "ontogenetic, performative and contingent", meaning that they are not fixed in nature, but are rather part of an emergent and unfolding process. For Lomborg and Kapsch [36],

> people's experiences of what algorithms are and what they can do to them is part of the shaping of the output of algorithmic operations. Different tactics for and acts of circumventing, say, algorithmic profiling on social media are manifestations of user agency.

This entails a recursive relationship between algorithms and people, one that

> [acknowledges] both the material underpinnings and mathematical logics of algorithms as 'technical entities' and the agency and sense-making of people who experience, valorize, and perhaps tactically try

> to resist algorithmic operations in their practical pursuits through daily life. [36]

Velkova and Kaun [53] share similar ideas on the topic of user agency and tactics for resistance:

> User agency is often neglected in the emerging discussion of the consequences of algorithmic culture. [...] As algorithms assume a dominant role in the mediation of power, it becomes increasingly important to consider to what extent and in what ways their power can be resisted.

This "resisting" to the power of algorithms starts from "acknowledging the *mutual co-constructions* of algorithms and their users"[6] [53].

2.4.2 A tactic against surveillance

Brunton and Nissenbaum [10] sum up and critique four cases of what are commonly believed to be good directions to limit the power of digital surveillance:

- **user opt-out**, the idea that a user is freely allowed to give consent, or take it back, over his or her online privacy is rooted in the idea that matters of data protection are an individual responsibility. More realistically, to depend entirely on personal choice entails that all but a few privacy-enthusiastic users are going to gloss over complicated privacy policies for ease of use, and give out their unquestioned consent.
- To rely on **corporate best practice** for a responsible use of tools for data surveillance would be a misguided approach, as it is simply not in the interest of big tech corporations to support general restraints on access to information.
- Imposing **laws and regulations** on the tracking of user data is a slow and gradual process, and cannot effectively tackle the fast speed to which technologies for surveillance develop.
- To count on **technological implementations** for the enhancement of digital privacy might impose technical challenges. Furthermore, despite tools such as Tor browser and proxy servers are very powerful for preserving one's own privacy, they are still out of reach for the majority of digitally unskilled users.

What the authors propose instead is a strategy of **data obfuscation**, a tactic both personal and political of *informational self-defense*, which can serve as a method for "informational resistance, disobedience, protest or even covert sabotage". Obfuscating means producing data that is misleading and ambiguous, plausible but

[6] My emphasis.

confusing, with the goal of covering one's tracks. Perturbing data with noise and sending out inaccurate data are means to enhance one's own privacy [25].

Obfuscation is, of course, not limited to the digital world. Brunton and Nissenbaum provide different types of obfuscation, together with their relative examples. According to a historically inaccurate story, for example, the population of Denmark wore the Yellow Star to make it impossible for the occupying Nazi forces to distinguish them from the Jews, adopting a strategy that is based on obfuscated and misleading information [10].

Brunton and Nissenbaum formulation of obfuscation does not refer directly to the context of personal data extracted by tracking algorithms, although their theory is further developed in the book *Obfuscation: A user's guide for privacy and protest* [9] to include the idea of *anti-profiling obfuscation*. In the specific co-constitutive relationship between users and algorithms, however, personal data obfuscation might not necessarily be a plan that is deployed having the conscious goal of sabotage in mind. As Velkova and Kaun point out [53],

> users also engage in negotiations of the meaning and functionality of technologies through multiple uses that may comply with or deviate from the 'original' meanings envisioned by designers. [...] These tactics can be based on users complying with algorithmic logics but resisting their output, [or on] 'tricking' algorithms to work toward unintended ends [...].

In this sense, tactics that make use of obfuscation or inaccuracy are already *embedded* into users' everyday interaction with technology. van der Nagel [43] draws on de Certeau's [11] formulation of *tactics* as calculated actions of resistance that lack a view of the whole. She provides the example of *screenshotting* as a tactic deployed by users to access certain kinds of content without giving profitable clicks and views to undesired websites.

2.4.3 Digital identities

> In a perfect world, all of us should be allowed some short vacations from our own identities.
> - Orson Wells [48]

The implications of deploying a strategy for counteraction based on obfuscation are not limited to the pollution of big tech corporation's databases. With reference to the centrality that a threat to "human nature" has for Zuboff's theory of surveillance capitalism (see Surveillance capitalism), personal data obfuscation might be a powerful tool to hinder the influence that algorithmically determined identities have over our own. In his book *We Are Data: Algorithms and the Making of Our*

Digital Selves [12], John Cheney-Lippold discusses the processes and politics that lead to the formation of users' "digital identities", the many layers of "who we are online" that are decided by advertisers, tech corporations, and governments. The logic of the formation of such identities is not one that actually takes into account the real, embodied individual it refers to, but has as its subject a purely "statistical body": "who we are in the face of algorithmic interpretation is who we are computationally calculated to be" [12]. Still, for as much as these identifications are not a direct reflection of our lives, they do affect our real-life identity and behavior. The mismatched relationship between digital identities and our own identities is thus described by Cheney-Lippold:

> The complexity of our individual histories cannot be losslessly translated into a neat, digital format. [...] In this algorithmic reality, there is instead a dependency on a data-based model of what it means to be 'famous,' 'not famous,' 'man,' 'woman,' 'gay,' 'straight,' 'old,' 'young,' 'African American,' 'Hispanic,' 'Caucasian,' 'Asian,' 'other,' 'Democrat,' 'Republican,' 'citizen,' 'foreigner,' 'terrorist,' or 'college educated.' [...] Each quotation-marked classification is an algorithmic caricature of the category it purportedly represents. These algorithmic caricatures, or what I call measurable types, have their own histories, logics, and rationales. But these histories, logics, and rationales are necessarily different from our own. Google's 'gender' is not immediately about gender as a regime of power but about 'gender' as a marketing category of commercial expedience.

The pigeonholing of people's lives into fixed categories by market ideology is not something new on the terrain of media activism. An example of how tactics of obfuscation have been put into place to disturb the notion of a stable identity can be found in Luther Blissett's[7] *Mind Invaders* [7], a fervidly imaginative book on "media guerrilla and sabotage". Writing at a time where internet usage had just recently broken through the mainstream, Blissett considered the potential for internet anonymity, and the threat of the "reactionary ideology" of stable identities. To counteract it, he proposed the implementation of a tactic of "multiple names", a word Blissett borrowed from the American avant-garde of the 70s and 80s, wherein different people would be wearing "the same *mask*":

> The ultimate goal for different people to use the same name is that of creating a situation in which no one is responsible in particular, and of further putting under practical examination the notions of identity,

[7] Luther Blissett was by itself a project of "collective name" for media activism that was put in place across different European countries during the 90s and early 2000. For a thorough overview of the Luther Blissett project see Deseriis [17].

individuality, originality, value and truth coming from Western philosophy.[8]

2.5 Digital materialism

The fundamentally *extractive* nature of surveillance capitalism has been noted, among others, by Couldry and Mejias [15] (see Surveillance capitalism). Mazzadra and Neilson [41], however, move beyond the mere *comparison* between material extraction, such as that found in the context of mines and plantations, and data extraction, to postulate how both phenomena are offshoots of the same global process of dispossession. To illustrate the implications of this twofold nature of extraction, Crawford and Joler [16] use the story of the "Mechanical Turk":

> In 1770, Hungarian inventor Wolfgang von Kempelen constructed a chess-playing machine known as the Mechanical Turk. His goal, in part, was to impress Empress Maria Theresa of Austria. This device was capable of playing chess against a human opponent and had spectacular success winning most of the games played during its demonstrations around Europe and the Americas for almost nine decades. But the Mechanical Turk was an illusion that allowed a human chess master to hide inside the machine and operate it.

Thus, the wondrous achievement of technological innovation turned out to be an illusion executed by a man hidden in a cramped space. According to Crawford and Joler, shifting our focus towards the material life of technological development stands in complete opposition to the preferred view of tech corporations, which rely on an imaginary of abstraction:

> [...] the *ethereal metaphor*[9] of 'the cloud' for offsite data management and processing is in complete contradiction with the physical realities of the extraction of minerals from the Earth's crust and dispossession of human populations that sustain its existence.

