The Political Micro-Targeting Threats to Democracy: *Building Legal Solutions to Preserve Fair Elections*

Juliana Augustinis Valle Machado Silva

Master's thesis submitted in partial fulfilment of the requirements for the Degree of Master of Laws

in

International Technology Law

Student number: 2658229
Submission date: July 15th 2020
Supervisor: prof. dr. Gareth T. Davies
Second assessor: Magdalena Jozwiak

Acknowledgements

I heartily thank my supervisor, prof. Gareth Davies, for embracing this challenging project and supporting me in every step of it - even in the out of ordinary circumstances of a global pandemic. As the in-person academic activities were momentarily paralyzed by the coronavirus, the warmness of the social contact was lost for a while. Still, Gareth made online supervision meetings possible, always helping to push the research further, and offering bright contributions.

Special thanks for my dear Kazu, for listening to my thoughts over and over, and sometimes engaging in discussions about possible solutions for election processes distorted by unethical uses of technology. I could not have a more dedicated, fond, and incredibly patient (I highlight) partner in my journey.

More than ever, I thank my parents and grandparents, eternal supporters in any aspect of my life. These times of closed borders made me even more grateful for the moments we can spend together. Soon the world will overcome these hard times and I will be able to hold you again.

Lastly and not least important, I would like to thank the Vrije Universiteit Amsterdam for offering me the conditions to do the fascinating master's in International Technology Law; all its professors and staff, for creating a stimulating academic environment which inspires me to pursue an academic career.

Table of Contents

‘We cannot evaluate the current trajectory of information civilization without a clear appreciation that technology is not and never can be a thing in itself, isolated from economics and society’. Shoshana Zuboff, The Age of Surveillance Capitalism. See (Zuboff, 2019).

‘When change is easy, the need for it cannot be foreseen; when the need for change is apparent, change has become expensive, difficult and time consuming’. David Collingridge, The Social Control of Technology. See (Collingridge, 1980).

List of Abbreviations

Ad	Advertisement
App	Application (computer program designed for a particular purpose)
CA	Cambridge Analytica
CFR	Charter of Fundamental Rights of the European Union
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EDPB	European Data Protection Board
EU	European Union
GDPR	General Data Protection Regulation
UK	The United Kingdom
US	The United States of America

Abstract

Political micro-targeting through big data analysis poses risks to modern democracies. This practice has been observed in several recent election campaigns worldwide. The 2016 Cambridge Analytica scandal exposed how millions of electors were targeted with political advertisement personalized through the algorithmic analysis of personal data. These specifically targeted messages can generate emotional responses and influence voters' behavior. Democracy relies, in its turn, on the assumption that people can make authentic decisions and are not inadvertently manipulated by particular groups. The nature of EU data protection regulations, centered on privacy protection, offers limited precaution against the political abuses performed in the online context using micro-targeting. This thesis claims that the prevention of political manipulation and distorted election processes requires the creation of regulatory norms dedicated to this matter. It proposes a potentially more effective transparency-based prevention mechanism to be mandatory in political online advertising. The formulated solutions encompass adding information about micro-targeting directly in the targeted ads and making targeted messages available to other groups of interest. Individuals should also have the option to be exposed to targeted political ads, which could be given through a reinforced consent. A prohibition of audience targeting is also evaluated. The Thesis outlines the implementation of a framework of novel legislative approaches discussing their advantages and setbacks in order to minimize the effects of political micro-targeting on electoral processes.

1. Introduction

1.1 Subject and problem definition

The European Commission acknowledges the phenomenon of political micro-targeting as one of the major contemporary challenges for European democracies (Nenadić, 2019). In a general sense, the practice involves the targeting of voters with political advertisements personalized through an algorithmic analysis of personal data. Boosted by the fast advance of computational technology, especially big data analytics, the collection of individual's data enables groups of interest to establish personality profiles of electors with increasing levels of sophistication. By identifying which groups are most likely susceptible to certain content, political actors can tailor their political ads in a manipulative manner to trigger individuals' emotional response. Simply put, some political agents use technological advancements to identify people's vulnerabilities, customize campaign advertising, and leverage political campaigns. Because of this aggressive technique of political propaganda, opinions can be subconsciously influenced towards or against a candidate or political party; voters can be mobilized or demobilized (Witzleb, Paterson, & Richardson, 2020).

This thesis assumes that the manipulative use of political micro-targeting can not only undermine individual political freedom of choice but also the human autonomy by swaying voters' political will on a 'micro' subconscious level. In addition, the practice endangers the right of voters to receive complete and fair information from all political parties and candidates. Under these perspectives, political micro-targeting deprives citizens of the right of making free and informed political decisions (Bayer, et al., 2019), distorting the democratic debate. It jeopardizes, furthermore, the fairness of elections by giving an 'undesirable advantage' to 'financially powerful groups' (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

Following the premise that elections should be fair and based on authentic political opinions, Chapter 2 starts from the analysis of the Cambridge Analytica data scandal to define the phenomenon of political micro-targeting. Furthermore, it examines in which sense this political tech-driven strategy threatens fundamental rights and democratic regimes. Having established these risks, Chapter 3 investigates how the current European Union (EU) legislative framework applies to micro-targeting and to what extent it is suitable to prevent the identified damages to democracy, considering it focuses on data protection and privacy aspects. From the analysis of the existing regulation and its shortcomings, Chapter 4 is dedicated to proposing different legislative tools and interpretative approaches to minimize or prevent the anti-democratic threats in micro-targeting, critically discussing the advantages and weaknesses of each perspective.

1.2 Research questions

We can phrase the central legal question of the study as ***what regulatory perspectives can be envisioned, in the European context, for the anti-democratic threats caused by political micro-targeting***?

This problem definition unfolds in the subsidiary questions listed below. Each sub-question will be answered, respectively, in one of the thesis chapters.

1. How can political micro-targeting be conceptualized from recent political data-scandals? What dangers does this practice present to human dignity, fundamental rights, and democracy?

2. What is the legal framework applicable in Europe to online political micro-targeting? Is it adequate to prevent political manipulation and ensure human autonomy and fair elections? What gaps can be identified in the current data protection norms for regulating political micro-targeting?

3. What regulatory approaches can be envisaged to prevent the democratic damages caused by online political micro-targeting and fill current legislative gaps? What would be the pros and cons of these proposals?

1.3 Research methods

The thesis performs a dogmatic analysis of socio-legal academic publications, legal norms, and European case-law to answer the proposed research questions, formulated on a theoretical level. Considering the multidisciplinary character of the technology law research area, it also uses academic contributions from other fields of knowledge such as political science and communication, sociology, and behavioral science.

In Chapter 2, as to conceptualize the socio-political and tech-oriented phenomenon of political micro-targeting, news articles regarding the Cambridge Analytica files, as well as scientific articles published in academic journals inside and outside the legal field are critically studied. To investigate the impacts of the described object of study for human dignity, fundamental rights, and democracy, Chapter 2 examines legal doctrine and articles from social sciences.

To answer the second sub-question, regarding the existing norms applicable to micro-targeting in Europe and regulatory gaps, Chapter 3 critically investigates the application of pertinent provisions of the European data protection framework to the case of political micro-targeting.

Chapter 4 examines the viability, benefits, and disadvantages of new legal perspectives to the established issue. Departing from the gaps in the existing EU data protection legislation, it builds upon socio-legal scholarship and in the case-law of European Courts on human rights and political micro-targeting.

1.4 Contribution to the field of technology law

The performed research contributes to the field of Technology Law by proposing new regulatory approaches for political micro-targeting that focus on the pressing democratic issues it poses. Although socio-legal scholarship addresses the topic especially since 2016, when the Cambridge Analytica case became public, this Thesis has identified a literature gap on legislative perspectives that emphasize the political risks raised to democratic regimes. These issues cannot be appropriately solved only through existing European privacy and data protection norms. Scholars point out that new regulatory approaches seem to be necessary to prevent the abuse of voter's choices and ensure the fairness of elections and democracy and 'more debate and research are needed on what lawmakers should do' (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019). European reports also recommend 'some regulation of micro-targeting for the purposes of protecting democratic public discourse, the fairness of elections and protection of personal data' (Bayer, et al., 2019). The approaches presented in this thesis serve, therefore, as a starting point for the development of new legislation and policies that prevent distortions in democratic disputes. Further legal-doctrinal, empirical, and experimental investigation can also assess the effectiveness, implementation, and enforcement of these proposed legal perspectives.

2. Political Micro-Targeting in Data-Driven Democracies

This chapter conceptualizes the digital phenomenon of political micro-targeting departing from the Cambridge Analytica (CA) case. Subsequently, it examines the democratic issues posed by this transgressive tactic of political manipulation that conjugates data analytics tools and behavioral psychology. The technique of algorithmic profiling used for manipulative purposes caught the world's attention since 2018 with the Cambridge Analytica data-breach scandal. On that occasion, whistleblowers and documents revealed that the 'strategic communication' company (Barry, 2018) misappropriated personal data of millions of social media users for promoting targeted political advertising according to psychological profiles. From this political data incident, Section 2.1 establishes the phenomenon of algorithmic political manipulation, or political micro-targeting, outlining the technique of psychographic segmentation used in data-driven political campaigns. Section 2.2 addresses human rights and democratic risks that emerge from online micro-targeting. Firstly, assessing in what sense this phenomenon differs from 'offline' political manipulative strategies. Secondly, it assesses the special damages caused by false targeted political messages, establishing how political micro-targeting represents a threat to individuals' autonomy in democratic societies and, ultimately, to human dignity.

2.1 Algorithmic profiling and manipulative political strategies

In the first months of 2018, a complex political-tech plot broke-out in the UK and US news. The facts comprehended a massive clandestine collection of personal data by a private company, and the use of computational technology to disrupt hugely influential political campaigns. In a nutshell, the political scandal involved the improper harvest of data of roughly 87 million[1] Facebook users by the data analytics and political consulting firm Cambridge Analytica (Kang & Frenkel, 2018). The aim: to develop individual 'psychographic profiles' (Meredith, 2018) and target voters with personalized messages during political campaigns. Legal scholars refer to the phenomenon as online political micro-targeting (Zuiderveen Borgesius, et al., 2018). 'The Guardian' reports from 2015 had already pointed to the use of Facebook data in the campaign of north American

[1] The first estimates indicated 50 million. After Mark Zuckerberg's declarations in a media conference call, the figure has risen to 87 million. See Ingram, David (2018, April 4). *Facebook says data leak hits 87 million users, widening privacy scandal*. Reuters. Retrieved from https://www.reuters.com/article/us-facebook-privacy/facebook-says-data-leak-hits-87-million-users-widening-privacy-scandal-idUSKCN1HB2CM. (Accessed 06 March 2020).

conservative Senator Ted Cruz (Davies, 2015). Nevertheless, the alleged connection of the 2018 news with the 2016 Brexit referendum and the 2016 US presidential elections was responsible for scaling up the political incident (Privacy International, 2019).

Close to two years after the CA news outbreak, further evidence emphasized the cross-border dimensions of the practice of data-driven political targeting and its potential damages to fundamental rights and democratic orders. In this regard, document leakages by a main whistleblower in the case encompassed documentary evidence of the firm's activities in elections over 68 countries, such as Brazil, Malaysia, and Kenya (Democracy Now, 2020). It is noteworthy that, even though most of the company's clients are from the South Hemisphere, CA influence in manipulating electoral processes in developing nations has been under-reported (Global Voices, 2019). In this context, Damian Collins, member of the Britain Parliament, and chair of its Digital, Culture, Media, and Sport Committee has highlighted the 'direct links between the political movements behind Brexit and Trump' (Mayer, 2018), and pointed to the coordination 'across national borders by very wealthy people in a way we haven't seen before' (Mayer, 2018).[2]

The expression '**algorithmic political manipulation**' comprises, in this vein, the core elements of the scenario under examination: the use of algorithmic analysis techniques for nudging human political consciousness and behavior (Absattarov, 2012) in a manipulative way, according to hidden political interests. Considering the terminology most employed by legal doctrine, the thesis will use the terms 'algorithmic political manipulation', 'online political micro-targeting', or simply 'micro-targeting' interchangeably to refer to the phenomenon.

Regarding the algorithmic analysis, CA made use of the method known as 'psychographic segmentation'. Through this method, the firm managed to establish detailed psychological profiles of internet users from their Facebook interactions alone (BBC News, 2018). In general terms, the technique involves the use of algorithms for establishing a correlation between a vast amount of personal data – for instance, Facebook 'likes', status posts, content shared, messages, and photos published – and personality traits. From this 'psychographic' analysis, internet users can be categorized, 'segmented', into personality profiles (BBC News, 2018). In the same way commercial businesses collect behavioral data of customers to segment them into groups of preferences through algorithmic analysis and target them with personalized ads to increase their sales. This same logic is now being applied for electoral purposes. Scholars in the field note that 'the objective of micro-targeting can be manifold: to

[2] As insightfully brought by Cees Plaizier in his Master Thesis, see (Plaizier, 2018), we can find in the 'Facebook Business' webpage the following declaration from Craig Elder, Digital Director from The Conservative Party in the UK, regarding the 2015 UK general election: '*The level of targeting we had available to us on Facebook—coupled with the research and data we produced internally—meant that we can say for the first time in a UK election that digital made a demonstrable difference to the final election result*'. See (Facebook, n.d.)

persuade, inform, or mobilise, or rather to dissuade, confuse or demobilise voters' (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

The data-company carried out this strategy initially by applying the so-called 'Big-Five' personality test, long used by psychologists, to some thousands of people (BBC News, 2018). According to the difficulty of accessing each group of respondents, CA would reward from 2 to 4 USD for filling the survey (Hern, 2018) in the application (app) called 'This Is Your Digital Life' (Meredith, 2018). To receive the payment, each respondent would necessarily have to agree on sharing their Facebook data with the app (Hern, 2018). The incident becomes even more intricate since the respondents also had to agree to share the data of their Facebook friends, that is, persons that had not even used the app (Chappell, 2018). The social media platform's policy in force at that moment would allow the sharing of friends' data with app developers to 'improve user experience in the app' (Cadwalladr & Graham-Harrison, 2018). Finally, the app transferred all the harvested data to a third party, CA, in this case violating Facebook guidelines (Rehman, 2019). Regardless of the existing Facebook terms and conditions, documents further revealed that the platform was aware by late 2015 of the transfer of users' data to Cambridge Analytica for commercial purposes. Nonetheless, it has failed to timely inform the affected data subjects or to engage in effective efforts to recover their data (Privacy International, 2019).[3]

It should be noted, moreover, that the existence of contractual terms is not lawful when these infringe higher legal norms. This Thesis claims that the Facebook policy allowing the sharing of the friend's data without their explicit consent was not compliant with European data protection norms. According to Guidelines of the European Data Protection Board (EDPB) on consent under the General Data Protection Regulation (GDPR),[4] consent can only be a lawful basis for the use of personal data when a data subject is offered a ***genuine*** choice to accept the terms or to decline them without significant negative consequences (European Data Protection Board, 2020). For meeting this condition, information necessary for consenting must be distinguishable from other contractual matters, be easily intelligible and accessible (European Data Protection Board, 2020). In the EDPB unambiguous wording: 'This requirement essentially means that information relevant for making informed decisions on whether to consent may not be hidden in general 'terms and conditions'. (European Data Protection Board, 2020). The Article 29 Working Party opinions on consent, which remain relevant despite being updated by the EDPB guidelines, point out that consent is presumed not to have been freely given whether it is bundled up as a non-negotiable part of terms and conditions (Article 29 Working Party, 2017). If consent is not obtained in full compliance with the GDPR, data subjects' control over personal data is illusory. It is not considered, hence,

[3] The case resulted in investigations in the US, UK, and Brazil, resulting in settlements and fines over $5 billion for improper sharing of user's data. See (Holt, 2019). In the UK, the Information Commissioner's Office concluded that the platform failed to comply with the UK data protection principles covering the lawful processing of data and data security. See (Information Commissioner's Office UK, 2019).