In their interactive work[10], Crawford and Joler take the Amazon Echo, a small voice-activated virtual assistant that can be purchased for a low price, and whose lifespan is not durable, as an example of how the "ethereal metaphor of the cloud" allows for the concealment of the long and laborsome production cycle of the small object. Not only is extraction a part of the life of the object *before* its sale (mining extraction) and *after* (data extraction), but the same continuity is to be found in its nature as a "black box" object: *after*, as a propriety device whose code and

[8] My translation from Italian.
[9] My emphasis.
[10] `https://anatomyof.ai/`

functioning cannot be known; and *before*, as the result of a long production cycle that is concealed and made impossible to trace[11]:

> [...] The very processes of creating, training and operating a device like an Amazon Echo is itself a kind of black box, very hard to examine and track *in toto* given the multiple layers of contractors, distributors, and downstream logistical partners around the world. [16]

Not only is land mining exploitation purposefully hidden inside the black box of the Amazon Echo, but so is the very real digital labor that is exploited for its production and training:

> Digital labor – the work of building and maintaining the stack of digital systems – is far from ephemeral or virtual, but is deeply embodied in different activities. The scope is overwhelming: from indentured labor in mines for extracting the minerals that form the physical basis of information technologies; to the work of strictly controlled and sometimes dangerous hardware manufacturing and assembly processes in Chinese factories; to exploited outsourced cognitive workers in developing countries labelling AI training data sets; to the informal physical workers cleaning up toxic waste dumps. [16]

[11] The impossibility to trace the cycle of production has been brought up by tech corporations as a legal defense against accusation of workers' exploitation, as it happened, for example, in the recent case of child labor employed for the extraction of cobalt in the Democratic Republic of the Congo [50].

Chapter 3

Methods

In the following sections, a theory of "digital humanism" is going to be introduced. Then, some considerations will be outlined as to how such a theory informed the design and production of a documentary film. Finally, the last section of the chapter is going to provide details on how testing was devised and performed in order to study the validity of three research hypotheses.

3.1 Design

3.1.1 Digital humanism

A shift towards "humanism" can be used as a framework to communicate and represent digital surveillance. Such a shift echoes and, to some extent, is a prosecution of the turn towards the material life of technology (Digital materialism) as a way to understand the twofold nature of global extractivism. In a similar way, the focus would be strategically shifted towards the concept of "human nature" as the starting point of, the referent for, and the end recipient of a discourse on surveillance. The focus is not going to be directed towards the *effects* of technology, nor to the *interaction* of humans and computers, but rather, from a phenomenological perspective, to the human subject's *experiencing* of technology, in the world.

The framework of contextual integrity (Models of privacy and digital surveillance) is useful in this situation to understand how concepts of privacy change from situations that are at reach for the human subject to comprehend, based on common sense, to a context that does not belong to everyday human life.

If AI algorithms employed for digital surveillance do constitute a "non-human" mode of reasoning (Representing complexity), then "human" life is the grounds onto which concepts of surveillance should be approached for communication. Furthermore, if human nature is fundamentally at stake due to a vast scale system of digital surveillance (Surveillance capitalism), one should first question what

does "human nature" fundamentally mean, and which part of it is at stake of being lost. A risk, for example, is the proliferation of statistically-determined digital identities (Digital identities) and the hold they can have over the lives of people. Against the splitting of different "layers" of identity, human nature should be understood in this context as *universal and unitary* in its essence.

The goal of digital humanism as a framing of digital surveillance, is the communication of the topic and the education of people. Its starting point is the understanding of algorithmic knowledge as a consequence of *human experiencing* of technology (Experiencing algorithms). If human agency is given importance in the co-constitution of users and algorithms (User agency), then it is possible to intervene in three sites where knowledge of algorithms is produced and located.

1) **Raising self-awareness of algorithmic knowledge**: for users that are *unaware* of their own degree of algorithmic knowledge, as a form of learning by doing that is developed instinctively, the educational goal would be for their self-awareness of the tactics of algorithmic knowledge to emerge. This seems consistent with D'Ignazio and Bhargava [21] proposed definition of what "big data literacy" means, as understanding and identifying the patterns of algorithmic manipulation.

2) **Engagement through folk theories**: since folk theories are understood as the common sense theoretical grounding that users develop from their experiential knowledge of algorithms (Folk theories), it follows that an educational project should engage with, and confront those very same theories that are believed, either erroneously or correctly, to guide algorithmic operations, instead of more sound, but complex scholarly notions [23]. That is, the same "language" that users develop to describe their own experience should be spoken.

   A further question, then, is how to place folk theories in relation to a proper scientific knowledge of the topic. As an answer to this question, this paper proposes a "non-authoritative" process of learning. Instead of serving as the recipient of information stemming from a more knowledgeable source, a viewer would learn by *observing* a behavior. This behavior, in turn, is one that is informed by an advanced study of surveillance and algorithms.

   The following diagram exemplifies the continuous interaction between how a user acts, how he or she develops instinctive knowledge of algorithms, and how this knowledge is grounded in a folk theory:

   Folk theory $\rightleftarrows$ **Algorithmic knowledge** $\rightleftarrows$ **Behavior**

   The educational process that is being here proposed can be therefore sketched as follows:

$$\textbf{Folk theory} \rightleftarrows \textbf{Algorithmic knowledge} \rightleftarrows \textbf{Behavior}$$

$$\textbf{Advanced education} \rightarrow \textbf{Algorithmic knowledge} \rightleftarrows \textbf{Behavior}$$

The bottom line shows how, for what concerns educated users, knowledge of algorithms is both the product of an experiential negotiation with them, and the consequence of a well-grounded scientific understanding of the topic. Then, the way an educated user behaves can be used as a point of reference for a person who is not educated on the topic of technology to reflect upon his or her own actions.

3) **Confronting personal data obfuscation**: obfuscation of personal data is seen as a tactic that is already embedded into the everyday interaction of users with algorithms. Unlike other proposals to counter digital surveillance (see A tactic against surveillance), obfuscation can not only be performed by a single individual, but it is also an instinctive tactic already normally enacted by people. Therefore, similarly as in point 1), educating on the topic of surveillance would mean for user self-awareness of his/her own, already-present tactics of obfuscation to emerge, so that they can be further put into practice in a self-critical manner to enhance personal privacy.

3.1.2 Documentary film

Inspiration was drawn from Eisenstein's theory of *cinedialectic* (Representing complexity) as a way to translate the communication theory of digital humanism into the form of a documentary film. Specifically, as Eisenstein considered the particular as a way to film a method, in a similar way particulars (e.g. singular people, life situations, beliefs and behaviors) were chosen to show instances of a greater concept of "human nature".

Interviews

Through filming and editing, this approach took shape in the "long interview" format. Interviewees were asked if they wanted to take part in a documentary film project that dealt with the topics of technology and surveillance. They were given a consent form to be signed[1], which informed them about which data would be gathered from the interviews, which control they had over it, and who was the data controller. Furthermore, they were informed that if something came up during the interviews that they did not desire to be shown, it would not have made the final

[1]The consent form was adapted from a standard model complying with GDPR rules provided by Aalborg University. For further information visit: `https://aaudk.sharepoint.com/sites/GDPR-for-students`.

cut of the film. The table below lists some of the features that were highlighted through the "long interview" format.

Human figure	The human figure was given a central role in the film. This often meant positioning the camera perpendicularly to the filmed subject.
Uninterrupted editing	Conventional editing of interviews gives primary importance to the content of an interviewee's speech. Therefore, to highlight an important phrase or word, the continuity of the talk is split into fragments of information. Here the opposite was tried out: to constitute a unity (which is in itself the unity of the human subject as a broader concept) that would draw attention to the speaker himself/herself, rather than to the information he/she provided. This meant, for example, not doing away with, but actually valuing flaws of speech and accents, stuttering, uncommon facial expressions and wandering gazes.
Discursive approach	The primary purpose of the interviews was not to reach the audience with once specific information. Rather, they were intended as an unfolding conversation between the filmmaker (sitting behind the camera) and the filmed subject, thus giving the impression to the viewer of being directly involved with the conversation, not as a third party listener (as it would be if both filmmaker and interviewee were shown together in the frame), but by assuming the role of the interviewer himself.
Focus on identity	The interviews allowed for isolated details about the interviewees' life to emerge occasionally. This served a double purpose: on the one hand, it gave the characters a life beyond their function as mere speakers; on the other, it was meant to bring attention to how certain specific details of people's lives are captured by tracking algorithms for the purpose of shaping a statistically determined identity.

Characters

A diverse set of characters was chosen as to represent the variety of human experiencing in relation to technology, drawing attention to how different ages, genders and professions are echoed by different attitudes on technology. This was not meant to imply a correlation between certain demographics and their opinions on technology, but rather the opposite: it meant giving value to the variety of human experiencing, and recognising that people reach different (and varied) life perspectives that cannot be predicted by the statistical application of a label. The characters included:

- four scholars, engaging with the topic of surveillance and algorithms from different perspectives relating to their own fields of study.
- Four teenagers, who were asked to tell and engage with their own beliefs and attitudes in relation to digital technology.
- A young, single mother and her 4 years-old child, who is already quite proficient in the use of a smartphone;
- Two lovers, who expressed their views on dating apps.
- The owner of a shop that specialises in products for smartphones.

Implementation of educational aims

Editing allowed for comparisons between different interviewees to be drawn by the spectator. The first of the three educational aims of the documentary film (raising self-awareness of algorithmic knowledge) was therefore accomplished through the juxtaposition of different people's retelling of their attitudes towards technology.

Secondly, and most importantly, valuing the opinions of scholars not as coming from an authoritative standpoint, but rather as the experience of any other person, meant that their attitude towards technology could be placed on the same conceptual level as to that of other, non educated characters of the movie. This allowed for an horizontal comparison regarding how our own knowledge influences our behavior in the context of technology. A key choice, in this sense, was doing away with the standard use of on-screen labels for name and profession that are usually used in documentary filmmaking to introduce a character. Not labelling professors as such meant that their opinion could be valued with the same standard as to that of the rest of the characters.

A scene titled "Teens with Phones" served as a key moment for the presentation and discussion of algorithmic knowledge and folk theories. In this scene, teenagers were asked to engage with their own beliefs as to why social medias

function the way they do. To film this scene, inspiration was drawn from the different techniques highlighted by Hairgattai et al. [28] to study people's algorithmic knowledge. These techniques included:

- sketching a diagram of which data do different social medias gather;
- mentally recalling how specific algorithmic habits are performed;
- inspecting the feed of one social media that is used frequently, and attempting at explaining why certain content is being recommended.

Lastly, the theme of personal data obfuscation was presented by an interviewee, who discussed it through the lens of "playfulness" towards algorithms. To elaborate the concept further, the playful attitude of a toddler (one of the characters) was then used as a metaphorical representation of "playfulness".

Production details

Shooting was carried out over the course of two months. A Canon XF100 was used as a primary camera, together with a Canon Legria for backup. Editing was done using DaVinci Resolve. The final edit of the film, which was titled "Data Surveillance", ran for 63 minutes.

Production of the documentary film was restricted due to limitations of time, personnel and budget. These limitations carried an important significance for the final form of the movie. For some indications on how different production resources could be used to further develop the project, see Discussion.

3.2 Testing

The research question is formulated as follows:

RQ: Can a documentary film produced according to the framework of digital humanism educate its audience on the topic of digital surveillance?

Three hypotheses were considered:

Hypothesis 1 The film is able to raise self-awareness of algorithmic knowledge.

Hypothesis 2 The film is able to involve the viewer using the "language" of folk theories on technology and algorithms. By affecting the interplay between behavior, algorithmic knowledge and folk theories, it is able to engage in the non-authoritative communication of complex topics.

Hypothesis 3 The film is able to raise self-awareness of and critical engagement with tactics of personal data obfuscation, which are already part of a user's algorithmic habits.

Test participants were recruited through convenience sampling by word-of-mouth. Test participants had to follow two criteria [23]:

1. Test participant must be a user of digital technologies.
2. Test participant should not have an advanced education in the field of digital media and technology. That is, an "insider" opinion cannot be taken into consideration.

Eight participants were recruited to take part in the testing phase. Six of them identified themselves as women, and two as men. Participants' ages ranged from 21 to 37 years old, with an average of 27. All participants but one had pursued, or were currently enrolled in university-level educations. The number of test participants was the maximum amount that could be recruited due to constraints of time and resources. After coding was performed, it remained unclear whether saturation was reached; that is, the point in data collection where no new issues are identified [29].

Testing was performed on a preliminary cut of the movie, which did not feature proper audio mixing. Screening of the movie was repeated four times: two test participants watched the film by themselves on separate dates, while the remaining 6 participants watched the film on two occasions in groups of 3. Screening formats included a laptop screen, a television and a projector. The researcher was physically present during all screenings but one.

Interviews were conducted shortly after the screening ended. Interviews were chosen as the preferred method for data collection as is often the standard for qualitative research of audience involvement with a film [13] (together with focus groups and written reports [52]). In this case, individual interviews were chosen as a method to study a spectator's own perception of the movie, and relate it to his or her own experience of technology.

Interviewees were made to sign a consent form[2] detailing which data was gathered from them, how they had control over it, and who was the data controller. They were also made aware that, for what concerns the discussion of their data in the research paper, pseudonyms would have been used to protect their identity.

All interviews were carried out in English, with the exception of one, which was performed in Italian. This last interview was carried out through an online video

[2]The consent form (see Appendix C - Consent form) was adapted from a standard model complying with GDPR rules provided by Aalborg University. For further information visit: `https://aaudk.sharepoint.com/sites/GDPR-for-students`.

call, while all the other interviews were made in the presence of both researcher and interviewee. Interviews lasted for about 15 minutes each.

Grounded theory was chosen as the design framework for the qualitative research, adopting an *interpretive* paradigm. This entailed giving importance to interpretation and observation. Study of the experiences of test participants was understood as starting from test participants *themselves* [29], valuing the subjective experience that they had while watching the movie. The experiences of test participants were discussed through a semi-structured interview. The interview was aimed at letting a viewer recollect her or his thoughts and opinions of the film, while at the same time guiding her or his focus towards matters of the film's form. Interviews were recorded and then transcribed using AssemblyAI[3]. To perform the analysis, the text of the interviews was slightly edited for greater readability. The interview that was performed in Italian was translated into English.

Axial coding was performed on the interviews' text, and it allowed for the isolation of some key central phenomena (namely viewing experience, critical engagement, knowledge and algorithms and personal data obfuscation) and to then identify the contexts they were brought up in, and delineate which consequences they carried [35]. Codes were devised following the inductive, bottom-up approach of grounded theory. Codes were chosen as non overlapping, and describing a unit of text roughly corresponding to one paragraph. The codes that were used had both an interpretive as well as a descriptive function. For a detailed description of the codes, see Appendix A - Codebook.

Percentual inter-rated reliability, performed by randomly sampling 25% of the test data, yielded a 60% agreement rate. After this process, a discussion between the author and the person who performed the reliability coding lead to a refining of the codes, so as to ensure a more precise formulation. However, a second round of coding to re-calculate inter-rated reliability was not performed due to time limitations.

Due to the low number of participants, testing was marked by significant limitations. Particularly, test data reflected a low degree of diversity among test participants in terms of age, gender and education.

[3] `https://www.assemblyai.com/app/`

Chapter 4

Results

Analysis of the interviews led to the formulation of codes that could describe the different perspectives that test participants assumed when engaging with the film. At the beginning and at the end of the interviews (see Appendix B - Semi-structured interview model), viewers were asked an open-ended question on which thoughts did the movie raise within them. From the analysis of their answers to these open-ended questions, there emerged a number of initial codes that were then grouped under three categories:

- "**viewing experience**" concerned how spectators engaged with the movie, whether they felt it having a "personal" approach, whether it was "relatable", if they perceived it as "effective", etc.;
- "**critical engagement**" grouped a few considerations some viewers expressed that went beyond the mere experience they felt when watching the movie, but that were the expression of a greater degree of intellectual involvement with its form and its topic;
- "**knowledge and algorithms**" included considerations the spectators made on their own algorithmic habits, whether they expressed a self-reflective attitude towards their own folk theories, and the criticism some raised against the model of non-authoritative knowledge described in the Digital humanism section.

The rest of the questions that participants were asked during the interviews went into greater detail about specific topics that the participant would have been unlikely to raise otherwise. An example was the topic of **personal data obfuscation**, which went on to form a codes' category on its own. In relation to this topic, participants expressed a diverse set of views, ranging from having never experienced a tactic such as that of personal data obfuscation, to having performed

similar actions in the past, to having formulated strong critical opinions in relation to how it was presented in the film.

Lastly, further specific questions were meant to probe the test participant about how he or she felt in relation to an isolated aspect of the movie. The answers they expressed in these contexts were grouped under the "viewing experience" category (concerning, for example, the degree to which viewers had been focusing on the "human" side of characters); under "knowledge and algorithms", if the participant made specific reference to the scene titled "Teens with Phones"; or they were understood as constituting a further "critical engagement" with the film.

4.1 Viewing experience

Results show that the audience experienced engagement and interest when watching the movie. Almost every participant remarked how relatable the topic of digital surveillance was for them. *"I think it's like thoughts that many people think about. Like for example, I think many can relate to ads popping up right after you've talked about a certain topic*", remarked Mette (for interviewees' demographics, see table 4.1), while Helena said the movie "*totally*" made her think about her own way of acting towards technology. Lucia added that the movie resonated well with her own thoughts, as she thinks "*quite obsessively*" about how her data is being used. Two people related to the character of the 4 years-old toddler, as she reminded them of children they know personally. Ana further added that the scenes with the child and the ones with teenagers felt "*comforting*", as they showed "*our daily lives*".