[4] Regulation EU 2016/679.

as a valid legal basis for data collection and processing, rendering these activities unlawful (European Data Protection Board, 2020). In the case under examination, Facebook users were likely ***not even aware*** of the possibility of having their data shared by their friends to third parties by merely consenting to the broad Facebook terms of service. This circumstance suggests that the social media platform contractual approach did not provide a lawful basis for valid consent, not exempting Facebook, therefore, from accountability in the terms of Article 5(2), and Chapter 8 (Articles 77 to 84) of the GDPR.

With the described dataset at hand, CA firstly cross-referenced the test-takers psychological results against their Facebook data. Based on the correlations, the company could then train algorithms for inferring personality types solely from Facebook interactions, categorizing, thereby, millions of individuals that did not have to take the test. See **Fig. 1**.

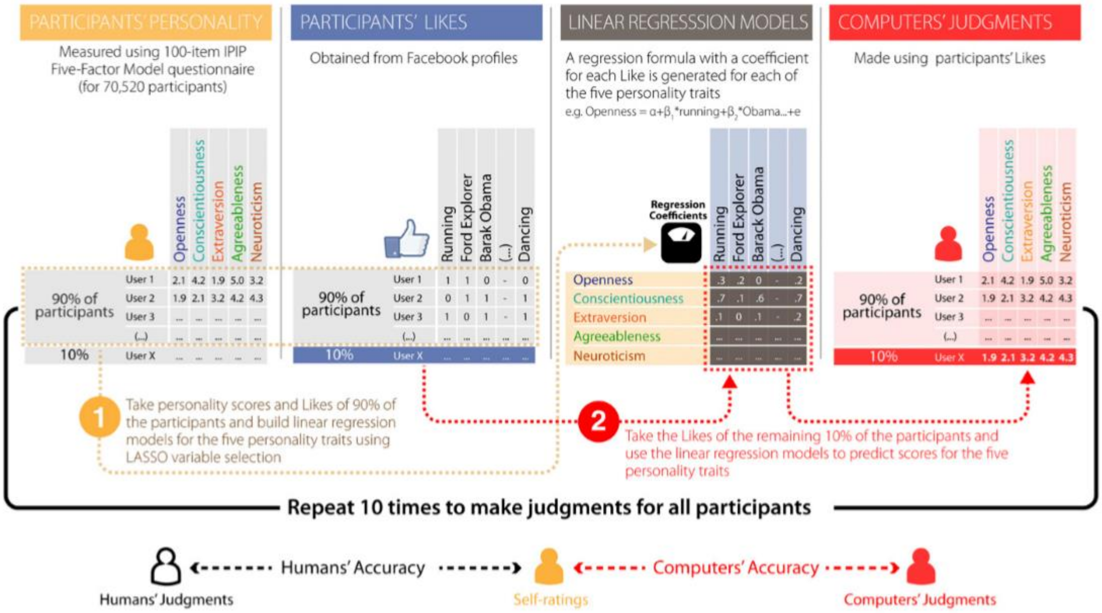

PARTICIPANTS' PERSONALITY
Measured using 100-item IPIP Five-Factor Model questionnaire (for 70,520 participants)
Openness
Conscientiousness
Extraversion
Agreeableness
Neuroticism
90% of participants
User 1 2.1 4.2 1.9 5.0 3.2
User 2 1.9 2.1 3.2 4.2 4.3
User 3
(...)
10%
User X
PARTICIPANTS' LIKES
Obtained from Facebook profiles
Running
Ford Explorer
Barak Obama
(...)
Dancing
90% of participants
User 1 1 1 0 - 0
User 2 0 1 1 - 1
User 3 1 0 1 - 1
(...)
10%
User X
LINEAR REGRESSSION MODELS
A regression formula with a coefficient for each Like is generated for each of the five personality traits
e.g. Openness = α+β1*running+β2*Obama...+e
Regression Coefficients
Running
Ford Explorer
Barack Obama
(...)
Dancing
Openness .3 .2 0 - .2
Conscientiousness .7 .1 .6 - .7
Extraversion .1 0 .1 - .2
Agreeableness
Neuroticism
COMPUTERS' JUDGMENTS
Made using participants' Likes
Openness
Conscientiousness
Extraversion
Agreeableness
Neuroticism
90% of participants
User 1 2.1 4.2 1.9 5.0 3.2
User 2 1.9 2.1 3.2 4.2 4.3
User 3
(...)
10%
User X 1.9 2.1 3.2 4.2 4.3
1
Take personality scores and Likes of 90% of the participants and build linear regression models for the five personality traits using LASSO variable selection
2
Take the Likes of the remaining 10% of the participants and use the linear regression models to predict scores for the five personality traits
Repeat 10 times to make judgments for all participants
Humans' Accuracy
Computers' Accuracy
Humans' Judgments
Self-ratings
Computers' Judgments

__Figure 1__: A study from the University of Cambridge shows the method used to obtain automated judgments of personality traits using linear regression models (machine learning algorithms). Diagram from (Youyou, Kosinski, & Stillwell, 2015).

Research from the University of Cambridge shows how computer-based personality judgments can be more accurate than those made by humans (Youyou, Kosinski, & Stillwell, 2015). The study has compared the accuracy of human personality judgments to those from computer-based analysis obtained using a sample of 86,220 volunteers who completed a 100-item personality questionnaire. The findings reveal that computer models need only 100 Likes to surpass an average human judge. With just 10 likes, the technology would be better at predicting a person's personality than an average coworker would be. With 70, 150, and 300 Likes, respectively, the algorithms outperform a cohabitant or friend, family member, and spouse (Youyou, Kosinski, & Stillwell, 2015). See **Fig.2**.

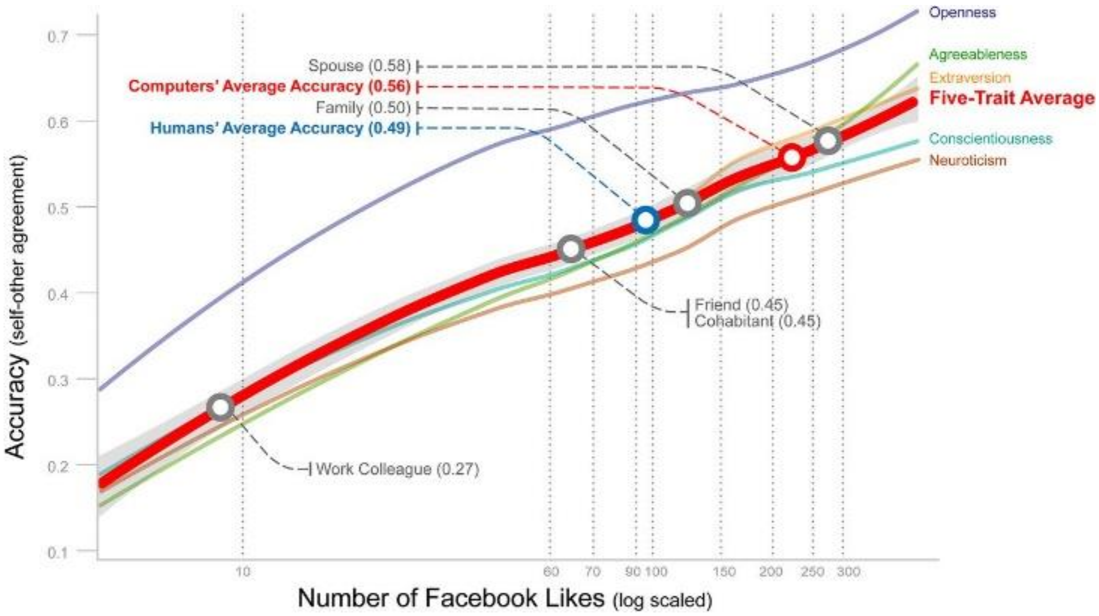


Figure 2: Computers' average accuracy across the Big Five traits (red line) steadily grows with the number of Likes available on the participant's profile (x-axis). Graph from (Youyou, Kosinski, & Stillwell, 2015).

The 'grand finale' of CA's algorithmic personality analysis and profiling was precisely its manipulative use. Through the profiles established from Facebook Likes, the data analytics firm could customize political messages exploring individuals' psychological vulnerabilities and target them accordingly. As openly phrased by the company:

> 'Analyzing millions of data points, we consistently identified the most persuadable voters and the issues they cared about. We then sent targeted messages to them at key times in order to move them to action' (Cudd & Navin, 2018).

Although the increasing sophistication and predictive precision of algorithmic analyses can unsettle, one might dispute the effectiveness of political micro-targeting to influence individuals' decisions. Whereas scientific research is not conclusive about how effective the practice of micro-targeting can be in steering political behavior, experiments suggest that predictive models could 'yield sizable and electorally meaningful gains to campaigns' (Nickerson & Rogers, 2014). Research shows that micro-targeting techniques can be 'a highly efficient way of managing campaigns in dynamic environments, as they allow for adaptive strategies relative to broad campaigning' (Madsen & Pilditch, 2018). Accordingly, the 'detailed knowledge of voter's identity' (Hersh & Schaffner, 2013) would enable political campaigns to concentrate their efforts and resources where they will be most effective (Nickerson & Rogers, 2014).[5]

[5] The assumption that micro-targeting can be an efficient tool for managing political campaigns, bringing electoral gains, is supported by declarations of political actors involved in the UK 2015 elections. In this regard, Tom Edmonds, Creative Director of the Conservative Party in the UK declared: '*We built an in-house agency to create the content we needed with the messages we knew would work—and used Facebook targeting to get that content in front of the voters who would decide the election. It was the first genuinely digital election in UK political history'.* See (Facebook, n.d.) Craig Elder, Digital Director from

Regardless of uncertainties on the persuasion power of political micro-targeting, scientific evidence suggests that:

> '(…) the improved efficiency gives data-savvy campaigns a competitive advantage. This has led the political parties to engage in an arms race to leverage ever-growing volumes of data to create votes' (Nickerson & Rogers, 2014).

An experiment on the persuasive power of pandering directed to groups also presents a thought-provoking finding: despite the effectiveness of targeted messages among the intended group, political candidates would likely lose support when voters outside the targeted group could see the targeted messages (Hersh & Schaffner, 2013). This indicates that election results would probably be diverse whether advertisements in campaigns were available for all citizens to see. These conclusions reinforce the relevance of transparency in political campaigns, as it is likely that political candidates would perform differently in elections if voters received full information over their platforms. The study also highlights the research limitations of testing campaign situations in artificial settings. The generalization of the findings to real-world scenarios would require, hence, replicability and further studies (Hersh & Schaffner, 2013).

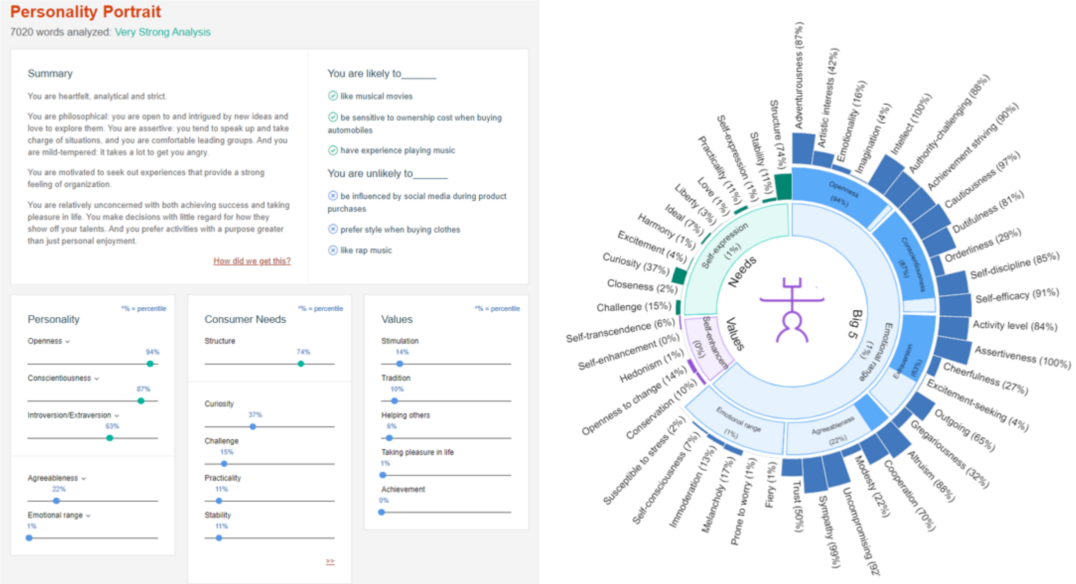
Considering the fast pace of technological development, one can also assume that the effectiveness of political micro-targeting can only become higher. Hence, groups of interest will be able to exploit individuals' psychological vulnerabilities and control their will, with increasing effectiveness. Media channels have reported research on a new generation of machine learning tools for behavior prediction that already overcome the Cambridge Analytica model (Vogels, 2016). An example of the continuous innovation process in the field is an IBM engine is capable of precisely infer personality traits from pieces of text. The company advertises its product in the following terms:

> 'Gain insight into how and why people think, act, and feel the way they do. This service applies linguistic analytics and personality theory to infer attributes from a person's unstructured text' (IBM, 2020).

The analysis, with a free demonstration available for the public, includes a detailed personality portrait, a list of the person's values and needs, and information on what kind of advertisement the subject would better respond to (Vogels, 2016). The application also enables the visualization of personality traits in a sunburst chart. This kind of plot represents character traits identified by the application like a doughnut chart but with different levels of hierarchy. Accordingly, the innermost circle or ring depicts the top of the hierarchy (Microsoft, 2020), see **Fig. 3**. The IBM demonstration system warns, however, that it uses no personal data, 'as it may not have the necessary controls

The Conservative Party in the UK: '*The level of targeting we had available to us on Facebook—coupled with the research and data we produced internally—meant that we can say for the first time in a UK election that digital made a demonstrable difference to the final election result*'. See (Facebook, n.d.)

in place to meet the requirements of the General Data Protection Regulation (EU) 2016/679'. Nevertheless, the system manages to deliver precise personality analysis outputs.