The way the movie structured the interviews was perceived as having a "personal" approach by 63% of viewers. One viewer experienced the film's structure as "*effective*". A few people remarked how the "interviews" felt more like an actual discussion. *"[The film's form] involved me quite a bit. [...] Not seeing the two of you within the frame helped me as if I was the interlocutor. Therefore, it was almost as if the*

Name	Age	Gender	Nationality
Ana	25	F	Portuguese
Emilie	28	F	Danish
Helena	28	F	Greek
Igor	34	M	Croatian
Lars	32	M	Danish
Lucia	25	F	Italian
Mette	21	F	Danish
Nico	22	F	Swedish

Table 4.1: An overview of the test participants' information.

person was speaking to me, and the fact that they did a continuous speech felt to me like an actual conversation" (Lucia).

Furthermore, 75% of the audience reported shifting their focus from the content of the interviews, to details of the person who was speaking, such as facial features and accents, especially during close ups. One viewer reported paying attention to a character's wrinkles, while, according to another, a character's face "*turned red*" during his interview. For Emilie, with reference to the scene with the owner of the shop for phone accessories, showing how different people related to technology gave her some considerations on the relationship between human nature and technology: "*I think [the scene] fit [in the context of technology] in that we are all made, like, the same thing in the eyes of technology. We are just data. We are all just information. But you actually zoomed in on the person and had him tell a personal story*".

Codes related to how viewers experienced watching the film were not all positive. In a few instances, participants formulated their answers using words that denoted uncertainty ("I guess", "maybe", "a bit" etc.). Often, the participant would extend his or her answer with tentative information that was not relevant to the question. Those cases were recognised as being indicative of a lack of proper engagement with the interviewer's question, and therefore a lack of a viewer's reflection to the specific aspect of the movie that the question was hinting at.

4.2 Critical engagement

Some spectators had a more structured critical engagement with the film. Two viewers argued about how different scenes of the movie related to each other conceptually. One of them, Lars, felt that the cameraman's/interviewer's approach and wording changed with respect to the person he was addressing, for example by assuming a more critical stance when speaking to the person labelled as "the optimist", while having a more positive tone when interacting with "the pessimist". Igor commented positively on the *durational* form of the movie, arguing that it felt uncanny with respect to the editing rhythm and speed of speech that are the standard for social media content ("*It's not as snappy*"). Despite being challenging, the film's rhythm managed to capture his attention, "*and that's quite hard to do!*".

The same person, a particularly savvy film spectator, draw two further links between how the film was structured, how it presented its content, and the topic of the film itself. Firstly, he commented on how the intrusiveness of the camera's optical zoom, rapidly moving closer to a person's face, reminded him of an actual invasion of privacy: "*In the beginning, when you had the first person's interview, the lady, the first one, and basically when she said the word "privacy", the camera sort of went into her face, like invaded her privacy*". However, the effects of this "breach of privacy" were not aimed at the characters on screen, but rather at the spectator

himself: "*I mean, it wasn't [perceived as intrusive] to the person who was into it [= in front of the camera] because you used the zoom instead of a physical camera movement, but it felt intrusive* to me *a little bit while I was watching*". Secondly, when discussing the importance that the film gave to minor details about the characters' speech, attitude, or face, he added: "*By how they look, you [= the spectator] can also collect some data*", thus drawing a comparison between the camera's and the algorithm's potential for "*data collection*".

4.3 Knowledge and algorithms

When discussing the movie, about 50% of participants mentioned their own interaction with algorithms in a self-reflexive manner. People showed willingness in discussing their own habits on social media, and provided some explanations as to why they believe they interact with the algorithm in a certain way. The most prevalent reasoning behind their algorithmic interactions was formulated in similar terms to what Eslami et al. [23] refer to as the "personal engagement theory", that is the belief that the more a user interacts with a certain type of content, the more that content will show up: "*There was a time, a couple of years ago, when I was a bit too much on Instagram and I was getting annoyed at what they showed me. And I started taking a more active approach to only actually click on what I thought was interesting, and not only what drove me to click on it, so that I would get shown things that I would have an actual interest in*", Nico commented. Lars stated that he turned off YouTube's search history in order not to be profiled: "*I have used YouTube a lot, but I switched up the algorithms and now it sucks, because it was too good, like I wasted too much time on it. I think I can turn off history now. It doesn't remember my choices. Basically that means that it doesn't update the belief [= my targeted preferences]*".

The scene titled "Teens with Phones", where teenagers were made to comment and reflect upon their usage of technology, was well received by most viewers. 63% of participants compared their own beliefs about smartphone usage to the theories that were mentioned on screen ("*I have the same beliefs*", Igor). Ana felt surveillance was creepy due to the fact that there is no way of knowing where one's data ends up: "*I don't have my camera recording, but sometimes I think if they are recording a video, there could be this world of black markets kind of selling this recorded video of people*". Lucia demonstrated a good degree of reflection on which kinds of data TikTok collects from her: "*Maybe the greatest source of data among all social medias is taken by TikTok, when it records the time, how long it takes to watch a video, if you press like or you save it, if you do both, if you scroll right away...*". Lastly, Lars commented on how relatable a scene was, where a boy and a girl exchange their phones and comment on each other's YouTube recommended videos, due to how personal that information feels: "*I liked the one [scene] where they were switching [phones with*

each other] because sometimes when you… for example, someone is putting on a song on YouTube or something and you're suddenly, oh, this is a lot of information, when you glance at the screen, about what they like, so suddenly you're like, this is a lot of animal videos, because they really like animals. Stuff like that. I thought that was a clever way of them noticing how personalized it actually was".

Three people made comments on how a character's knowledge on the topic of algorithms and surveillance influenced the character's own behavior in relation to technology. This was coded positively as "knowledge VS behavior". Mette commented how she "*[thinks] it's interesting how the people who might know a bit more about [technology], they choose not to use it that much. [...] People like Mark Zuckerberg, for example, he also doesn't share that much information on his phone*". The difference in knowledge was mostly attributed to age and profession. However, the teenagers were considered quite savvy with respect to how their social medias' algorithms' function: "*So, they are not maybe as knowledgeable as the people who have worked with this, on the side effects that [technology] might have, but they* know *what is happening*" (Helena).

Perhaps surprisingly so, at least three people had a strong negative reaction to the absence of on-screen labels that could show a person's name and professional role. This issue was raised only in regards to the interviews with professors, and, in opposition to the positive code described above, these answers were negatively coded as "profession VS experience". Particularly, this affected viewers by making them doubt the professionalism of the interviewees. In this sense, speaking on behalf of one's own experience, and speaking from a professional perspective were two conflicting positions ("*I felt like there was a lot of that personal touch to it that took a little bit away of [credibility]*", Ana). For Mette, as these characters were shown in a professional setting, they should not have spoken as individuals: "*I think that they [were still meant to] spoke as professionals, if that makes sense. So I think it would have made it more clear to me, I guess, if I knew what their profession was*". Ana stated: "*I take it less seriously if I don't know [the professional context]. It's just a thinking [= They are merely expressing their thoughts]. I tend to take more seriously opinions from people that have more experience*". She went on to question the source of the information she was receiving: "*One thing that I take seriously is also like the sources of data that you take to form your opinion, right? [...] For me, it's hard to see if a person is really knowledgeable on something if they don't show, for example, the two sides of the coin*". Lastly, she added that observing these people's experiences and hearing their opinions was not enough for her to gain any new knowledge: "*One thing that it would be nice, it would be nice to see some statistics, actually. I think I would take the statistics more seriously than the opinion of someone that is maybe a professor*".

4.4 Personal data obfuscation

Participants had a different set of reactions to the topic of personal data obfuscation (PDO), as it was presented in the movie. When questioned about it, most participants answered in terms of critical engagement with the topic of PDO (coded as "critique of PDO"), rather than recollecting their personal experience with similar tactics. The concept of PDO seemed to raise interest, and it encouraged viewers to evaluate its feasibility. When questioned about it, 75% of viewers expressed varying degrees of approval or disapproval of a tactic for privacy enhancement based on PDO. A majority of them were skeptical of its practical feasibility, or of the ethics it implied. Ana stated that "*the idea is not completely, like, decent*", further adding that users have neither the time nor the willingness to engage with a similar tactic: "*We can live our lives without thinking about this. And there's so many other interesting things to think about. I think most people would just ignore this possibility*". Mette remarked the actual need to put truthful information on the internet: "*If I order something, I still have to put in my real address and, I don't know, it seems like a lot of work still, to manipulate it*".

Two people had heard of a similar tactic before. Lars admitted to having download an app, some years ago, that would send out purposefully wrong information with every Google query. However, he deleted the app after a short while, as it was impairing his browsing behavior: "*I just installed the app because I thought, oh, I like my privacy. But then I had to press "I'm not a robot" all the time... I need to google shit!"*. Igor, who had also heard of a similar tactic before, stated that "*it actually seems an enormous effort. And also maybe like a futile effort, unless masses of people are doing it*", further stressing how a collective effort is necessary to limit surveillance. Nico and Lucia proposed a more constructed criticism of PDO. Nico wondered how many users would need to adopt such a strategy before data becomes no longer a valuable metrics for real life phenomenon: "*That made me wonder which percentage of people would you need to do that in order for the data to become useless. I'm sure because even if it's 3%, 4%, 5%, then that introduces such a discrepancy between real life and the data that they collect*". Lastly, Lucia made a connection with another scene of the movie, in order to make her point: "*The interview you made with the professor [appearing somewhere around the middle of the film] contains an answer to this behavior [that is being proposed by scholar discussing PDO towards the end], meaning that creating a fictional character as a solution, is something you cannot do forever, constantly. You cannot spend all of your time trying to deconstruct the filter*".

Differently than what was previously theorised, participants did not seem to have engaged much in tactics of PDO. This prompted the need to ask more precise questions as to understand whether a past action that a participant might recollect could be actually labelled as a PDO tactic. Three participants denied having ever

acted in such a way when online (coded as "no PDO"). When asked about it, they clarified they feel compelled to tell the truth when online. One of them, Ana, after having discussed her criticism of PDO tactics, recognised having put wrong information on websites before, but for very narrow purposes. This prompted the need to acknowledge two new codes to differentiate between degrees and purposes of a PDO tactic.

- The code of **partial obfuscation** refers to a tactic of obfuscation that has a reduced scope, and is often enacted with a specific motive in mind, such as avoiding age limitations, or pulling off a prank. Furthermore, a "partial" tactic of obfuscation also refers to the case wherein a user rejects an algorithm's profiling of himself/herself. However, this rejection is not performed by sending out wrong data, but rather by using standard tools for content curation (i.e., pressing the "Hide" button)[1].

- **Playful obfuscation**, on the other hand, has no clear motive besides that of being performed with an attitude that is both playful and adversarial with respect to the algorithm's profiling function.

Three participants fell under the category of "partial obfuscation": one had created an entirely pretense Facebook profile as a joke; one used to state a wrong age to bypass age restrictions; and the last one, a woman, rejected the algorithm's profiling of herself, as pregnancy-related content was being shown to her more and more frequently, which annoyed her.

Among all test participants, only Lucia could recall an instance in which she performed an act of "playful obfuscation". According to her account, when visiting e-commerce websites that are more often aimed at women rather than men, she is prompted to select her gender. Not wanting to comply with this further stereotyping, she states she is a man: "*Sometimes I'm looking for products, on some websites, maybe stuff I have to buy, household appliances, stuff like that. Sometimes I end up on websites that are asking for information on whether I am a woman, or man. Usually to those type of stuff I lie*".

[1]This case still counts as a technique of "obfuscation", in that *some* wrong data is being sent out. The user is "rebelling" to a highly precise profiling of his or her persona, and therefore he or she sends out the information that that profiling, despite actually being accurate, does not refer to *his or her* profile.

Chapter 5

Discussion

With reference to **Hypothesis 1**, results seem to confirm its validity. Despite the fact that only in about half of the interviews was a self-critical awareness of algorithmic knowledge explicitly stated, this number bears a smaller significance in the context of qualitative analysis. Instead, codes labelled under "viewing experience" and, most importantly, "critical engagement" effectively pointed to ways in which the film was able to raise viewers' considerations on the topic of algorithmic knowledge, sometimes in unexpected ways. Still, in retrospect, Gran et al. [26] research study on algorithmic awareness might have proven useful to formulate a more precise theory on where algorithmic awareness is located, as, during testing, differences emerged between the three contexts that the authors discuss (algorithm-driven recommendations, algorithm-driven advertisements, algorithm-driven content), which were not taken into consideration during the development of the methodology.

Reflecting on how different demographics are proven to have differing degrees of algorithmic awareness, as reported by Gran et al., a further matter for consideration, which this study did not include, could have been the conditions that enable the emergence of algorithmic knowledge in the first place and, consequently, algorithmic awareness. Swart [49], for example, discusses how algorithmic knowledge is predicted by levels of education, and how it has a negative correlation with age. Most importantly, as Cotter and Reisdorf [14] point out, algorithmic *experiencing* itself is *already* dependant on factors such as socioeconomic status: "Those with more resources experience greater opportunities for encountering, attending to, and retaining information more than others". In this sense, this study understood algorithmic experiencing and knowledge as unquestioned concepts that were not marked by a social significance. Thus, the way a communication strategy based on the theory of digital humanism could be able to engage with algorithmic awareness, while at the same time taking into consideration its nature as a product of social forces, is still an open question.

As formulated in **Hypothesis 2**, the film managed to feature and engage constructively with common folk theories that users develop to give reason to their algorithmic knowledge. Participants showed a great amount of interest in the scene in which teenagers were made to discuss their own beliefs on technology through different techniques, such as that of recounting out loud what their algorithmic habits consisted of. As the filming of that particular scene was limited by production constraints, future attempts at deploying similar educational strategies as those that were used in the case of this movie, should give greater relevance and visibility to content that features a self-reflexive engagement with folk theories.

However, the film did not prove successful in its juxtaposition of the algorithmic experience of "educated users", and that of "common users". A tendency among viewers to engage negatively with scenes in which scholars were made to discuss their fields of expertise, while giving relevance to their own subjective experiences (highlighted by the absence of name and profession labels), emerged from the usage of the code "profession VS experience". Such scenes were seen as unhelpful or even confusing by test participants. There are two considerations that can be drawn from this criticism. On the one hand, it is clear that the link between the scholar's "human" experiencing of technology and their professional knowledge was not developed appropriately, so as to convince the spectator of its importance. Partially, this can be attributed to production constraints: further filming of the subjects (i.e. the professors) beyond the context of the mere "long interview" might have proven useful for the spectator to get more acquainted with the character, and acknowledge the value of their "human" side as recipients of scholarly knowledge. On the other hand, however, it becomes clear that, in the context of education, self-awareness of algorithmic knowledge cannot entirely replace the acquisition of new information, as was the case with a test participants lamenting the absence of "statistics", which she needed in order to be properly convinced by what she was being shown. Furthermore, a greater degree of "grounding" of algorithmic knowledge with technical information might be also proven useful for the viewer to bridge his or her knowledge of social media algorithms for content curation, to how algorithms function in a broader sense, and not just in the context of social medias. As such, the following statement by Lomborg and Kapsch [36] could be reversed:

> We suggest that mobilizing data literacy should focus not only on technical knowledge and skills, but also on showcasing real life examples of algorithmic work in different contexts, relatable to the life of ordinary people.

In the context of this research project, it becomes clear that showing how algorithms work in the "life of ordinary people" is not enough. Without additional

technical information, the risk is that of relying on an educational model that understand knowledge as coming solely from *within* the individual, and that is thus radically inbound.

Lastly, with respect to the topic of personal data obfuscation (**Hypothesis 3**), the film seemed to achieve the opposite to what it did with folk theories: the idea of PDO was effectively communicated to viewers, who engaged in criticism of it; but it did not stimulate the desired degree of self-reflection on how they might themselves have put such a tactic into practice. Most importantly, the starting consideration, according to which PDO is a strategy *already embedded* into users' algorithmic habits, who then need to be made self-aware of it, was put into question by test data, wherein only 12.5% of participants (one out of eight) acknowledged using a tactic of PDO. However, this result is not necessarily at odds with existing literature: according to Min's study [42] surveying 3441 technology users, about 13.2% of them fell under the "activist" type:

> The activist group provided answers considered very active algorithmic engagement such as, "I sometimes click and follow things I don't agree with, because I don't like algorithms".

However, the testing of participants, including questions on the obfuscation tactics they might have employed, highlighted a further conceptual problem: if obfuscation techniques are understood as *embedded* in the constant process of negotiation of privacy that users enact habitually, the isolation of a "standalone" instance of obfuscation might be proven difficult or even nonsensical, as such acts cannot be disentangled from all the other acts of compliance, semi-compliance or semi-resistance that users enact in relation to algorithms. This difficulty prompted the necessity to use the codes of "partial obfuscation" and "playful obfuscation", which proved helpful in understanding how different user behaviors might be understood as being part of a scale of differing obfuscation tactics. However, future research might consider how to structure such a division already in its premise.