Personality Portrait

7020 words analyzed: Very Strong Analysis

Summary

You are heartfelt, analytical and strict.

You are philosophical: you are open to and intrigued by new ideas and love to explore them. You are assertive: you tend to speak up and take charge of situations, and you are comfortable leading groups. And you are mild-tempered: it takes a lot to get you angry.

You are motivated to seek out experiences that provide a strong feeling of organization.

You are relatively unconcerned with both achieving success and taking pleasure in life. You make decisions with little regard for how they show off your talents. And you prefer activities with a purpose greater than just personal enjoyment.

How did we get this?

You are likely to______

- ✓ like musical movies
- ✓ be sensitive to ownership cost when buying automobiles
- ✓ have experience playing music

You are unlikely to______

- ☒ be influenced by social media during product purchases
- ☒ prefer style when buying clothes
- ☒ like rap music

Personality *% = percentile

Openness ⌄ 94%

Conscientiousness ⌄ 87%

Introversion/Extraversion ⌄ 63%

Agreeableness ⌄ 22%

Emotional range ⌄ 1%

Consumer Needs *% = percentile

Structure 74%

Curiosity 37%

Challenge 15%

Practicality 11%

Stability 11%

>>

Values *% = percentile

Stimulation 14%

Tradition 10%

Helping others 6%

Taking pleasure in life 1%

Achievement 0%

Figure 3: The images show the outputs from the IBM 'Personality Insights' engine for Barack Obama's personality traits, based on the textual analysis of a 2012 debate. The scores, shown in percentiles, compare the person's characteristics to a sample population of Twitter users whose personalities were calculated using the IBM model. Source: (IBM, 2020)

In the same manner as e-commerce, no technical obstacle seems to prevent the use of this personal knowledge of voters as a means of gradually changing political perceptions, beliefs, and voting intentions. Technology, as a tool each day more intertwined with society, might serve for multiple human ends. These are certainly not neutral but guided by social, political, and economic interests. The big challenge arises, thus, for ethicists, academics, and lawmakers in assessing harmful uses and defining limits for disrupting technologies in a way it can benefit collective interests. Previous research indicates that micro-targeted campaigns become more frequent in politics (Madsen & Pilditch, 2018) and academics in the field assume that micro-targeting techniques associated with data-driven campaigning will increasingly permeate the European political conjuncture (European Data Protection Supervisor, 2018).[6] The European Commission has recognized, furthermore, the micro-targeting of voters based on the unlawful processing of personal data, along with the exposure of citizens to online disinformation, as one of the major challenges for European democracies (Nenadić, 2019).

From a legal perspective, this constantly evolving technological scenario poses substantial risks to individuals' autonomy and democratic regimes, which will be detailed further in this work. It demands, consequently, urgent regulation. The North

[6] A study requested by the European Parliament brings to light that 'The use of political micro-targeting was also reported in Italy, Germany and the Netherlands, although to a lesser degree. The situation in other EU Member States is less researched and thus less clear'. See (Bayer, et al., 2019).

American case illustrates this situation, where ‘relatively loose data-protection regulation may have facilitated the rapid development and adoption of the technique’ (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019) of political micro-targeting.

In summary, the scientific indication that micro-targeting is a mechanism that can work and become more efficient with rapid technological enhancement. This particular use case poses serious risks for society. These factors combined should be more than sufficient to trigger regulatory initiatives. Here, the analogical application of the precautionary principle might well be argued to regulate technological applications. This is often claimed as a principle of international law (Davies G. T., 2009) and vastly used to reject the need for scientific ***certainty*** for regulating environmental issues. Whether the notion of precaution pervades popular wisdom and is enunciated in sayings as ‘better safe than sorry’, the precautionary principle can be formulated as ‘*not* having scientific certainty is *not* a justification for *not* regulating’ (Hanekamp, Vera-Nava, & Verstegen, 2007).

A central issue for the regulation of emerging technologies is, therefore, how long governments should wait to attack the new threats presented to fundamental rights. The so-called ‘**Collingridge Dilemma**’ is pertinent to this matter, as it explains the difficulties of regulatory timing when there are legislative gaps associated with potential new harms or risks. While at an early stage, regulation is problematic for the lack of full information about the innovation’s likely impacts, at a later point the technology becomes more entrenched in society, making legal and policy changes harder to implement (Bennett Moses, 2013). The dilemma underlies ‘the importance of taking regulatory control over situations when the first signs of problems start emerging’ (Ranchordás & van ‘t Schip, 2019). See **Fig. 4**

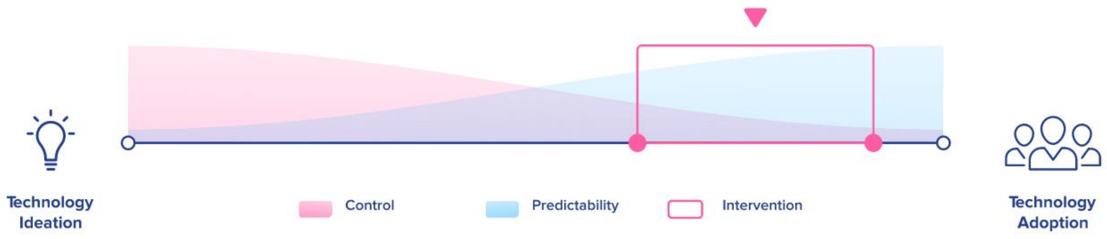


Figure 4: The chart represents the hardships in regulating a new technology according to the degree of diffusion in society and a suggestion of when intervention would be ideal. While in the first stages predictability is lower, with time control over technology decreases, making it harder to regulate. Source: (Besti, 2019).

A solution for the impasse might be the implementation of a flexible regulatory framework, for instance, based on general principles and rules. Another example could be a goal-based solution which could adapt to new scenarios. Although very detailed regulations might be easier to carry out, the setback is that these can be easily overcome by new technological features unpredictable at the time of its edition. In its turn, more general norms may be harder to implement, promote less certainty, and may inhibit innovation if too broad. They can, however, be applied to new situations, being **'future-proof'** in a sense. The non-foreseeability of technological advances requires the regulator to balance the interests between the adaptability of broader norms, the legal certainty and enforceability of detailed regulations. Other proposals for a future-proof legislative approach involve the formulation of experimental legislation to test the effectiveness of new solutions and the assessment of the regulatory impact considering short- and long-term risks of a novel regulation or policy (Ranchordás & van 't Schip, 2019). Further investigations on regulatory instruments are discussed in Chapter 4.

2.2 The anti-democratic threats in algorithmic political manipulation

In general lines, manipulation can be described as the attempt to induce a person to behave according to the manipulator interests by exploiting their decision-making vulnerabilities (Susser, Roessler, & Nissenbaum, 2019). Some scholars place the **hidden nature** of the influence as a central element to the definition, that is, the target unawareness of the manipulation. Others argue that the phenomenon is characterized by a **non-rational, emotional influence**, but not necessarily hidden. For the latter, 'manipulating someone means influencing them by **circumventing** their **rational**, deliberative **decision-making** faculties' (Susser, Roessler, & Nissenbaum, 2019). Regardless of the definition adopted, algorithmic political manipulation described in Section 2.1 may fit the definition as a special case of manipulation through information technology. In the case of the CA scandal, the individuals targeted were completely unaware of the political influence exercised through the exploitation of their data. This thesis argues, however, that even if people had given consent for the processing of their data for psychological categorization purposes, algorithmic political micro-targeting could still influence an individual's behavior in a manipulative way, due to its level of precision. This Section will explore the special features of this online manipulative phenomenon that make it a special threat to democracy.

2.2.1 Offline political micro-targeting

Human minds are constantly swayed by a diffuse web of inputs. Biological factors, environmental conditions, educational, political discourses, and ultimately every social interaction contributes to shaping differing personalities and understandings of the world. 'Diversity is a biological fact, continually reproduced in each generation, regardless of anyone's intentions. Diversity is also a cultural product' (Greenwood & Levin, 2007). If multiple biological and societal factors make each of us different in 'who we are', it stands to reason that individuals will react distinctly to messages or actions, according to their personality, perceptions of reality and beliefs.

The use of personal characteristics for influencing human behavior is not a new phenomenon in society. It can be perceived, for instance, in the ordinary situation of the salesman that makes product's suggestions for clients based on the way they present themselves; or, likewise, tries to influence an extrovert consumer by being more talkative to them, or by letting more reserved one take their time in the store. Similarly, in the political sphere, the use of the voter's traits for targeted political propaganda is not a novel experience either. In the US, the practice of canvassing illustrates how political micro-targeting developed in the offline context (Zuiderveen Borgesius, et al., 2018), long before algorithmic political targeting could be put in practice. Through door-to-door contact with electors, political parties and intermediaries could 'hold extremely detailed information about possible voters' (Zuiderveen Borgesius, et al., 2018) and use it to plan their campaign strategies.

2.2.2 How algorithmic political micro-targeting is different

New methods of data collection and analysis gave rise to a much more sophisticated form of political targeting (Zuiderveen Borgesius, et al., 2018), referred here as micro-targeting. In the online context of growing availability of user's data, the personalization of political propaganda has reached the micro-level, being able to exploit the 'unique weaknesses of individual brains' (Harari, 2018). Previously, voters could be profiled and targeted according to personal traits directly noticeable by the human perception - as in the practice of canvassing, for example. In the tech-democracy era, computational technology enables the creation of fine-grained psychological profiles to tailor-make political messages. Political advertisements are evolving into 'precision-guided munitions' (Harari, 2018). This unique element of exploiting individual vulnerabilities in a psychological and intimate level is, therefore, what makes algorithmic political micro-targeting an especially powerful tool of manipulation. In this respect, Yuval Harari claims that:

> 'in recent years some of the smartest people in the world have worked on hacking the human brain in order to make you click on ads and sell you stuff. Now these methods are being used to sell you politicians and ideologies, too' (Harari, 2018).

The distinguished power of subconsciously influencing decision-making processes compromises, therefore, the capacity of individuals to make independent choices, representing insidious harm to individual autonomy (Susser, Roessler, & Nissenbaum, 2019). From a political perspective, the 'target of particular voters with tailored information that maximises, or minimises, voter engagement' (Zuiderveen Borgesius, et al., 2018) represents a sensitive threat of manipulation. In this vein, individuals could be exposed to information that aggravates their bias or intolerances. A party could use social media, as an example, 'to expose xenophobic voters to information about the high crime rates amongst immigrants' (Zuiderveen Borgesius, et al., 2018). Micro-targeting could also lead to a biased perception of a political party program, for highlighting or distorting different proposals according to the voter profile. Therefore, 'leading to a lack of transparency about the party's promises' (Zuiderveen Borgesius, et al., 2018).

Other manipulative practices in political campaigns might aim at suppressing political participation. The example of the infamous ‘Do So’ (‘Don’t Vote’) campaign in Trinidad and Tobago is remarkable in this respect. In this case, CA, working on behalf of the majority-Indian United National Congress (UNC) party, promoted a campaign for micro-targeting young black voters. By sending tailored messages that associated refusal to vote as ‘a sign of resistance against politics’ (Global Voices, 2019), the firm managed to demobilize potential electors of their political competitor, the People's National Movement (PNM). The UNC was victorious in the elections (Global Voices, 2019). Another example relates to the reported ‘dark posts’, or ‘unpublished posts’[7] during the 2016 Donald Trump campaign that targeted African American voters with messages of Hillary Clinton calling African American males ‘super predators’ (Zuiderveen Borgesius, et al., 2018). For William A. Gorton, micro-targeting ‘undermines the public sphere by thwarting public deliberation, aggravating political polarization, and facilitating the spread of misinformation’ (Gorton, 2016).

Karen Yeung also highlights the power of algorithmic decision-guiding techniques and their ‘troubling implications for democracy’ (Yeung, 2016) posed by its intensive use. A big-data-driven nudge, or an algorithmic ‘hypernudge’, would consist in the systematic use of big data techniques for ‘altering human choices in a predictable way without forbidding any options or significantly changing their economic incentives’ (Yeung, 2016). Unlike traditional forms of nudging human behavior, like speed bumps or highlighting vegetables in grocery stores, big-data-driven nudge would have the potential to ‘directly affect millions of users simultaneously’ and produce unprecedented surveillance and control of individuals (Yeung, 2016). These features reinforce the anti-democratic threats posed by political uses of ‘hypernudge’.

Some scholars warn, moreover, of the democratic risks associated with the creation of ‘virtual echo chambers’, or ‘filter bubbles’ through algorithmically personalized advertisement, meaning spaces where a limited set of ideas is constantly reinforced (Wilson, 2017). This so-called ‘resonance effect’ (Helbing, et al., 2017) can harm the core democratic value of the plurality of ideas and jeopardize the political debate. Through political micro-targeting, internet users can be deluded into believing that the received information is objective and universally encountered by other people, when in fact, it specifically targets them (Bayer, et al., 2019). A study published in Scientific American captures the social consequences related to ‘social polarization, resulting in the formation of separate groups that no longer understand each other and find themselves increasingly at conflict with one another’ (Helbing, et al., 2017). It remarks that:

[7] According to Facebook Business Help Center, ‘Unpublished Page posts allow Page admins to manage delivery of ad content through audience filters. These scheduled or draft posts are delivered on a future publication date or through promotion within an ad set’. See Facebook for Business. *Fundamental Beginner’s Guide*. Facebook.
Retrieved from https://www.facebook.com/business/help/835452799843730 (Accessed 06 March 2020).

> 'In this way, personalized information can unintentionally destroy social cohesion. This can be currently observed in American politics, where Democrats and Republicans are increasingly drifting apart, so that political compromises become almost impossible. The result is a fragmentation, possibly even a disintegration, of society' (Helbing, et al., 2017).

Other democratic risks relate to the unfair advantage that the use of micro-targeted campaigns would give to larger and wealthier political parties in detriment of less funded ones. In this sense, financial power could distort democratic processes and undermine the 'free flow of political ideas', since rich parties would have conditions to hire specialized firms as intermediaries to run modern tech-driven political campaigns (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

2.2.3 Algorithmic political micro-targeting and fake news

If political micro-targeting is problematic from the perspective of individual autonomy and democracy, it gets even more alarming when the targeting of voters involves the spread of false content. This phenomenon might fit the notion of deception as a special case of manipulation, consisting of the planting of false beliefs to covertly influence someone (Susser, Roessler, & Nissenbaum, 2019). The manipulative potential of algorithmic targeting could be amplified, in this regard, by its combination with disinformation networks that emerged through social media for disseminating false messages.[8] If, on the one side, algorithmic micro-targeting is a powerful manipulative mechanism for giving partial views that are most likely to affect individuals' behavior, on the flip side, the targeting through fake-news could be defined as an 'enhanced algorithmic political manipulation', for providing false tailored representations of the political debate to provoke emotional reactions. Fake news became popular, in this regard, for instigating irrational fears and bias in the voters, often disseminating hate speech. Micro-targeted 'deep-fakes', consisting of incredibly convincing manipulated audios or videos, are illustrative of the use of online micro-targeting to misinform specific groups (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

The European Commission observes that there is empirical evidence that false news is spreading significantly 'faster, deeper, and more broadly' than the true ones (European Commission). A study published in Science indicated that falsehood tends to reach more people than the truth and are diffused faster on the web. It found that the top 1% of false news cascades reached from 1000 to 100,000 people, whereas the true ones rarely reached more than 1000 people (Vosoughi, Roy, & Aral, 2018). The report claims that 'the degree of novelty and the emotional reactions of recipients may be responsible for

[8] The Avaaz Investigative Report 'Far Right Networks of Deception' uncovers the tactics of disinformation networks for systematically spreading misleading or false political content during elections across Europe. See (AVAAZ, 2019).

the differences' (Vosoughi, Roy, & Aral, 2018). Possibly, 'the information that makes us angriest becomes the information least likely to be questioned' (Ghosh & Scott, 2018). The combination of micro-targeting and false messages maximizes, therefore, the impacts of each personalized message (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019). Considering these findings and the more intense emotional appeal of **fake targeted messages**, it can also be reasonably assumed that its **power** to influence individuals is **even higher**, with even more serious **impacts** on the democratic debate.

2.2.4 Micro-targeting, human autonomy, and democracy

The potential of algorithmic micro-targeting to steer individuals' political behavior in a pervasive manner raises questions about the meaning of human autonomy in democratic societies. Yuval Harari claims that human minds are continuously subjected to external influences, highlighting that the liberal notion of 'free will', as complete freedom of choice, would be a myth inherited from the Christian theology (Harari, 2018). Considering this as a valid premise, what does autonomy mean in contemporary democracies? This thesis assumes that rather than an illusory and arguably sad conception of complete independence from others, democracy requests that human beings are not converted into extensions of others' will. Democracy does not presuppose a liberal notion of free-will as complete freedom of choice, but it requires that citizens do not have their will controlled by others, their brains 'hacked', without notice. Although every person is a product of their environment and influenced by multiple determinations, democracy relies on the assumption that persons are not intentionally controlled in their will without realizing it. This is precisely the risk posed in the current technological stage of development of algorithmic micro-targeting: the 'hacking' of the human brain (Harari, 2018), where the will of the targeted electors become the will of the 'micro-targeters'. This democratic notion can be visualized, for instance, in existing norms that prohibit anyone from paying other people to vote for a certain candidate or that ensure voting confidentiality. The influence exerted by parents over children is also an illustrative example of the thin line between influence and control over other people's minds. While an infant's personality is certainly affected by the relationship with their parents, this influence cannot be regarded as a democratic problem, as parents' interventions do not have the power to determine their child's political preferences and decisions. As much as a father or mother could wish to shape a child's personality, they cannot turn them into exactly what they desire.

The Kantian formulation of human dignity as a formula of humanity is pertinent in this context, meaning that no person should be treated as a means to other peoples' ends. As 'the deepest justification for human rights' (Andorno, 2009), the general principle of human dignity is often interpreted in law as a protection against degradation and objectification of human beings, even when there is consent to their own instrumentalization (van Beers, 2012). In this light, taking control of other people's decisions would make them instruments for the controller's purposes. Just like marionettes, steered human beings would represent the will of their puppeteers. Micro-targeting and the quick development of this technique (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019) represent, in this regard, a threat not only for individuals’ autonomy in democratic societies but ultimately to human dignity.

3. The European regulatory framework for algorithmic political micro-targeting

Given the risks for violating human rights and fairness in electoral processes, discussed in Chapter 2, Chapter 3 examines the strengths and shortcomings of the law applicable to political micro-targeting in the European Union (EU) context. Considering the gap of specific legislation, the EU data protection set of rules is deemed as the main legislative body fitting. In this field, the consent-based approach, verified both in the GDPR and in the ePrivacy Directive, is a major object of criticism. This chapter will address some of these critiques, related to the model's inefficiency to provide effective control over personal data and prevent manipulative threats posed by political micro-targeting.

Within the scope of Member States legislation, rules on political advertising could also regulate micro-targeting in political campaigns.[9] However, due to time limitations, this thesis limits its scope to the analysis of European provisions, not examining the national legislations of EU member states.

3.1 The existing law in Europe: Benefits and gaps in data protection framework

As a novel phenomenon supported by developing technologies, it is not surprising that algorithmic micro-targeting does not yet find specific legislation in Europe. Since the practice of targeting voters using data analytics methods inherently involves the collection and processing of personal data, the EU privacy and data protection set of rules applies to this case.

This Section discusses strengths and weak points of the general data protection norms applicable in the EU context, in order to formulate a more functional and future-oriented regulatory approach to political micro-targeting, as proposed in Chapter 4.

3.1.1 Data Protection in EU primary law

In the European *primary* Law,[10] the protection of personal data is granted the status of an autonomous fundamental right (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019). The Charter of Fundamental Rights of the European Union (CFR) sets this right forth

[9] In this regard, the European Union (EU) has no specific competence to regulate national elections (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

[10] Together with unwritten general principles of law, the provisions of the Charter of Fundamental Rights of the European Union constitute the EU primary law, applied when Member States are implementing Union Law. See (European Union Agency for Fundamental Rights, 2018).

in Article 8, apart from the right to respect for private and family life provided for in Article 7. Article 8 of CFR establishes that personal data must be processed fairly, for specified purposes, and upon consent or another legitimate legal basis, besides referring to the rights to access, rectification, and erasure of data by the concerned persons. In its last item, the legal provision outlines that an independent authority must control the compliance with these rules.

The special protection granted to personal data in Europe, as a human right, is precisely what makes it a non-tradable element, hindering, for instance, the direct purchase of voter's personal data by political parties (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

3.1.2 Data Protection in EU secondary law

The EU *secondary* data protection law consists of two main legal instruments: the GDPR, in effect since 2018, and the 2002 EU ePrivacy Directive, at the time of writing still being revised by the EU.[11] The GDPR delineates specific rules for the fair and transparent processing of personal data. Personal data is understood as any information relating to an identified or identifiable natural person ('data subject'). The ePrivacy Directive, for its part, sets out rules for the use of cookies to trace behavior of internet users (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019). Nevertheless, both regulations rely on the notion of ***informed consent***. In the GDPR the informed consent is required for the lawful processing of personal data, whereas the EU ePrivacy Directive requires it to set tracking cookies on someone's computer.

3.1.3 Data Protection as part of the solution for micro-targeting

Despite the existence of data protection norms in Europe since the 1995 Data Protection Directive, the GDPR entry into force in 2018 brought a new quality of enforcement to this matter. In the context of the 1995 Directive, dependent on the implementation of each EU Member State, scholars note that 'some states passed weak laws and signalled a business-friendly environment with weak-wristed Data Protection Authorities' (Hoofnagle, van der Sloot, & Zuiderveen Borgesius, 2019). The GDPR, as a modern legal instrument legally binding and self-executing, i.e., immediately applicable across the EU, has harmonized national data protection and privacy laws. In terms of enforcement, the GDPR has established not only relevant fines for non-compliance and a workable data-breach notification system, but also empowered Data Protection Authorities with considerable enforcement competences. 'Such authorities broad

[11] The EPD's eventual replacement, the ePrivacy Regulation (EPR), will build upon the EPD and expand its definitions. (In the EU, a *directive* must be incorporated into national law by EU countries while a *regulation* becomes legally binding throughout the EU the date it comes into effect.) https://gdpr.eu/cookies/

powers to investigate, to intervene and even halt data processing, and to bring legal proceedings' (Hoofnagle, van der Sloot, & Zuiderveen Borgesius, 2019).

Regarding the lawful processing of personal data, the GDPR requires from organizations that control the processing of personal data (data controllers) observance of the Fair Information Principles (FIP) under Article 5. The processing of personal data should be done under one of the six legitimate grounds provided in Article 6, being the consent of people whose data are used (data subjects) one of the most relevant cases. As a rule, the use of special categories of data is prohibited, except for the cases expressly mentioned in Article 9 (2).

Although not created to regulate micro-targeting, many of these rules may be regarded as useful to limit the practice. With no pretension to be exhaustive in the analysis, some pertinent rules to the case under examination are the requirement of ***informed, freely given, specific, unambiguous*** and ***through a clear affirmative action*** consent, as established in Article 4 (11). The GDPR defines, therefore, a high standard for consent to be regarded as a lawful basis for the collection and processing of personal data. This represents a big change for consenting mechanisms in practice, making it significantly harder for organizations to obtain valid consent from data subjects (Morris, 2019). The ***principles of transparency and purpose limitation***, set out respectively in Article 5(1) (a), (b), GDPR, and ***the rights of the data subjects to access and to obtain the erasure*** of its personal data, set forth in Articles 15 and 17 of the Regulation, are also relevant

The combination of the ***GDPR conditions for lawful consent*** require that data controllers obtain a clear, fully informed, and affirmative, consent from data subjects for each specified purpose, before collecting and processing personal data. In the case of political micro-targeting, a reasonable interpretation of these dispositions is that data controllers must previously inform the concerned subjects, in a transparent manner, that personal data are collected for the personalization of political advertisements. More than that, 'freely-given' consent means that the interested person must be able to easily deny consent for a processing operation that is not necessary for the performance of a contract or service.

In practice, both the existence of consent in the GDPR terms and the effectiveness of these conditions to avoid political manipulation are disputable and dependent on empirical confirmation. Nonetheless, a full compliance of data controllers with those rules could at least enable data subjects to actively deny and restrict the use of their data for political marketing. Non-compliance with the regulation can entail, furthermore, administrative fines up to up to 20 million euros or, in the case of an undertaking, up to 4% of the total worldwide annual turnover of the preceding financial year, whichever is higher.[12] Moreover, the ***GDPR transparency requirements*** enable other actors, such as journalists and researchers, to discover how personal data are used by companies,

[12] Article 83(5), GDPR.

political parties, or other groups of interest in society (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

The ***rights of access to personal data, erasure, and restriction of processing*** are also relevant in this context. This set of possibilities empowers the data subjects to know what information was collected by them, to request the deletion of the data, or to restrict its processing. It **can limit**, therefore, the processing of personal data for political micro-targeting. The right of access is depicted in the documentary 'The Great Hack' on the CA scandal when a North American professor files a legal claim against the company (Fischer, 2019) after having received a poor response to a request of access to his data (White, 2019).

Another point of interest about micro-targeting is the ***special protection granted to sensitive data***. The GDPR prohibits in its Article 9 the processing of special categories of data when this activity could create significant risks to fundamental rights and freedoms of the data subjects, also listing exceptions. The general prohibition includes, for instance, the processing of personal data revealing **political opinions,** racial or ethnic origin, religious or philosophical beliefs, or trade union membership.

A reasonable interpretation of the legal rule is that ***all data, from which this sort of opinions and beliefs could be concluded, should be considered as sensitive data*** for GDPR application purposes. In this regard, political micro-targeting makes use of algorithms and big data analytics to infer people's personalities from apparently 'innocent' data not related to politics. Such personal traits include, or might be a proxy to, one's **political opinions**, racial or ethnic origin, religious and philosophical beliefs. Based on these premises, this thesis supports that any data that are used by micro-targeting algorithms should be considered, *per se*, sensitive. Therefore, the GDPR limitations to the collection of sensitive personal data to any data collected with micro-targeting purposes would be directly applied. Furthermore, the conditions under which sensitive data could be processed for political micro-targeting would be set out, in Article 9 (2) (a) and (d) of the GDPR. These involve the ***explicit consent*** of the data subject, or the use of sensitive data by a not-for-profit body with a **political**, philosophical, religious or trade union aim. In the latter case, the processing is restricted by law to sensitive data of members, or former members of the organization who are in regular contact with it. The processing should relate, in addition, to the institution purposes and not be disclosed outside that body without the consent of the data subjects. Hence, this line of thinking conditions the use of personal data for political micro-targeting to the explicit consent of the data subjects, or to the data of members of a non-profit organization for the entity purposes. Pursuant to the EDPB guidelines on consent, considering that the threshold to obtain 'regular' consent is already high in the GDPR, the extra effort required in ***explicit consent*** refers to an express statement of the data subject. This might be obtained, for example, through an electronic form, an email, the upload of a scanned document carrying the signature of the data subject, or by using an electronic signature (European Data Protection Board, 2020).

From the analysis of the aforementioned GDPR provisions, this Thesis concludes that the existing EU data protection framework surely plays a relevant role in the limitation of the amount of personal data collected and processed, which includes micro-targeting purposes. However, some points of criticism lie in the limitations of a privacy-focused legislative approach and the GDPR consent-based model for regulating political micro-targeting, as explained in Section 3.1.4. It is worth mentioning in this vein that, in the time of its drafting, the European lawmakers did not have in mind the special context of algorithms used for political micro-targeting. Thus, a complete solution for this matter calls for rules that consider the specific impacts of this practice and aim to prevent the threats for elections, individual autonomy, and the democratic rule of law. Despite technology regulation requiring flexible norms as a rule, the resolution of ***privacy issues***, or ***democratic issues*** related to the processing of personal data may require different legislative approaches.

3.1.4 The limitations of the consent-based model

Grounded on the individual control of personal data, the GDPR consent-based model establishes the informed and freely given consent as one basis for the lawful processing of personal data. From this perspective, if a person is clearly informed about what she is consenting to and has the option to freely deny it, the processing for a purpose such as political micro-targeting would be regarded lawful.

Two major criticisms impend over this legislative design. The *first* critique concerns the existence or viability of a ‘meaningful consent’ in reality, considering informational and/or power asymmetries between data subjects and data controllers. The *second* one connects to the fact that even the existence of a fully informed and freely given consent does not eliminate the underlying dangers of manipulation posed by this phenomenon, and thus it is directly related to the case of political micro-targeting.

Regarding the *first* criticism, from laypeople to scholarly literature argue that users would not have adequate information nor a real choice concerning the processing of their data. Benjamin Bergemann states in this regard that:

> ‘Both groups argue that it is hard for users to comprehend what they are consenting to. Moreover, they criticize that users often do not have a choice but to consent because they rely on products such as social network services or smartphones’ (Bergemann, 2018).

The author brings about the idea of the ‘**consent paradox**’, consisting of the prominent role assigned to consent in data protection at the same time it is subject to numerous criticisms (Bergemann, 2018). To address this apparent contradiction, many scholars ‘emphasize that consent continues to be an essential part of data protection policies despite its perceived limits’ (Bergemann, 2018).

Some authors will propose a reform of the current consent model, for instance, by enhancing transparency duties to mitigate ***information asymmetries***. Others seem to question the notion of free and informed consent in general, in a more fundamental objection. Bergemann's critique calls into question whether consent can be considered freely given from the perspective of ***power asymmetries***. From this point of view, power disparities between internet users and digital platforms would not allow freely given consent in essence. As users are increasingly dependent on these service providers, there would be few alternatives or room for negotiation of privacy policies (Bergemann, 2018).