With respect to a critical engagement with the concept of personal data obfuscation, ideas expressed by some test participants are echoed in the relevant literature. Brunton and Nissenbaum [10] take on the ethical issues related to a strategy of obfuscation, especially in regard to the question of whether it can be thought of as a form of "free riding", a situation in which an actor gains an advantage at the cost of others. Lastly, criticism on how PDO can be thought of as an individual effort, and on what is the role of collective action, can be understood through the words of Draper and Turow [19] dealing with the life of critical theorist Theodor Adorno:

> [Adorno] implied that his decision not to engage directly with powerful cultural institutions was informed in part by his belief that individual action cannot address social problems. These one-off approaches rarely

> result in broad social change not necessarily because they fail to elicit widespread engagement, but because individual responses seldom succeed in undermining powerful systems.

In conclusion, going back to the research question, the film proved successful in communicating the topic of digital surveillance to its audience. However, it was shown how the approach of digital humanism should be coupled with an appropriate technical education in the field of computer sciences in order for it to be effective. Moreover, utilizing digital humanism as a framework to educate on the topic of surveillance is not limited to the production of documentary films. Its focus on user agency, on the production of algorithmic knowledge and folk theories, and the consideration of PDO as a tactic against surveillance, are angles that could find a suitable application in the context of educational courses and workshop aimed at developing young people's digital literacy. As previously stated (see User agency), a common approach in these educational contexts is that of treating problematic technology usage as an "addiction". By setting this consideration against the business model of big tech corporations, it seems that an unfair degree of blame is put on the shoulders of teenagers. Recognising user agency, and discussing individual tactics of resistance might prove to be fruitful ways forward.

Chapter 6

Conclusion

Systems for digital surveillance, datafication and behavior prediction are becoming more and more pervasive. The general public shares feeling of disconnection to the topic of privacy, or even of *resignation* and powerlessness. In this context, this paper proposed a framework for non-tech literate users to approach, understand and, possibly, care about the subject of surveillance. A core tenet of this theory, which was termed "digital humanism", was to give importance to a user's agency in the co-constitution of users and algorithms, in order to engage the user with a self-critical reflection on his or her own algorithmic knowledge, and the folk theories that are believed to guide algorithmic operations.

A documentary film titled "Data Surveillance" was produced as a communication tool to educate the viewer on the topic. The film took on both scholarly and popular debates around the concept of privacy and surveillance. For example, it engaged with the idea that the global infrastructure of "surveillance capitalism" is implementing a system of "behavior modification", which poses serious risks for the autonomy of the human subject. The film also gave relevance to how "algorithmic identities" are formed by algorithms, and how they relate to our own. Starting from these considerations, the film was devised to give primary importance to the human figure, so as to question how do humans experience technology, and what is special about "human nature" that is at stake of being modified. The film included a diverse set of characters having different relationships to technology, so as to encourage the viewer to observe and take a critical stance towards his or her own attitude in relation to it. Lastly, as a strategy against surveillance, the film introduced the idea of "personal data obfuscation" as a way to deceive algorithms by sending out wrong information about oneself.

Test screenings of the film showed that viewers were highly engaged to the topic of digital surveillance, as it touched on their daily lives. Albeit the film

was effective in engaging the viewer in a critical self-reflection of his or her own algorithmic habits, some spectators lamented the absence of a more "authoritative" source of knowledge on the topic, which could provide more technical details (e.g. numbers and statistics) on why surveillance is problematic. A future development of the project might take this criticism into consideration and widen the scope of the movie with different sources of information. Finally, the topic of personal data obfuscation was met with mixed reactions. For most viewers, the idea seemed interesting and thought-provoking, but few of them acknowledged having ever had any similar behavior on the internet. Furthermore, the idea was criticised for being too much of a burden to be put on a user's responsibility. Instead, some viewers emphasised the importance of collective action to tackle the issue of surveillance.

Bibliography

[1] Alejandro Barredo Arrieta et al. "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI". In: *Information fusion* 58 (2020), pp. 82–115.

[2] *Bad Bot Report*. Tech. rep. Imperva, 2022. URL: https://www.imperva.com/resources/resource-library/reports/bad-bot-report/.

[3] Kirstie Ball. "Review of Zuboff's The Age of Surveillance Capitalism". In: *Surveillance & Society* 17.1/2 (2019), pp. 252–256.

[4] Michael S Bernstein et al. "Quantifying the invisible audience in social networks". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2013, pp. 21–30.

[5] Pietro Bianchi. "Filming Capital". In: *The SAGE Handbook of Marxism*. Ed. by Wendy Matsumura. 2022, pp. 1320–1336.

[6] Susan Bidel. *Poor-Quality Ads Cost Marketers $7.4 Billion Last Year*. Tech. rep. Forrester, 2017. URL: https://www.forrester.com/press-newsroom/poor-quality-ads-cost-marketers-7-4-billion-last-year/.

[7] Luther Blissett. *Mind Invaders. Come fottere i media: Manuale di guerriglia e sabotaggio culturale*. Castelvecchi, 2000.

[8] James Bridle. "Machine Learning in Practice". In: *Medium* (2017). URL: https://medium.com/intersections-arts-and-digital-culture-in-the-uk/james-bridle-machine-learning-in-practice-d7cb58cd20cb.

[9] Finn Brunton and Helen Nissenbaum. *Obfuscation: A user's guide for privacy and protest*. Mit Press, 2015.

[10] Finn Brunton and Helen Nissenbaum. "Political and ethical perspectives on data obfuscation". In: *Privacy, due process and the computational turn*. Routledge, 2013, pp. 185–209.

[11] Michel de Certeau. *The Practice of Everyday Life*. Berkeley: University of California Press, 1984.

[12] John Cheney-Lippold. *We Are Data: Algorithms and the Making of Our Digital Selves*. NYU Press, 2017.

[13] Christine Cornea. “Introduction: Interviews in film and television studies”. In: *Cinema journal* 47.2 (2008), pp. 117–123.

[14] Kelley Cotter and Bianca C Reisdorf. “Algorithmic knowledge gaps: A new horizon of (digital) inequality”. In: *International Journal of Communication* 14 (2020), p. 21.

[15] Nick Couldry and Ulises A Mejias. *The Costs of Connection: How Data is Colonizing Human Life and Appropriating it for Capitalism*. Stanford University Press, 2020.

[16] Kate Crawford and Vladan Joler. “Anatomy of an AI System”. In: *Retrieved September* 18 (2018), p. 2018.

[17] Marco Deseriis. “Lots of money because I am many: The Luther Blissett project and the multiple-use name strategy”. In: *Cultural Activism*. Brill, 2011, pp. 65–93.

[18] Nicholas Diakopoulos. “Algorithmic accountability: Journalistic investigation of computational power structures”. In: *Digital journalism* 3.3 (2015), pp. 398–415.

[19] Nora A Draper and Joseph Turow. “The corporate cultivation of digital resignation”. In: *New media & society* 21.8 (2019), pp. 1824–1839.

[20] Francesco D’Abbraccio and Andrea Facchetti. “Introduzione”. In: *AI & Conflicts. Volume 1*. Ed. by Francesco D’Abbraccio and Andrea Facchetti. Krisis Publishing, 2021.

[21] Catherine D’Ignazio and Rahul Bhargava. “Approaches to building big data literacy”. In: *Bloomberg data for good exchange* (2015).

[22] Motahhare Eslami et al. “" I always assumed that I wasn’t really that close to [her]" Reasoning about Invisible Algorithms in News Feeds”. In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2015, pp. 153–162.

[23] Motahhare Eslami et al. “First I" like" it, then I hide it: Folk Theories of Social Feeds”. In: *Proceedings of the 2016 cHI conference on human factors in computing systems*. 2016, pp. 2371–2382.

[24] Rafael Evangelista. “Review of Zuboff’s The age of surveillance capitalism”. In: *surveillance & society* 17.1/2 (2019), pp. 246–251.

[25] Gloria González Fuster. “Inaccuracy as a privacy-enhancing tool”. In: *Ethics and information technology* 12 (2010), pp. 87–95.

[26] Anne-Britt Gran, Peter Booth, and Taina Bucher. “To be or not to be algorithm aware: a question of a new digital divide?” In: *Information, Communication & Society* 24.12 (2021), pp. 1779–1796.

[27] Kevin Hamilton et al. “A path to understanding the effects of algorithm awareness”. In: *CHI'14 extended abstracts on human factors in computing systems*. 2014, pp. 631–642.

[28] Eszter Hargittai et al. “Black box measures? How to study people's algorithm skills”. In: *Information, Communication & Society* 23.5 (2020), pp. 764–775.

[29] M. Hennink, I. Hutter, and A. Bailey. *Qualitative Research Methods*. SAGE Publications, 2020. ISBN: 9781473944251. URL: https://books.google.dk/books?id=_InCDwAAQBAJ.

[30] Jeroen van den Hoven et al. “Privacy and Information Technology”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2020. Metaphysics Research Lab, Stanford University, 2020. URL: https://plato.stanford.edu/archives/sum2020/entries/it-privacy/ (visited on 05/11/2023).

[31] Tim Hwang. *Subprime Attention Crisis: Advertising and the Time Bomb at the Heart of the Internet*. FSG originals, 2020.

[32] Rob Kitchin. “Thinking critically about and researching algorithms”. In: *Information, communication & society* 20.1 (2017), pp. 14–29.

[33] Spyros Kokolakis. “Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon”. In: *Computers & security* 64 (2017), pp. 122–134.

[34] Yevhen Laniuk. “The Goliath with the Microchip: Interpreting Shoshana Zuboff”. In: *23rd International Conference on "Ethical Thinking - Past & Present" (ETPP 2021/23)*. Ed. by V. Gluchman. FF PU, 2021, pp. 174–191.

[35] Sarah Lewis. “Qualitative inquiry and research design: Choosing among five approaches”. In: *Health promotion practice* 16.4 (2015), pp. 473–475.

[36] Stine Lomborg and Patrick Heiberg Kapsch. “Decoding algorithms”. In: *Media, Culture & Society* 42.5 (2020), pp. 745–761.

[37] Jens-Erik Mai. “4 Situating Personal Information: Privacy in the Algorithmic Age”. In: *Human rights in the age of platforms*. Ed. by Rikke F. Jørgensen. Knowledge Unlatched MA, 2019.

[38] Jens-Erik Mai. “Big data privacy: The datafication of personal information”. In: *The Information Society* 32.3 (2016), pp. 192–199.

[39] Lev Manovich. *Defining AI Arts: Three Proposals*. 2019. URL: http://manovich.net/index.php/projects/defining-ai-arts-three-proposals.

[40] Michael Mercier. *Broad Solutions to Smartphone Addiction*. Tech. rep. Screen Education, 2021. URL: https://www.screeneducation.org/broad-solutions-to-smartphone-addiction.html.

[41] Sandro Mezzadra and Brett Neilson. "On the multiple frontiers of extraction: Excavating contemporary capitalism". In: *Cultural studies* 31.2-3 (2017), pp. 185–204.

[42] Seong Jae Min. "From algorithmic disengagement to algorithmic activism: Charting social media users' responses to news filtering algorithms". In: *Telematics and Informatics* 43 (2019), p. 101251.

[43] Emily van der Nagel. "'Networks that work too well': intervening in algorithmic connections". In: *Media International Australia* 168.1 (2018), pp. 81–92.

[44] Helen Nissenbaum. "Privacy as contextual integrity". In: *Wash. L. Rev.* 79 (2004), p. 119.

[45] Elleen Pan et al. "Panoptispy: Characterizing audio and video exfiltration from android applications." In: *Proc. Priv. Enhancing Technol.* 2018.4 (2018), pp. 33–50.

[46] Luci Pangrazio and Neil Selwyn. ""It's not like it's life or death or whatever": Young people's understandings of social media data". In: *Social Media+ Society* 4.3 (2018), p. 2056305118787808.

[47] Frank Pasquale. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, 2015.

[48] Jonathan Rosenbaum. *Afterword to THE BIG BRASS RING, A Screenplay by Orson Welles (with Oja Kodar)*. Blog. 2021. URL: https://jonathanrosenbaum.net/2021/12/afterword-to-the-big-brass-ring-a-screenplay-by-orson-welles-with-oja-kodar/.

[49] Joëlle Swart. "Experiencing algorithms: How young people understand, feel about, and engage with algorithmic news selection on social media". In: *Social media+ society* 7.2 (2021).

[50] "Tesla, Apple, Google, Microsoft dodge Congo cobalt class-action". In: *Mining.com* (2021). URL: https://www.mining.com/tesla-apple-google-microsoft-dodge-congo-cobalt-class-action/.

[51] Sabine Trepte et al. "Do people know about privacy and data protection strategies? Towards the "Online Privacy Literacy Scale"(OPLIS)". In: *Reforming European data protection law* (2015), pp. 333–365.

[52] Tiina Tuominen. *The Art of Accidental Reading and Incidental Listening. An empirical study on the viewing of subtitled films*. Tampere University Press, 2013.

[53] Julia Velkova and Anne Kaun. "Algorithmic resistance: Media practices and the politics of repair". In: *Information, Communication & Society* 24.4 (2021), pp. 523–540.

[54] Shoshana Zuboff. *Surveillance capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs, 2019.

[55] Shoshana Zuboff. "The Secrets of Surveillance Capitalism". In: *Frankfurter Allgemeine Zeitung* (2016). URL: https://www.faz.net/aktuell/feuilleton/debatten/the-digital-debate/shoshana-zuboff-secrets-of-surveillance-capitalism-14103616.html.

Appendix A

Appendix A - Codebook

CODEBOOK	
VIEWING EXPERIENCE	
Effectiveness	*Definition*
	Viewer describes the movie in terms of how *effectful* it was, or it was not.
	Examples
	"I think actually it was a good effect."
Focus on human	*Definition*
	The spectator's attention during the movie has shifted towards "human" details of the person being interviewed (e.g., facial expression, accent). Test subject might make a broader consideration on human nature. This highlights the engagement the spectator had with the interviewee beyond the mere information that was being provided.
	Examples
	"I noticed everyone had a different accent." "There were parts when I started focusing very much on their faces [...] That made me feel like I related a bit more to them."

Personal	*Definition*
	The viewer has perceived the "long interview" format as being more personal in nature.
	Examples
	"It felt like the interviewed person was talking to me." "It felt like a conversation." "It felt like getting to know them."
Relatable	*Definition*
	Code refers to situations in which test subjects express the fact that they related greatly to the characters; to the topics that were shown in the movie in a general sense; or simply to the movie itself. No mention of how folk theories discussed in the movie relate to one's own.
	Examples
	"I think it's like thoughts that many people think about. Like for example, I think many can relate to ads popping up right after you've talked about a certain topic or something." "It's our daily life." Words used: a lot; something that many people relate to/think about; very interesting; all of us (do this); I often think; I liked ... it reminded me of (life situation/acquaintance).

Variety	*Definition*
	The code refers to how the speaker is recognising and/or valuing the variety of people and experiences presented in the movie. This is important, as it highlights that the characters in the film were not taken simply as interviewees, but as people that had different lives and experiences.
	Examples
	"You spoke with interesting people, like, a lot of different age groups."
No engagement with question	*Definition*
	The test subject expressed little or no engagement **to what the interviewer suggested** in the question on how was the viewing experience (and not to the viewing experience itself).
	Examples
	I guess; a bit; I don't know.

KNOWLEDGE AND ALGORITHMS	
Algorithmic habits	*Definition*
	The test subject self-reflects on his or her own way of utilizing the platform to achieve a certain goal (e.g. only liking a certain content etc). The test subject might or might not provide a reasoning (folk theory) as to why the algorithm reacts in this way. This code does not refer to the situation in which a participant discusses his or her own use of social medias in different contexts (e.g. "I scroll Instagram while I take a bath").
	Examples
	"I started taking a more active approach to only actually click on what I thought was interesting and not only what drove me to click on it" "I hide stuff..." "I unfollow..."
Folk theory	*Definition*
	The viewer shows a great degree of engagement with the part of the movie where folk theories are discussed (i.e. "Teens with Phones" section) and self reflects on his or her own **beliefs**. Test subject does not just explain his or her behavior (*Algorithmic habits*), but reflects upon it and gives an explanation, a ratio, through engagement with the film. A participant might also just express a belief about the way algorithms works which is not supported by evidence
	Examples
	"I try never to press on ads on social media or anything like that **because** I don't want it to control what I see."

Knowledge vs behaviour	*Definition*
	The code points to a reflection the participant made on how people with different knowledge about technology interact with it differently. This is important as it shows that a reflection has taken place on the tight link between knowledge and the way we use technology. Knowledge can not only refer to professional knowledge, but also folk knowledge. People that know a lot about social medias might decide not to use them, for example. There **must** be a connection between knowledge and behavior, a reflection on knowledge alone is not enough.
	Examples
	"So I think it's interesting how the people who might know a bit more about it, they choose not to use it that much." "I also think that it's interesting that people like Mark Zuckerberg, for example, that he also doesn't share that much information on his phone."