Criticisms focused on '***information asymmetries***' warn against the fact that most people do not read or understand privacy policies and conditions for personal data collection and usage. Possible reasons for that involve the excessive and unreasonable time consumed for reading consent notices, Also, their lack of clarity, often including legal and technical jargon. In addition, the frequency that consent banners happen to appear in the current online environment. Although improved transparency may not solve all the limitations in the consent-model, this Thesis agrees with the premise that higher information favors better decisions by data subjects (Bergemann, 2018). Thus, transparency should be considered at least a part of the solution to any problem with respect to the collection and processing of personal data.

Despite these general critiques of the GDPR privacy self-management model, a ***specific remark*** can also be addressed to political micro-targeting. This criticism refers to the fact that even though full information can, in theory, enable the data subjects to deny the consent for the personalization of political advertisement, ***consenting to this practice does not necessarily avoid the manipulative threat in micro-targeting***.

In the GDPR terms, consent aims to give people control over their personal choices, enabling them to decide for each purpose their data can be used. Therefore, consenting to a process that might entail the loss of control over political decisions would be a contradiction in these terms. Eventually, giving consent to micro-targeting would equal to giving up control over one's political will, or to the delegation of one's own civic will to third parties. Such a process would not come to terms with the democratic principle, which relies on no level of control of others to steer their political behavior. The functioning of a democracy essentially requires that people can form authentic political opinions to choose their policymakers. If individuals desire to give up control and let themselves be swayed by others interests, it distorts not only their own political decision-making but also the collective interest in a fair democratic system, in a sense. From a democratic perspective, this common interest should override individual decisions to give up control over political choices. In this light, ***individual consent*** to data processing for political micro-targeting **would not be an adequate mechanism** to avoid the anti-democratic threats posed by algorithmic political manipulation. This thesis argues that regulation of micro-targeting through data protection norms and consent is inherently limited, requiring *complementary approaches* that focus on the prevention of political manipulation. The legal challenge of defining a more suitable approach to this question is examined in Chapter 4.

4. Regulatory approaches for political micro-targeting: Building legal solutions

In Chapter 3, it is argued that the existing EU norms applicable to political micro-targeting are a useful mechanism for giving individuals more control over their data. Notwithstanding, a closer look at the EU data protection legal framework reveals that these rules are not enough to prevent the manipulation of individuals and the harms caused by this practice. In this chapter, legal solutions are proposed to regulate political micro-targeting with the aim of preventing manipulations of the electoral processes.

Primary EU legislation grants to personal data the status of a fundamental right, placed, therefore, out of the trade. This legal position constrains, for instance, the possibility (in Europe) of someone legally buying voters' data (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019), which could be used for profiling and micro-targeting purposes. The GDPR, for its part, empowers data subjects with a range of rights, such as access to personal data, rectification, and erasure. This can offer transparency and control over personal information and play a role in limiting the amount of data collected and processed for political targeting. Finally, under the GDPR, data controllers must find a lawful basis for using personal data for political purposes. In this regard, the informed and freely given consent is one of the most common legal grounds. Controllers are obliged to provide full information about which data is collected, for which purposes and have extra obligations such as giving the chance to revoke consent (GDPR.eu, n.d.). Such transparency requirements can benefit not only the data subjects but also enable societal control about political campaign strategies and marketing.

The drawbacks of the existing EU data protection norms in preventing the use of micro-targeting techniques for undesirable political purposes relate to heterogeneous factors. For instance, the **wide** privacy and data protection scope of the GDPR (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019). On one side the regulation brings up detailed rules regarding the processing of personal data, obligations of data controllers, and rights of the data subjects. On the other side, these norms are not specifically directed to micro-targeting. This non-specific approach makes it harder to prevent the harmful uses of personal data for steering individuals' political will by applying the GDPR.

Another weakness of EU data protection legislation relates to information asymmetries, i.e., the circumstance that in the online environment people are often not effectively informed of the purposes of collection and processing of their data. For instance, for political advertising.

Last, but definitely not least, the political manipulation threat remains present even in a hypothetical situation of GDPR full compliance. Targeted messages could influence voters' behavior even in a scenario where the data subjects were adequately informed and not impelled – directly or indirectly – to agree to the processing of their data for targeted political advertising. Put directly, consenting with manipulation does not necessarily avoid it.

In short, the weaknesses of the current regulations can be summarized as:

- The data protection nature of the GDPR can only indirectly prevent the practice of political micro-targeting (by limiting the usage of personal data).
- The protection of data via informed consent is not effective in practice. Consent is not properly informed and is not explicitly requested for political micro-targeting.
- Even if properly obtained, informed consent enables political micro-targeting and subsequently voter's manipulation.
- The protection of personal data via informed consent is not enough to make individuals less susceptible to political micro-targeting.

The use of micro-targeting techniques to politically manipulate individuals is on a collision course with democratic standards and the principle of human dignity. Considering the shortcomings of European data protection regulation to prevent these distortions, this chapter proposes regulatory approaches that specifically address the human and democratic issues of political online micro-targeting. These are:

(i) **Transparency approaches** in digital political campaigns, ideally combined with a reinforced consent perspective. The first transparency approach proposed connects with providing detailed and easily understandable information about targeted political ads to the concerned subjects. The second one relates to the duty of making political ads available to the public.

(ii) **The prohibition of audience targeting in elections**, banning the act of political micro-targeting itself.

The ***pros*** and ***cons*** of these approaches are explicitly assessed in the following sections.

4.1 Transparency guidelines for political micro-targeting

Political micro-targeting disrupts fundamental rights and social values in multiple ways. From an ***individual perspective***, it causes two kinds of damages: it allows undercover interference in someone's political will, jeopardizing human autonomy; moreover, micro-targeting violates the fundamental right of the non-targeted electors to receive complete information about the candidates and parties in the electoral dispute (Bayer, 2020). In both cases, the targeting practice is deeply problematic for depriving citizens of the right to make ***free and informed political decisions***, either by attacking voters' emotional weaknesses upon the use of personal data or by hiding pieces of political information from them.

The manipulative use of political micro-targeting jeopardizes ***the principle of human dignity*** in line with the Kantian second formulation of the categorical imperative. According to this ‘formula of dignity’ human beings should never be treated as a means, but always as an end in themselves (Andorno, 2009). This emphasizes the intrinsic worth of human beings and leads to two distinguishable and complementary aspects to ensure respect for human life: a subjective and an objective aspect. The ‘subjective’ dimension of human dignity corresponds to the respect for human freedom of making autonomous choices (Andorno, 2009). Political manipulation through micro-targeting violates this dimension of human dignity by diminishing human autonomy to make judgments about political parties and candidates.

From a ***collective perspective***, the practice of political micro-targeting harms ***democracy***, considering no one should be capable of taking complete control over anyone else’s political decision-making processes. Democracy presupposes individuals not being converted into an extension of another person's will. It also fragments the public discourse by hampering access to a shared information foundation composed of multiple perspectives (Bayer, et al., 2019). Finally, it harms the fairness of elections, for favoring the political actors that have most (financial) resources to exploit this tech-based advantage, possibly enabling a disproportionate influence on public opinion.

Authors stress that the European approach to the challenges of micro-targeting, political manipulation and disinformation requires further development, ‘primarily to include additional layers of transparency’ (Nenadić, 2019). For instance, during election campaigns that use audience targeting tools from social media platforms. In this light, transparency duties in political advertising could give back to the power of choice to individuals in electoral processes, rebalancing the equation of freedom of speech of political actors and the voters’ rights to autonomy and political self-determination. A study requested by the European Parliament on the impacts of disinformation and propaganda on the functioning of the rule of law in the EU and its Member States acknowledges that:

> ‘the act of micro-targeting – without the knowledge and understanding of the targeted individual, and informed and freely given consent – violates human dignity and the right to freedom of (truthful) information, and it destroys public discourse’ (Bayer, et al., 2019).

This Thesis proposes three transparency measures to tackle the harms caused to democracy and human dignity, discussing the advantages and drawbacks of each of them:

- The first one refers to imposing transparency duties in political targeted advertisements, aiming to restrain the manipulative power for those who receive targeted propaganda.
- The second comprises a prohibition of hidden ads in political processes. This is focused on the violation of the right to information of those groups of citizens unaware of the content their fellow citizens are exposed to (Bayer, et al., 2019).
- The third one, complementary to the others, refers to a reinforced consent for data processing for political micro-targeting purposes.

4.1.1 Transparency obligations in targeted ads

Micro-targeting is implemented through a range of actors in a complex digital advertising ‘ecosystem’. These actors might be social media networks, advertising platforms, data analytic services, and others. Social media companies can make, for instance, a vast amount of data available for political advertisers. These can provide additional data to the existing dataset[13] to profile and target audiences through data analytics. After testing variations of messages and defining the layout and content that maximizes engagement, the tailored messages are often placed as sponsored content in digital platforms, [14] or as banners on websites to reach the targeted audience.

Considering this multi-player advertising environment, a conceivable approach to limit the aggressive informational practice of micro-targeting is to **create a transparency rule prescribing that any political advertisement specifically targeting a person must plainly and clearly declare it**. Rather than banning the practice of micro-targeting itself, this approach focuses on increasing the autonomy of citizens for participating in the collective sphere. This perspective aligns, thus, with the notion of human dignity as empowerment.

In addition to clearly stating whether a political advertisement promoted in the online environment is targeted or not, politicians and social media should enable the targeted person to obtain **detailed information on micro-targeting**. In line with the EDPB Statement 2/2019 on the use of personal data in political campaigns, the core idea is that any targeted political advertisement needs to provide adequate information, for instance, on ***why*** that individual is receiving a certain message, ***who*** created the ad, ***how*** the ad was promoted, and ***how*** the person can exercise their rights as data subjects (European Data Protection Board, 2019). From this point of view, any person should be able to click on a button and find out that an advertisement has targeted them based on a specific set of online preferences (personal data collection) or because an automated process considered they should fit a set of characteristics (profiling).

[13] Whether directly collected from individuals, by technologies that track user online activity, or from third parties as data brokers, data marketing services, or online campaigning platforms. See (Bayer, et al., 2019).

[14] These include social media platforms, or other digital platforms, like Google and YouTube. Facebook, Google, Twitter, and Snapchat are most heavily used for political advertising. See (Bayer, et al., 2019).

While the consequences to the public are unpredictable, such a measure could increase public ***awareness*** about politically manipulative strategies in election campaigns. Beyond that, one can also imagine that the experience of facing plain information about the use of personal information, and automated profiling could be overwhelming. Individuals could feel disturbed when confronted with the amount of personal data collected with their given (informed) consent (or not), or with the categories in which they have been profiled, something that not necessarily would make someone proud or happy. Political advertising would no more claim to be an 'objective truth' but be presented as a 'personal truth', or better saying, a 'personalized truth'. This level of transparency could be a game-changer in the digital political campaigns.

Eventually transparency obligations could be extended to platforms providers, which should be responsible for 'informing users about which of their data are used for content selection or micro-targeting, and offering them the chance to exclude some (or all) personal data from this process' (Bayer, et al., 2019).

4.1.1.1 The pros of *transparency obligations in targeted ads*

This approach goes in line with the subjective dimension of the principle of human dignity, as discussed before, according to which individuals should have the autonomy to make choices. It also aligns with recent recommendations of the European Commission in the sense that European member states should encourage the disclosure of information concerning any targeting criteria used in the dissemination of paid online political advertisements (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019). In the Netherlands, the Dutch government has proposed a new Political Parties Act, recommending new transparency obligations for political parties regarding digital election campaigns and micro-targeting (van Hoboken, et al., 2018).

Another advantage of this perspective relates to its compatibility with the liberal, market-based, general philosophy of the GDPR, according to which, people should be informed and empowered to make their own choices. Unlike a more restrictive approach banning micro-targeting and prohibiting consent for such purposes, imposing a transparency obligation in this fashion would not disturb the idea that voters are autonomous free actors. More importantly, it would **not compromise** the fundamental right to freedom of speech of political parties or any agent promoting political ads, which are also protected by the European human rights framework both in Article 11 of the EU Charter of Fundamental Rights, and Article 10 of the European Convention on Human Rights (ECHR). Furthermore, this transparency obligation can and should be combined with the obligation of making political online ads available to the public and of reinforced consent, which are discussed in Sections 4.1.2 and 4.1.3.

Although transparency cannot guarantee that people will not be swayed by the granularly personalized political content, it would allow individuals to critically reflect on their fragilities and blind spots. In addition, unveiling the mechanisms of political ad personalization could also serve as a starting point for a critical public debate on the topic. Dan Ariely demonstrates in his research on psychology and behavioral economics how human behavior can be predictably irrational,[15]stating that people should become aware of what makes them susceptible to irrational thinking to fight it (Ariely, 2009). Applying this reasoning to the case of political micro-targeting, people should be conscious of what leads to irrational political behaviors and makes them vulnerable to manipulation. In this sense, the access to transparent information on the practice of micro-targeting could contribute to reducing the irrational consequences from it.

The discussed transparency obligations would, finally, be an interesting, flexible, and future-proof approach for regulating micro-targeting in digital campaigns. By imposing the detailed description of targeting criteria, the rule could be applied to all kinds of technological means for targeting voters. To illustrate: imagine a political ad using 'if… then' campaigns in the same manner as the ones created by e-commerce. This means that if a user clicked on an ad of a certain candidate, then they would see another ad in favor of this same candidate, and so forth (Vogels, 2016). Here, even though the targeting technique differed from the one based on the collection of personal data, profiling, segmenting, and targeting, people would still have to be informed of why they have received those specific political ads.

4.1.1.2 The cons of *transparency obligations in targeted ads*

A downside of this transparency approach is that in a similar way as consenting to micro-targeting does not remove the manipulation risk, being aware of the targeting criteria and manipulative uses of your data does not guarantee that individuals will not be subconsciously influenced by it. A study on the effects of personalized advertising on Facebook shows that, while the awareness about the personalization of a political ad made voters less likely to distribute and share the post, 'the perceived trustworthiness of the political party that disseminated the Facebook post did not appear to be affected' (Kruikemeier, Sezgin, & Boerman, 2016). Interestingly, it also shows that information on the usage of personal data has not significantly altered the user's responses to the political ad (Kruikemeier, Sezgin, & Boerman, 2016). The correlation between people's awareness about political micro-targeting methods and its manipulative effect on voters' decisions still requires further investigation. Nevertheless, the existing findings seem to indicate that even with all the information available, individuals could still be influenced in their voting intentions.

[15] By conducting a series of experiments, the author illustrates how our choices are, more often than we could expect, naïve, based on first impressions, or random. In this vein, understanding our irrational behavior could be a starting point for improving our decision making. Ariely proposes that *'in terms of our personal lives, we can actively improve on our irrational behaviors. We can start becoming aware of our vulnerabilities'*. See (Ariely, 2009)

Also pertinent in this regard is the analogy with the so-called non-deceptive placebos. It is widely believed that placebo treatment requires concealment or deception of the patient to produce significant effects (Kaptchuk, et al., 2010). Nonetheless, studies have shown that placebos pills administered without deception, i.e., with the patient's knowledge that they were not taking real medicine, might as well be an effective treatment and considerably improve patient's symptoms (Kaptchuk, et al., 2010). In those cases, neither expectation of improvement nor an association between getting a pill and getting better played a role (Gimlet, Science Vs, 2019). These studies have indicated that there should be new subconscious ways through which the placebo effect might work, unveiling unknown aspects of the human mind (Gimlet, Science Vs, 2019). Bringing these findings for the discussion on political micro-targeting, it is also unclear how the subconscious mechanisms of the human brain can work. Similarly to non-deceptive placebos, being conscious that a political ad was tailored according to intimate preferences and personality traits and can be manipulative does not necessarily mean that the personalized ad will not sway the targeted individual entirely or to some extent. The mysteries of the human brain are certainly far to be completely understood.

Another drawback relates to the fact that many individuals might not be interested in checking this kind of information upon receiving a political ad. Regardless of the conceived disadvantages, this approach would provide a vast dataset and elements for empirical research on its effectiveness in the political process. As a suggestion, this approach could be provided for in experimental legislation, as a temporary norm to test its effects on practice. In case of a positive outcome, the temporary norm could be converted into permanent legislation.

4.1.2 Publicizing hidden political ads: exposing 'dark posts'

Besides the manipulative effects produced by political micro-targeting, this practice also limits the audience of the campaigning content sent only to targeted groups (Bayer, 2020). The flip side of political campaigns directing tailored advertising to targeted groups is that non-targeted voters are excluded from 'political communication that is supposed to be public and inclusive in a democracy' (Bayer, 2020). This restriction of political content violates, at once, informational rights of the non-targeted citizens and the collective right to the public discourse, which harms the democratic process (Bayer, 2020). The right to receive and impart information and ideas is, furthermore, recognized by Article 10, ECHR, as a component of the right to freedom of expression.

As shown by the Study requested by the European Parliament, one can imagine local voters being targeted with two series of posts on social media. While national minorities were exposed to messages calling on them to affirm their power, the rest of the electorate would see posts focusing on the importance of a dominant nation (Bayer, et al., 2019). Although not presenting false facts, both messages would not be susceptible to proof (Bayer, et al., 2019) and would disseminate divisive content, which illustrates the manipulative characteristic of this informational practice. In the 2016 US Presidential Elections, Donald Trump's campaign was also known to use sponsored Facebook posts visible only by users with specific profiles. The so-called '**dark posts**' were used to micro-target groups of voters, including over 40,000 variations of ads per day.

As mentioned in item 2.1, the damaging impacts of micro-targeting in election processes were empirically demonstrated by a study on the effectiveness of micro-targeting. Evidence shows that a candidate is likely to lose support when voters outside the targeted group can see the message (Hersh & Schaffner, 2013). This indicates that ***limiting political content to targeted groups presumably affects electoral processes***. If political messages were available for the public, the results would probably be diverse.

Considering this scenario, a second transparency approach, not opposed but complementary to the one presented in Section 4.1.1, relates to imposing **an obligation in digital campaigns that micro-targeted political ads are available to all groups of interest in the online platform used to share it**. According to this rule, although political campaigns could still target certain audiences, they would no longer be allowed to hide political ads from other groups of interest. In practical terms, a regulation could establish, **first**, that all political advertisers making use of digital means such as websites and social media platforms for sending micro-targeted messages to a group of electors would have the legal obligation to **make the targeted ads publicly available**, in that same digital mean, to any other group of voters. To ensure complete transparency, a **second** obligation for the advertisers would be to **list the criteria for deciding to whom the ads get sent** when publishing the micro-targeted ad. As pointed by scholars in the field, 'so far, platforms do not give much information on how and to whom political ads are targeted' (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

Under this framework, the use of Facebook 'dark posts', for instance, would not be banned, but their content would have to be publicized on the official page of the candidate or political party that sent the message to a particular group. To enforce this sort of regulation, websites and social media platforms could also be accountable for not making political micro-targeted ads available with these requirements. Although micro-targeted political messages could still be sent to particular groups, the availability to the general public would take away its exclusive, hidden, or 'dark' feature.

4.1.2.1 The pros of *publicizing hidden political ads*

The approach advantage refers to implementing the ***voter's right to receive information*** about political parties and candidates. Moreover, it would establish in practice ***a right to know that people and voters are differently targeted***. By requiring targeted political ads to be available to the electorate, the rule would ***favor public control*** of political campaigns and advertisements disseminated online. Individuals, other parties, and the media would be able to verify what content was being sent exclusively for which groups and to identify the targeting strategies of political actors. More than acknowledging how individuals are targeted, society would have information on micro-targeting political strategies and enable it to keep a record. This transparency requirement would contribute, furthermore, to ***attract media's attention*** on the political micro-targeting issue.

The requirement to disclose that ads were sent exclusively for certain demographic groups, capitalizing on implicit biases of these audiences (Agarwal, 2020), would also ***bring to light on what prejudices*** political candidates and parties play. Examples of biased micro-targeted political ads, such as racist and anti-minorities messages aiming only nationalist and white audiences can be only imagined today. Information on the real cases is currently scarce precisely by the lack of transparency on how political micro-targeting is implemented. Revealing what targeted messages are sent to those groups would bring another level of information to the democratic debate and eventually reveal an 'iceberg' of distasteful campaign methods and messages of which we can only see the tip. As quoted by Isaac Newton and popularized by the German tv series 'Dark': 'what we know is a drop, what we don't know is an ocean' (Tüzemen, 2020). The publicizing of the micro-targeted messages with information on the aimed groups would likely raise serious concerns for political actors making use of this campaign strategy. This could fundamentally change the practice of micro-targeting.

The approach would favor the ***collective accountability of political practices*** in democratic regimes.

4.1.2.2 The cons of *publicizing hidden political ads*

Weaknesses of the proposed approach, common to all transparency approaches, relate to the remaining possibility of political advertisers to send micro-targeted political ads. While access to information would increase, the political will of targeted groups could still be influenced. Once again, further research could help to elucidate the effectiveness of this approach in practice.

4.1.3 The refined consent approach

Complementary to the other transparency approaches discussed so far, the reinforced consent refers to obtaining the GDPR notion of meaningful consent specifically for the case of political advertising. Based on articles 9 and 22 of the GDPR, this thesis upholds that the consent given by data subjects for the collection and the processing of personal data for political micro-targeting should be ***explicit***.

Upon Article 9, GDPR, it is argued on item 3.1.3 that any data used for micro-targeting purposes should be considered **sensitive**, by definition. The data used for micro-targeting can reveal political opinions, racial or ethnic origin, religious or philosophical beliefs, or act as a proxy to these pieces of information. As a consequence, the **use of**

personal data for political micro-targeting would require that the data subject has given ***explicit* consent** to the processing for this specified purpose.

The rule set out in Article 22 of the GDPR reinforces the need for explicit informed consent for the processing of personal data for political micro-targeting. Under that regulation, when a decision based solely on automated processing produces legal effects for the data subject or significantly affects them, it is only regarded as lawful upon explicit consent. This clearly includes profiling. In this regard, affecting a person's vote in an election should be considered as a legal effect generated by automated decision-making (European Data Protection Board, 2019). The European Commission has also recognized that:

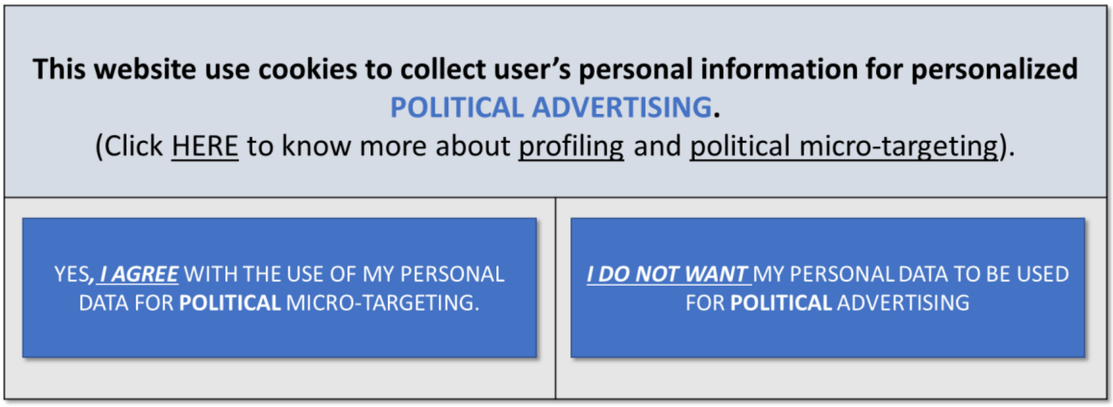
> 'given the significance of the exercise of the democratic right to vote, personalized messages which have for instance the effect to stop individuals from voting or to make them vote in a specific way could have the potential of meeting the criterion of significant effect' (European Commission, 2018).

A study requested by the European Parliament also supports the view that profiling connected to targeted campaign messaging should be considered as a solely automated decision-making that produces significant effects on individuals (Bayer, et al., 2019). Accordingly, it argues that where micro-targeting is based on collected, observed or inferred special categories of personal data, it should only be allowed based on the informed, explicit and freely given consent by the individual, or where significant public interest based on EU or national law merits so (Bayer, et al., 2019).

This perspective goes in line with the EDPB view that the minimum requirement for the act of micro-targeting should be the opt-in explicit consent by the user with a possibility to unsubscribe (European Data Protection Board, 2019). The explicit consent would require, moreover, an ***explicit statement*** of the data subject (European Data Protection Board, 2020). For that, this approach suggests that data controllers should show the data subjects an **independent consent notice** bringing specific and unambiguous information on the collection of personal data and the use of algorithmic profiling to personalize political advertising. In this separate electronic form, users would be presented two evidently clear options to agree or not agree with that collection and processing of their data. By this approach, any collection or processing of personal data for political micro-targeting could only be allowed upon this ***specific*** and ***explicit*** opt-in procedure. As a consequence, any omission of consent should lead to the prohibition of data usage for this end, implementing the logic of data protection by default.

The proposed ***reinforced consent approach*** would require, therefore, that consent for using personal data to tailor political ads is ***distinguished*** from other targeting criteria. Through this logic, data subjects could easily set their preferences, saying 'no' to the use of personal data for political micro-targeting and 'yes' to other targeting criteria they do not consider prejudicial. For instance, data analytics (used to improve the website), or personalization (used to adapt the website content according to the user interests).

Another relevant condition relates to the ***standardization*** of the consent banners model to all personal data processing. Although each website or digital platform could still personalize the banner with their design preferences - for instance, with the website or company colors, logo, or font - the basic content should be essentially the same, for the sake of transparency. Instead of endless ways of requiring the consent of data subjects as it verifies in the present online environment, data processing actors would be obliged to inform data subjects following the same structure and avoiding the risks of unclear messages – intentionally or not. A suggestion of a consent banner for providing data to be used in political micro-targeting is shown in **Fig. 5**.



This website use cookies to collect user's personal information for personalized POLITICAL ADVERTISING.

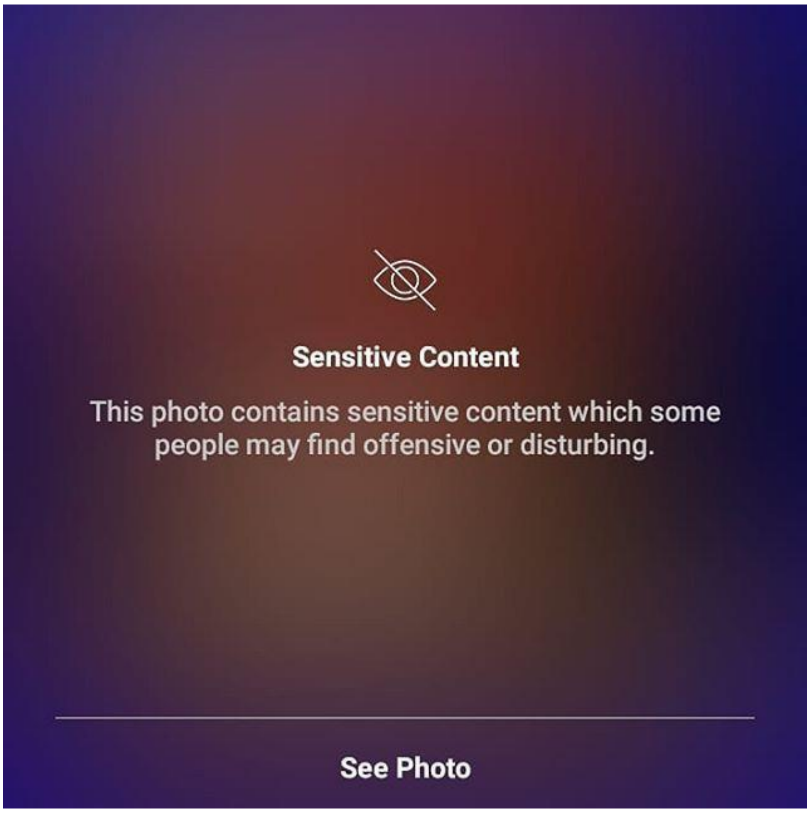
(Click HERE to know more about profiling and political micro-targeting).

YES, ***I AGREE*** WITH THE USE OF MY PERSONAL DATA FOR **POLITICAL** MICRO-TARGETING.

I DO NOT WANT MY PERSONAL DATA TO BE USED FOR **POLITICAL** ADVERTISING

Figure 5: A suggestion of a model for consent banners following the reinforced consent proposed approach. In addition to being independent to cookies consent banners and requiring explicit consent for political purposes, the data subject would be able to request more information about profiling and micro-targeting. It would be required by the regulation to clearly present this information, explaining, for instance, ***which*** *personal data could be used to infer* ***what*** *sort of characteristics and* ***how*** *this could be used to politically target voters.*

The logic of meaningful, reinforced consent could also be applied to the ***visualization*** of micro-targeted ads. This could be implemented in the same manner as in Facebook and Instagram sensitive posts which first appear blurred with a warning screen stating that the content may be disturbing or sensitive (Facebook, 2020) (see **Fig. 6**). Following this suggestion, political micro-targeted ads could have their visualization initially restricted to the (targeted) audience with a warning that it is a (personalized) political ad and allowing the user to decide whether to reveal its content or not. For instance, "this advertisement was directed to you based on your personal information and online preferences. Click here to know more details or here to view it". The visualization would require, therefore, a distinct action of the user, confirming their will to see that ad, and at the same time preventing subconscious effects from accidental visualizations of the ad. This measure could contribute to building a much safer political ad environment in social media platforms.