Profession vs experience	*Definition*
	Some participants raised the issue that the filmed interviews did not have labels with name and profession. Some people found this choice interesting, for others this made the film more difficult to follow. In this last case, the goal of transmitting a scholar's opinion as his or her own personal experience informed by his or her profession, did not succeed entirely. Instead, the viewer was caught in between a limbo: is this person a random guy? Is he or she an expert? In extreme cases, the viewer doubted the knowledge of the professional completely.
	Examples
	"I think that they still spoke as, like, professionals, if that makes sense. So I think it would have made it more clear to me, I guess, if I knew what their profession was." "I did notice that you didn't put a name to the people, which I thought was interesting and that you didn't even really tell us exactly what their role were and why they were speaking. I get that he's a professor, but you didn't really say professor of what?"
Previous knowledge	*Definition*
	The code refers to a previous source of knowledge (talking with friends, reading about it etc.) the test subject is telling about.
	Examples
	"It is a topic we discuss at lunch at work."

PERSONAL DATA OBFUSCATION	
No obfuscation	*Definition*
	The person acknowledges what personal data obfuscation is, but he or she has never tried it. Might add that he/she feels compelled to tell the truth.
	Examples
	"That is not something I really do."
Partial obfuscation	*Definition*
	The test subject recounts past instances in which he or she put into practice an obfuscation technique (i.e. he or she lied on the internet). This can also be just one, simple instance. Partial obfuscation is reduced in scope, and often has a clear motive (e.g. avoiding age limitations, trolling). Partial obfuscation also refers to the situation in which the algorithm is not deceived with false data, but its own functions for content curation (for example the "Hide content" button) are used to limit highly targeted content. This case also counts as an adversarial tactic because a user is resisting the implication according to which "a content is very relevant to you" -> "therefor you must be enjoying it". What the user is saying instead is "a content is very relevant to me" -> "therefore I hide it, because I find this relevance creepy"
	Examples
	"I have tried not to press on the information that was given to me."

Playful obfuscation	*Definition*
	The test subject recounts past instances in which he or she put into practice an obfuscation technique (i.e. he or she lied on the internet). This can also be just one, simple instance. Playful obfuscation has no clear motive but that of refusing to comply with the algorithm, feeding it with wrong information with the goal of deception. It can be enacted by the user without a clear idea of why he/she is doing it.
	Examples
	"In this situation, I lie."
Critique of PDO	*Definition*
	The test subject raises questions on the feasibility of personal data obfuscation as a individual and political project of privacy enhancement. It denotes a critical engagement of the test subject with the implications of obfuscation.
	Examples
	“It seems like a lot of work.”
CRITICAL ENGAGEMENT	
Connections	*Definition*
	The test subject critically draws connection between different parts of the movie. This might mean making a link between what two different interviewees said, or how the form of the film changed between two or more interviews, or how different scenes related to each other as organic parts of the movie.
	Examples
	"In the interview with the optimist, you seemed to act like you were the pessimist. And with the pessimist, you seemed like the one who is an optimist."

Timing	*Definition*
	Participant makes a critical consideration on the *durational form* of the movie, i.e. **not** how he/she engaged with time ("The film felt slow/rushed"), but rather how timing itself was performed, which function it served, and how it related to the film's main topic.
	Examples
	"It's not like the standard nowadays, on social media, timing wise. It's not as snappy."
Privacy metaphor	*Definition*
	Test participant might use certain **formal aspects** of the movie (e.g. zoom) as being a visual example of how privacy and/or surveillance function.
	Examples
	"When the person said the word "privacy", the camera went into her face, as if it was sort of invading her privacy."
Profiling camera	*Definition*
	Code refers to how viewer's attention was caught by the way the camera and the editing allow him/her to focus on certain details of the characters being interviewed. These "capturing" of details is similar to the one made by profiling algorithms. Therefore, the camera was perceive as an instrument of "data collection", thus drawing a comparison between film and algorithms.
	Examples
	"By how they look, you can sort of collect some *data*."

Code not assigned	**Instance**
	• The test subject repeats a viewpoint he or she has expressed previously without any addition in content or wording. • Participant mentions the specific situation or use he or she makes of a social media ("I use Facebook only to catch up on concerts"), without mentioning habits related to interaction with algorithms. • Test subject gives his/her opinion about how the movie should have been; or criticises aspects of it that were due to its production ("There were no shots of people walking"). This is not related to how the form of the movie impacted the viewer's engagement with core topics of the movie (such is the case, instead, when the lack of profession label is criticised). • Descriptive attitude: participant gives description of formal aspects of the movie related to shooting/editing (as he/she thinks interviewer is expecting) without telling how he/she related to it. • The test subject misinterprets the meaning of the question and talks about something else entirely. • Generic statements ("The film was good"). Compliments. • Test subject wanders off topic and discusses other issues related to technology, but not taken on in the movie (AI, attention deficit etc.).

Appendix B

Appendix B - Semi-structured interview model

1) Did you like the movie?
Do you have any comments about it?

2) ...I'm sure you noticed that this movie is made in quite an unusual way in that it has these very long interviews with very few cuts, people talking a lot all the time. This was done to explain the topic of digital surveillance. How did that feel to you?

3) With respect to the scene in which teenagers discuss about their beliefs and exchange their phones... I think it's interesting to raise the point that what they are making are, to some extent, quite trivial statements. Like, for example: "When I want to see more of some content, then I press 'like'". It's very trivial, but at the same time not so trivial, maybe? And the fact that it's shown in the movie gives to it a certain importance. Did that make you think in some way about your beliefs about what happens when you use your phone/social medias?

4) This movie is also a lot about faces. You get a lot of "big faces" of people while they speak. Did it happen to you during the movie, that maybe you were shifting your focus from what the person was saying to the actual person itself? Maybe they were makeing weird facial expressions or maybe they had a weird accent? Did you consider that?
If yes, can you recollect when you noticed that?

5) Here and there in the movie there are some bits of information that these people give out about themselves. Like, we see a person that is dressed

in a very professional way, and then he says ‘‘You know, I’m a 40 years old boring professor...’’. So, in a way, these details we catch a glimpse of tell something about people themselves. And these details are the same that then algorithms use. Did that make you think about something? Did these detials made you feel like you were getting to know a character more?

6) About the last person that was interviewed, the one who sat next to the water, he had an idea about "playing" with algorithms. What did you think about that? Can you remember any time you have performed anything similar?

7) Was there a scene or moment that you remember particularly? Anything that raised a thought in you?

Appendix C

Appendix C - Consent form

Consent text - Interview

This is a request for your consent to process your personal data. The purpose of the processing is for academic research.

You consent to the processing of the following data about you: standard demographics information (age, gender, profession); content of interview with researcher regarding the experience of watching the documentary film ‘‘Data Surveillance’’.

I, Arturo Fabbro, is the data controller of your data.

Your data will be stored securely, and I will solely use the data for the above purpose.

You always have the right to change your consent. If you wish to change your consent later on, you can send an email to xxxxxxxxxxx@xxxxxxx, stating as the email’s subject: ‘‘Retraction of consent on personal data’’.

The General Data Protection Regulation entitles you to obtain information that you can request by reaching out to the above stated email address.

I hereby consent to Arturo Fabbro processing my data in accordance with the above purpose and information.

Date:
Name:
Signature

How I process your data

The data controller
Arturo Fabbro
Xxxxxxxx XXX
XXXX XXXXXXXX
XXXXXXX

The purpose of processing your data
The data processing will serve the purpose of a research project on the topic of digital surveillance. The project is carried out within the institution of Aalborg University.

I process the following personal data:
General personal data (see Article 6(1) (a)):
- name;
- age;
- profession;
- information on technology usage.

Sensitive personal data (see Article 9(2) (a)):
- content of structured interview.

How I store your data
I will store your personal data for as long as necessary for the data processing purpose for which I are obtaining your consent and in accordance with the applicable legislation. I will then erase your personal data with the exception of the video and audio recording.

Your rights

When I process your personal data, you have several rights under the General Data Protection Regulation. For example, you have a right to erasure and a right to data portability.

In certain cases, you have a right of access, a right to rectification, a right to restriction of processing and a right to object to our processing of the personal data in question.

Be aware that you cannot withdraw your consent with retroactive effect.

Do you want to complain?

If you believe that I do not meet my responsibility or that I do not process your data according to the rules, you may lodge a complaint with the Danish Data Protection Agency at dt@datatilsynet.dk.

However, I encourage you also to contact us, as I want to do me utmost to accommodate your complaint.

Disclosure to and from third parties

Your data (or parts of your data) may be transferred to Aalborg University.