Sensitive Content

This photo contains sensitive content which some people may find offensive or disturbing.

Figure 6: Example of Instagram warning screen for possible sensitive or disturbing content for users. Source: Instagram app. One possible implementation of the reinforced consent approach would be to apply this kind of screening to political ads, requesting a definite consent to view political ad, micro-targeted or not.

4.1.3.1 The pros of the *refined consent*

Instead of a legislative proposal, the reinforced consent approach offers guidelines for the practical implementation of the rules already provided by the GDPR, which prioritize individual autonomy in the online environment. Furthermore, it does not exclude other transparency duties in political ads and campaigns, being complementary to them. Even if people explicitly consented to the use of their data for political purposes, there would still be a right to receive information on targeting criteria used in political ads (Section 4.1.1) and a right to know what different messages are sent to targeted groups (Section 4.1.2).

4.1.3.2 The cons of the *refined consent*

In the same direction of the criticism directed to the transparency approaches and the GDPR consent-based model, the fact of giving fully informed consent to political micro-targeting does not necessarily avoid the manipulative power embedded in it. In this vein, the effectiveness of this approach to avoid political manipulation would require confirmation through empirical studies on voters' behavior. Notwithstanding, even if this viewpoint is not enough to eliminate the threat posed by micro-targeting, it makes the requirements for the practice of micro-targeting stricter based on the current data protection regulation. It strengthens, therefore, transparency duties in micro-targeted political campaigns.

4.2 The restrictive approach: Banning audience targeting

A common weakness of the proposed transparency approaches refers to the persisting threat of political manipulation. Even if a person were clearly and fully informed, first, when consenting to the use of their data for political micro-targeting and, later, when receiving a targeted message, the manipulative effect of this practice on the subconscious level of the human mind can still be present. Moreover, at least in theory, a person could deliberately accept to become an instrument of others' will by desiring to receive micro-targeted political propaganda. By this line of argument, one could argue that a more restrictive regulatory approach, encompassing a full prohibition of micro-targeting, could be required to prevent the manipulative dangers raised by it. A regulatory approach opposed to the transparency ones would be to **establish a legal norm prohibiting audience targeting in elections**.

The ***objective*** dimension of the principle of human dignity could serve as a justification for legislation that prohibited consent for political micro-targeting purposes and banned the practice of micro-targeting. 'Dignity as constraint' is interpreted in law as a requirement of protecting people against degradation, even if the subjects agree with their objectification and violation of dignity (van Beers, 2012). In brief, this means that individual freedom can be restricted to ensure respect for human dignity. It could be argued that as an implication of the democratic principle, people should not be subjected to techniques with increased manipulative power, such as micro-targeting. Another possible ground for prohibiting audience targeting in elections could be the legal provision of a ***right to be unpredicted*** by technological means as algorithms, considering that automated decision processes are often considered 'black boxes'.

4.2.1 The pros of *banning*

The main advantage of a legal prohibition and enforcement of audience targeting would be the prevention of political manipulation and democratic damages through micro-targeting.

4.2.2 The cons of *banning*

Although the approach is conceivable, one drawback relates to a more severe restriction of political freedom of speech of political parties, which is also provided for as a fundamental right in the European human rights framework. A second negative aspect refers to its difficulty to find a proper definition on what to ban.

Regarding the first aspect, a consistent argumentative effort, possibly higher than to justify transparency approaches, would be required. Nevertheless, the jurisprudence of the ECtHR signals that it

> '(…) will accept, in some circumstances, that outright bans on political advertising may be consistent with freedom of expression, in order to prevent the risk of distortion of public debate by wealthy groups' (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

In Animal Defenders International v. the UK (Animal Defenders International v. UK, 2013), the court held that a ban on paid political advertising on television did not violate article 10, ECHR. The case involved an animal rights group seeking to broadcast a political ad outside an election period, and not a political party. However, the judgment provided judicial guidelines for 'hard cases' affecting freedom of expression, which could apply to an eventual law prohibiting political micro-targeting. In this occasion, the court has recognized that the prevention of **public debate distortions exerted by groups with unequal access to political advertising**, and the 'immediate and powerful effect of broadcast media' were acceptable legislative choices underlying the ban. It also considered the quality of parliamentary review of the measure was adequate and that there was a reasonable risk of abuse from wealthy political bodies if the ban was relaxed (Dobber, Fathaigh, & Zuiderveen Borgesius, 2019).

Regarding the issues in defining the problem, it is difficult to establish what criteria (if any) would be acceptable for ad audience targeting. Serving ads based on public voter records and general political affiliations (left-leaning, right-leaning, and independent), as Google used to do in the U.S. (Google, 2019) could be said to offer **more clear political manipulation risks**. Conversely, the manipulative impact of general categories as age, gender, and location (postal code level), as currently admitted by Google terms of service (Google, 2019) would be more controversial. The underlying question here is: when does political targeting become 'micro' and illegal? It could become common sense that sending targeted discriminatory or defamatory content to a group susceptible to this content would be prejudicial for democracy and therefore, be banned. But on the other hand, people could disagree whether targeting a female audience with political ads regarding policies aiming at gender equality would be problematic for democracy. Many people could not see a big democratic change in this kind of targeting. This illustrates some challenges in implementing a broader micro-targeting prohibition.

5. Conclusion

Technology is a human tool embedded in social relations in all its complexity. As so, the use of innovations of any sort is not neutral but guided by economic and political interests. Inevitably, the technology's application requires ethical, human rights and impact assessments to ground regulations. When it comes to politics and electoral processes, profiling algorithms, fueled by personal data of internet users, have been used with the purpose of micro-targeting voters in election campaigns. This thesis has shown, in Chapter 2, that this aggressive technique for the dissemination of customized political advertisements has been employed in recent election processes. The manipulative potential of this granular tech-oriented identification of personality traits raises serious concerns about human autonomy and the fairness of elections. As mentioned in Section 2.2.2, political micro-targeting can delude internet users into believing that they receive information universally encountered by other people, when in fact, it is specifically directed to them. Moreover, as discussed in Chapter 2, scientific studies show that this technique is highly effective in obtaining political gains. Although more research is needed on this matter, the ongoing technological developments in the field reinforce the idea that the precision of algorithmic profiling can only increase with time. That is to say, if the technique is not yet perfect, it can only become more effective in the coming future. In this context, the 'Collingridge Dilemma' illustrates the hardships of regulating new technologies between its ideation and adoption. Whereas in its first stages of technology development the future effects are less predictable and uncertain, with time technology becomes deep-rooted in society, making regulation and the prevention of undesired societal damages harder.

The political manipulation through micro-targeting endangers human autonomy to make independent political choices and the subjective dimension of the principle of human dignity. This also threatens democracy, which requires the human mind not to be influenced on a level that it is swayed, or 'hacked' by undercover political interests. The practice of sending targeted political ads that can only be seen by the voters more susceptible to it violates the voter's right to have access to full information about political parties and candidates, distorting the political debate as a whole. In this sense, political micro-targeting disrupts the citizens' democratic right to make a free and informed political choice. Finally, micro-targeting can unlawfully favor groups of interest that can afford to make use of this manipulative strategy, unbalancing the fairness of competition between political parties in electoral campaigns.

Having established the phenomenon of political micro-targeting and the threats it poses for democratic societies; Chapter 3 examines the existing norms applicable to this practice in the EU. Considering that the practice is based on the collection of personal data to profile individuals and target them with political personalized ads, the European data protection legislative framework applies to the case. By identifying the norms that could partially prevent the described issues, it concludes that the EU data protection norms ***can help*** individuals to have ***more control over their data*** and limit the collection for any purposes, including micro-targeting. Nonetheless, in the way the regulations are currently designed and applied, they are ***insufficient to prevent the manipulation*** that can be accomplished by political micro-targeting.

Data protection norms are helpful to obstruct the monetization and commercialization of personal data, which possesses the status of a fundamental right in the European legal regime. It also limits the collection of personal data and empowers data subjects with the rights of access, rectification, and erasure of personal data. Finally, it establishes a consent-based model for the collection and processing of personal data, which, at least in theory, can increase transparency and limit the use of personal data for micro-targeting purposes.

Some of the current regulatory issues concern the fact that, in practice, people seem not to be adequately informed about the uses of personal data for political micro-targeting purposes. In addition, the manipulative power could persist even if people were aware and could give free, meaningful consent to micro-targeting. In this way, the application of the GDPR consent-based mechanism, focused on the data protection of individuals, presents limitations to regulating the democratic threats presented by political micro-targeting.

Motivated by the issues presented in Chapters 2 and 3, Chapter 4 proposes novel approaches to prevent political manipulation and distortion of the public debate caused by micro-targeting, discussing advantages and drawbacks of each one. I sum these up in **Table 1**. The first legislative suggestion refers to ***transparency approaches (Section 4.1)*** to political advertising in the online context. An alternative legislative policy is the ***prohibition of audience targeting in digital electoral campaigns (Section 4.2)***. The core advantages of transparency approaches refer to enabling voters to make informed political decisions. Three transparency approaches are proposed. They combat the misinformation produced by political micro-targeted political ads on different levels.

The ***first transparency approach (Section 4.1.1)***, requires detailed information in micro-targeted advertisements, in order to make voters aware that they are receiving personalized political ads and how they were profiled. It focusses, therefore, in reducing the manipulative potential of micro-targeting through awareness. It does not disturb, however, either the notion of political freedom of speech of political agents or the liberal GDPR conception of free-will to make informed choices.

The ***second transparency approach*** consists of the disclosure of micro-targeted ads to other groups of interest, providing transparency about political platforms in election campaigns and empowering voters to make informed and ***authentic*** choices. More than that, it exposes how political forces target specific groups, how they try to persuade different audiences, and how they can abusively explore their characteristics and vulnerabilities. It unveils, therefore, strategies and bias of political parties and candidates, being a potential game-changer in electoral fairness. By potentially generating the loss of supporters and reducing the effectiveness of political micro-targeting, this approach could discourage the practice and even contribute to its end. Furthermore, it does not limit freedom of expression of political agents.

The ***third transparency approach*** encompasses two directives. First, it defines an explicit, separate, and standardized way of consenting to the collection of personal data for political micro-targeting purposes. Second, it requires a confirmation consent for visualizing micro-targeted ads in social media. These rules compose the concept of a reinforced consent which would also contribute to preventing the manipulation threats by offering users an effective choice not to be in contact with micro-targeted political ads.

This Thesis claims that the transparency approaches should, ideally, be combined to optimize the protection of the fundamental rights of voters and democracy. The major disadvantage is not entirely eliminating the risks of subconscious political manipulation and greatly relying on the interest of individuals in being informed. In this regard, studies to assess the effectiveness of these measures to prevent political manipulation would be required. Eventually the promotion of public debate about political micro-targeting is an indispensable measure to fight misinformation and deception in political campaigns.

Considering the uncertain potential of the transparency approaches to defeat political manipulation, ***the restrictive approach of micro-targeting*** is posed as an alternative solution. If successfully enforced, the model of prohibiting the use of audience targeting in society could directly eliminate the manipulation and democratic threats of political micro-targeting. The shortcomings relate, however, to the more severe interference with the fundamental right of freedom of expression of political agents. Although surmountable, the ban of political micro-targeting would probably require a higher argumentative effort than the implementation of transparency approaches. To justify the restriction of targeted political ads, policymakers would need to gather consistent evidence of how the practice of political micro-targeting substantially harms democracy in concrete election processes. This Thesis offers some pieces of information about the damaging potential of micro-targeting. However, in-depth case-studies would be also needed to ground a more restrictive regulation. The restrictive approach also offers hardships in the definition of when the practice of targeting voters becomes 'micro' and, therefore, illegal. A public debate would be necessary, in this regard, to establish whether political micro-targeting should be allowed, for instance, when based on age or gender of voters. All these factors considered, the banning approach seems harder to implement compared to transparency approaches.

This Thesis supports that both the combination of the **three transparency approaches** are an ideal starting point to regulate this issue, but a **banning approach** should not be ruled out. Nonetheless, a choice between them inevitably requires a balance of interests by legislators and policymakers. In this regard, the use of experimental legislation to test the efficacy of the proposed approaches might consist of a short-time solution for the urgent matter of political manipulation through micro-targeting. Transparency approaches should be adopted as a first attempt to solve the existing human autonomy and democracy issues, given the possibly easier legislative implementation, before the formulation of a more restrictive solution.

Table 1: pros and cons of regulatory approaches for political micro-targeting

Transparency approaches (complementary)		
Transparency obligations in targeted ads	**Key points**	Any political ad specifically targeting someone must: (i) plainly inform that it is a targeted message and (ii) provide details about the targeting (*why* the person was targeted, based on which information and automated decision-making, *who* creates and promotes the ad, and *how* data subjects' rights can be exercised).
	Pros	• Gives detailed and clear information about micro-targeting, increasing public awareness and possibly reducing manipulation risks. • Aligns with the GDPR liberal idea. • Flexible approach, applicable to other means of political targeting.
	Cons	• Does not completely eliminate the risk of political manipulation. • Relies on the individual interest in being informed and not being susceptible to political manipulation.
Publicizing hidden political ads	**Key points**	Targeted ads must be: (i) publicly available in the same digital means used to target groups and (ii) list the criteria for deciding who the ads get sent to.
	Pros	• Implements the voter's right to receive information about political parties and candidates. • Establishes a right to know that people and voters are differently targeted. • Favors the collective accountability of political practices. • Contributes to raise media attention on the issue of political micro-targeting. • Potentially unveils bias in political campaign marketing strategies.
	Cons	• Targeted groups could still be subconsciously influenced

Table 2: continued; pros and cons of regulatory approaches for political micro-targeting

Reinforced consent	**Key points**	Consent for the *collection and processing* of personal data for political micro-targeting should be specific, explicit, and separate. (i) Data controllers must show an independent and standardized consent notice with specific information on the use of personal data for political micro-targeting. No implicit opt-in. (ii) Require consent for the *visualization* of political ads. Ads could have their visualization initially limited to a warning that it is a (personalized) political ad. (iii) Visualization would require an express action of the user, confirming they want to see that ad.
	Pros	• Offers guidelines for the implementation of the GDPR existing rules. • Favors individual autonomy in the online environment. • Complementary to transparency duties in political ads and campaigns.
	Cons	• The GDPR consent-based model does not necessarily avoid the manipulation threat embedded in it.
Restrictive approach (alternative)		
Banning audience targeting	**Key points**	• Legal prohibition or restriction of audience targeting in elections.
	Pros	• Eliminates or restricts the practice of micro-targeting, preventing the manipulative dangers raised by it, not fully addressed by the transparency approaches.
	Cons	• Restricts freedom of expression of political parties and actors more severely. • Issues in the definition of what should be considered "micro"-targeting

Bibliography

Primary sources (legislation, case-law, and policy documents)

Legislation

Charter of Fundamental Rights of the European Union of 26 October 2012, 2012/C 326/02.

Convention for the Protection of Human Rights and Fundamental Freedoms of 03 September 1953, ETS No.005 (European Convention on Human Rights)

Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications).

Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation) of 27 April 2016.

Case-law

Animal Defenders International v. UK, Application No. 48876/08 (European Court of Human Rights April 22, 2013).

Policy documents

Article 29 Working Party. (2017, November 28). *Guidelines on Consent under Regulation 2016/679.* Retrieved from European Commission, Justice and Consumers: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051

Bayer, J., Bitiukova, N., Bard, P., Szakács, J., Alemanno, A., & Uszkiewicz, E. (2019, February 1). Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States. *SSRN*. doi:http://dx.doi.org/10.2139/ssrn.3409279
European Commission. (2018). Free and Fair elections. Brussels: European Commission.

European Comission. (2020). *White Paper on Artificial Intelligence A European approach to excellence and trust.* Brussels: European Comission Website.

European Commission. (n.d.). "Fake News" and Disinformation - Knowledge For Policy.
European Data Protection Board. (2019). *Statement 2/2019 on the use of personal data in the course of political campaigns.* EDPB.

European Data Protection Board. (2020, May 06). *Guidelines 05/2020 on consent under Regulation 2016/679.* Retrieved from European Data Protection Board Website: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-consent-under-regulation-2016679_en

European Data Protection Supervisor. (2018). *Opinion 3/2018 EDPS Opinion on online manipulation and personal data.* European Data Protection Supervisor. Retrieved from https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en .pdf

European Union Agency for Fundamental Rights. (2018). *Applying the Charter of Fundamental Rights of the European Union in law and policymaking at national level.* Luxembourg: Publications Office of the European Union.

Secondary sources (academic literature)

Absattarov, R. (2012). *Political Manipulation: Concept and Definition.* Retrieved from http://www.rusnauka.com/12_KPSN_2012/Politologia/3_107760.doc.htm

Agarwal, P. (2020). *Sway: Unravelling Unconscious Bias* . Bloomsbury Sigma.

Andorno, R. (2009, April 22). Human Dignity and Human Rights as a Common Ground for a Global Bioethics. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine Volume 34, Issue 3,* pp. 223-240. doi:https://doi.org/10.1093/jmp/jhp023

Ariely, D. (2009). *Predictability Irrational* . New York: HarperCollinsPublishers.
Barry, E. (2018, April 20). Long Before Cambridge Analytica, a Belief in the ‘Power of the Subliminal’. *The New York Times*. Retrieved from https://www.nytimes.com/2018/04/20/world/europe/oakes-scl-cambridge-analytica-trump.html

Bayer, J. (2020, March 31). Double harm to voters: data-driven microtargeting and democratic public discourse. *Internet Policy Review 9(1)*. doi:10.14763/2020.1.1460

Bennett Moses, L. (2013). How to Think About Law, Regulation and Technology: Problems with 'Technology' as a Regulatory Target. *(2013) 5(1) Law, Innovation and Technology,* pp. 1-20.

Bergemann, B. (2018). The Consent Paradox: Accounting for the Prominent Role of Consent in Data Protection In: Hansen, Marit Kosta, Eleni Nai-Fovino, Igor Fischer-Hübner, Simone (Ed.): Privacy and Identity Management. The Smart Revolution, ISBN 978-3-319-92925-5, Cham. *Springer International Publishing*, pp. 111-131. doi:http://dx.doi.org/10.1007/978-3-319-92925-5_8

Cudd, A. E., & Navin, M. C. (2018). *Core Concepts and Contemporary Issues in Privacy.* Switzerland: Springer.

Collingridge, D. (1980). *The Social Control of Technology.* London: Frances Pinter.

Davies, G. T. (2009, January 29). Law and Policy Issues of Unilateral Geoengineering: Moving to a Managed World. *SSRN*, p. 15. doi:http://dx.doi.org/10.2139/ssrn.1334625

Dobber, T., Fathaigh, R. Ó., & Zuiderveen Borgesius, F. J. (2019, December 31). The regulation of online political micro-targeting in Europe. *Internet Policy Review 8(4)*. doi:10.14763/2019.4.1440

Gorton, W. (2016, February 5). Manipulating Citizens: How Political Campaigns' Use of Behavioral Social Science Harms Democracy. *New Political Science: Vol. 38, No. 1*, pp. 61-80. doi:10.1080/07393148.2015.1125119

Greenwood , D., & Levin, M. (2007). *Introduction to Action Research.* Thousand Oaks, CA: SAGE Publications, Inc. doi:10.4135/9781412984614

Hanekamp, J. C., Vera-Nava, G., & Verstegen, S. W. (2007, February 18). The historical roots of precautionary thinking: the cultural ecological critique and 'The Limits to Growth'. *Journal of Risk Research, 8:4,* pp. 295-310. doi:10.1080/1366987042000265056

Helbing, D., Frey, B., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., . . . Zwitter, A. (2017, February 25). Will Democracy Survive Big Data and Artificial Intelligence? *Scientific American*.

Hersh, E. D., & Schaffner, B. F. (2013, April 9). Targeted Campaign Appeals and the Value of Ambiguity. *The Journal of Politics*, pp. pp. 520-534. doi:10.1017/s0022381613000182

Hoofnagle, C. J., van der Sloot, B., & Zuiderveen Borgesius, F. (2019, February 10). The European Union general data protection regulation: what it is and what it means. *Information & Communications Technology Law*, *28:1*, pp. 65-98.

Kaptchuk, T. J., Friedlander, E., Kelley , J. M., Sanchez, M. N., Kokkotou, E., Singer, J. P., . . . Lembo , A. J. (2010, December 22). Placebos without Deception: A

Randomized Controlled Trial in Irritable Bowel Syndrome. *PLoS One. 2010; 5(12): e15591*. doi:10.1371/journal.pone.0015591

Kruikemeier, S., Sezgin, M., & Boerman, S. C. (2016). Political Microtargeting: Relationship Between Personalized Advertising on Facebook and Voters' Responses. *Cyberpsychology, Behavior, And Social Networking Volume 19, Number 6*. doi:10.1089/cyber.2015.0652

Madsen, J. K., & Pilditch, T. D. (2018, April 10). A method for evaluating cognitively informed micro-targeted campaign strategies: An agent-based model proof of principle. *PLoS ONE 13(4): e0193909*. doi:https://doi.org/10.1371/journal.pone.0193909

Nenadić, I. (2019, December 31). Unpacking the "European approach" to tackling challenges of disinformation and political manipulation. *Internet Policy Review 8(4)*. doi:10.14763/2019.4.1436

Nickerson , D., & Rogers, T. (2014). Political Campaigns and Big Data. *Journal of Economic Perspectives Volume 28, Number 2,* pp. 51-74. doi:10.1257/jep.28.2.51
Ranchordás, S., & van 't Schip, M. (2019, August 9). Future-Proofing Legislation for the Digital Age. *SSRN*.

Rehman, I. u. (2019). Facebook-Cambridge Analytica data harvesting: What you need to know. *Library Philosophy and Practice (e-journal).*

Susser, D., Roessler, B., & Nissenbaum, H. (2019, June 30). Technology, autonomy, and manipulation. *Internet Policy Review 8(2)*. doi:10.14763/2019.2.1410
Tüzemen, S. (2020). *The Quantum and Cosmic Codes of the Universe.* Cambridge Scholars Publishing.

van Beers, B. (2012, May 19). TV Cannibalism, Body Worlds and Trade in Human Body Parts: Legal-Philosophical Reflections on the Rise of Late Modern Cannibalism. *SSRN*, pp. 65-75.

van Hoboken, J., Appelman, N., Ó Fathaigh, R., Leerssen , P., McGonagle , T., van Eijk , N., & Helberger, N. (2018). *De verspreiding van desinformatie via internetdiensten en de regulering van politieke advertenties.* Amsterdam: Instituut voor Informatierecht (IViR), Universiteit van Amsterdam.

Vosoughi, S., Roy, D., & Aral, S. (2018, March 9). The Spread of True and False News Online. *Science*, pp. 1146-1151.

Wilson, D. G. (2017, October). The Ethics of Automated Behavioral Microtargeting. *AI Matters*, pp. 56-64. doi:https://doi.org/10.1145/3137574.3139451

Witzleb, N., Paterson, M., & Richardson, J. (2020). *Big Data, Political Campaigning and the Law - Democracy and Privacy in the Age of Micro-Targeting.* Routledge.

Yeung, K. (2016, May 22). ‘Hypernudge’: Big Data as a Mode of Regulation by Design. *Information, Communication & Society,* pp. 118-136. doi:10.1080/1369118X.2016.1186713

Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computers Judge Personalities Better than Humans. *Proceedings of the National Academy of Sciences 112(4),* (pp. 1036-1040). doi:10.1073/pnas.1418680112

Zuboff, S. (2019). The Age of Surveillance Capitalism. London: Profile Books.

Zuiderveen Borgesius, F. J., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., . . . de Vreese, C. (2018, February 9). Online Political Microtargeting: Promises and Threats for Democracy. *Utrecht Law Review*, pp. 82-96. doi:https://doi.org/10.18352/ulr.420

Monographs

Curto, R. M. (2018). Political micro-targeting: a European information war story.

Plaizier, C. (2018, January). Micro-Targeting Consent: A Human Rights Perspective on Paid Political Advertising on Social Media.

Other sources (news articles, websites, reports)

AVAAZ. (2019). Far Right Networks of Deception. AVAAZ.

BBC News. (2018, April 09). Retrieved from BBC: https://www.bbc.com/news/av/technology-43674480/facebook-data-how-it-was-used-by-cambridge-analytica

Besti, F. (2019). *Self-driving society*. Retrieved from The Collingridge Positioning: https://selfdrivingsociety.fondazionebassetti.org/self-driving-society-design-tools/

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. Retrieved from The Guardian: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Chappell, B. (2018, April 10). How To Check If Your Facebook Data Was Used By

Cambridge Analytica. *NPR*. Retrieved from https://www.npr.org/sections/thetwo-way/2018/04/10/601163176/how-to-check-if-your-facebook-data-was-used-by-cambridge-analytica?t=1581873280955

Davies, H. (2015, December 11). Ted Cruz using firm that harvested data on millions of unwitting Facebook users. *The Guardian*. Retrieved from The Guardian: https://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data

Democracy Now. (2020, January 06). New Cambridge Analytica Leaks to Expose Election Manipulation in 68 Countries. *Democracy Now!* Retrieved from Democracy Now!: https://www.democracynow.org/2020/1/6/headlines/new_cambridge_analytica_leaks_to_expose_election_manipulation_in_68_countries

Facebook. (2020). *13. Violent and graphic content*. Retrieved from Community Standards: https://www.facebook.com/communitystandards/graphic_violence

Facebook. (n.d.). *The Conservative Party*. Retrieved from Facebook for Business: https://www.facebook.com/business/success/conservative-party

Fischer, W. (2019, August 18). We talked to the professor who fought Cambridge Analytica to get his data back in Netflix's 'The Great Hack' about why privacy rights in the US are lagging behind the rest of the world. *Business Insider*. Retrieved from Business Insider Nederland: https://www.businessinsider.nl/netflix-great-hack-david-carroll-interview-data-rights-cambridge-analytica-2019-8/

GDPR.eu. (n.d.). *GDPR checklist for data controllers*. Retrieved from GDPR.eu: https://gdpr.eu/checklist/

Ghosh, D., & Scott, B. (2018, March 19). Facebook’s New Controversy Shows How Easily Online Political Ads Can Manipulate You. *Time*. Retrieved from Time: https://time.com/5197255/facebook-cambridge-analytica-donald-trump-ads-data/

Gimlet, Science Vs. (2019, May 9). *Placebo: Can the Mind Cure You?* Retrieved from Gimlet, Science Vs: https://gimletmedia.com/shows/science-vs/5whgzd

Global Voices. (2019, August 5). Retrieved from Global Voices: https://globalvoices.org/2019/08/05/netflixs-the-great-hack-highlights-cambridge-analyticas-role-in-trinidad-tobago-elections/

Google. (2019, November 20). *An update on our political ads policy*. Retrieved from Google: https://blog.google/technology/ads/update-our-political-ads-policy/

Harari, Y. N. (2018, September 14). Yuval Noah Harari: the myth of freedom. *The Guardian*. Retrieved from The Guardian: https://www.theguardian.com/books/2018/sep/14/yuval-noah-harari-the-new-threat-to-liberal-democracy

Hern, A. (2018, May 6). Cambridge Analytica: how did it turn clicks into votes? *The Guardian*. Retrieved from https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie

Holt, K. (2019). Facebook Fined Yet Again Over Cambridge Analytica Scandal. *Forbes*. Retrieved from https://www.forbes.com/sites/krisholt/2020/12/30/facebook-fined-yet-again-over-cambridge-analytica-scandal/

IBM. (2020). *Personality Insights*. Retrieved from IBM Watson: https://personality-insights-demo.ng.bluemix.net/

Information Commissioner's Office UK. (2019, October 30). *Statement on an agreement reached between Facebook and the ICO*. Retrieved from ico.: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/10/statement-on-an-agreement-reached-between-facebook-and-the-ico/

Kang, C., & Frenkel, S. (2018, April 4). Facebook Says Cambridge Analytica Harvested Data of Up to 87 Million Users. *The New York Times*. Retrieved from The New York Times: https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html?auth=login-email&login=email

Mayer, J. (2018, November 18). New Evidence Emerges of Steve Bannon and Cambridge Analytica’s Role in Brexit. *The New Yorker*. Retrieved from https://www.newyorker.com/news/news-desk/new-evidence-emerges-of-steve-bannon-and-cambridge-analyticas-role-in-brexit

Meredith, S. (2018, April 10). Facebook-Cambridge Analytica: A timeline of the data hijacking scandal. *CNBC*. Retrieved from CNBC: https://www.cnbc.com/2018/04/10/facebook-cambridge-analytica-a-timeline-of-the-data-hijacking-scandal.html

Microsoft. (2020). *Create a sunburst chart in Office*. Retrieved from Office Support: https://support.microsoft.com/en-us/office/create-a-sunburst-chart-in-office-4a127977-62cd-4c11-b8c7-65b84a358e0c

Morris, W. (2019). GDPR, ePrivacy and cookies: an update . *Lexology*.

Privacy International. (2019, April 30). Retrieved from Privacy International: https://privacyinternational.org/news-analysis/2857/cambridge-analytica-gdpr-1-year-lot-words-and-some-action

Vincent, J. (2020, June 17). *Facebook and Instagram will let users 'turn off' all political ads for the 2020 election*. Retrieved from The Verge: https://www.facebook.com/communitystandards/graphic_violence

Vogels, R. (2016, 16). Trump, Micro Targeting And The Mechanisms Of Data Capitalism. *Huffpost*. Retrieved from Huffpost.

White, S. (2019, July 24). Interview: Behind The Great Hack, with David Carroll. *PrivSec Report*. Retrieved from https://gdpr.report/news/2019/07/24/interview-behind-the-great-hack-with-david-carroll/

List of figures and tables