

Why Does Artificial Intelligence Challenge Democracy?

A Critical Analysis of the Nature of the Challenges Posed by AI-Enabled Manipulation

Caroline Serbanescu*

The Cambridge Analytica scandal has shown how Artificial Intelligence ('AI') applications can be used to influence electoral decisions. The apparent discomfort which such AI applications have in doing so triggered perhaps best illustrates the systemic nature of the disruption posed by AI applications to legal institutions in general, and to democracy in particular.

In an attempt to find the root causes of such systemic legal disruption, this paper investigates why AI applications challenge democracy pursuant to the problem-finding approach. It argues, akin to a hypothesis, that the reason why AI applications disrupt democracy lies in the new forms of manipulation which they enable and which this paper calls "AI-enabled manipulation". This paper thus presents a critical analysis of the nature of the challenges posed by AI-enabled manipulation to democracy by showing that such new forms of manipulation disrupt three of the main principles or assumptions on which democracy relies. These three democratic assumptions are citizens' autonomy, the principle of equal participation and the public forum. Throughout the paper, suggestions for regulatory responses to AI-enabled manipulation are also made. Furthermore, this paper exposes potential new problems and new regulatory concerns which its analysis generates, thereby opening to potential further research. Lastly, this paper suggests a shift of regulatory focus in the face of AI-enabled manipulation.

* LL.M. student in IP and ICT Law at KU Leuven (Belgium)
[caroline.serbanescu@gmail.com]

Introduction

The Cambridge Analytica scandal has shown how Artificial Intelligence ('AI') applications such as fake news, fake accounts and algorithmic profiling can be used to influence electoral decisions. The apparent discomfort which such AI applications have in doing so triggered perhaps best illustrates the systemic nature of the disruption posed by AI applications to legal institutions in general, and to democracy in particular.

In an attempt to find the root causes of such systemic legal disruption, this paper investigates why AI applications challenge one of the foundations of many legal systems: democracy. My thesis, akin to a hypothesis, is that AI applications disrupt democracy due to the new forms of manipulation that they enable, which I call "AI-enabled manipulation". Because AI-enabled manipulation disrupts three of the main presumptions or principles on which democracy relies, these applications pose structural challenges to democracy. The approach I am taking is thus internal to individuals, focussing on AI applications' ability to shape individuals' decision-making processes, and not external to them, as I do not discuss AI applications' potential to make decisions about them.¹

I understand democracy in this paper in its etymological sense, i.e. the rule by the people. More precisely, I focus on liberal democracy, a form of government based on citizen representation and respect for individual freedoms and choices, due to its wide application.² Furthermore, this paper studies why AI applications challenge citizens' representation in parliaments, and thus the legislative branch of government. It does therefore not examine legal disruption resulting from AI applications to the executive and judicial branches of government and to constitutional review. Accordingly and because of my

¹ Daniel Susser, 'Invisible Influence: Artificial Intelligence and the Ethics of Adaptive Choice Architectures' (2019) <<https://dl.acm.org/doi/pdf/10.1145/3306618.3314286>> accessed 18 May 2020, 403.

² Amartya Sen, 'Democracy as a Universal Value' (1999) 10(3) *Journal of Democracy* 3, 4–5; Russell Hardin, *Liberalism, Constitutionalism, and Democracy* (OUP 1999), 6; Anders Westholm, José Ramón Montero, Jan W van Deth, 'Introduction: Citizenship, Involvement, and Democracy in Europe' in Jan W van Deth, José Ramón Montero, Anders Westholm (eds), *Citizenship and Involvement in European Democracies: A Comparative Analysis* (Routledge 2007) 1, 4; William A Galston, 'The Populist Challenge to Liberal Democracy' (2018) 29(2) *Journal of Democracy* 5, 9–10.

etymological understanding of democracy, I decided to focus on citizens' participation in democratic processes by way of voting for their representatives. Despite that there are other ways for citizens to participate in democracy, my view is that studying the challenges posed by AI applications to elections is the most illustrative of the systemic nature of such challenges to democracy.

Furthermore, this paper takes a problem-finding approach, in contrast to the problem-solving approach generally characterising legal thinking. Pursuant to this approach, this paper seeks to find questions – not definitive solutions – arising from an ill-defined problem, AI-enabled disruption of democracy, for which there are “no known methods of solution” and if solutions are found, there are no criteria for assessing their correctness. Therefore, this paper seeks to open up to new questions in relation to AI challenges to democracy in order to help drafting regulatory responses without gaps and without generally threatening regulatory efforts.³

In this paper, I defend my thesis that AI-enabled manipulation is the root cause of or the thread behind AI challenges to representative democracy as it impairs three main principles or assumptions underlying democracy. These assumptions are citizens' autonomy (Section 1); the principle of equal participation in democratic processes (Section 2); and the public forum (Section 3). Section 1 also discusses AI-enabled manipulation as new forms of manipulation whereas Section 4 proceeds with a suggestion for a shift of regulatory focus.

³ Patricia Kennedy Arlin, 'Wisdom: The Art of Problem-Finding' in Robert J Sternberg (ed), *Wisdom: Its nature, origins, and development* (CUP 1990) 230, 231, 235, 239; Hin-Yan Liu, 'From the Autonomy Framework towards Networks and Systems Approaches for 'Autonomous' Weapons Systems' (2019) 10 *Journal of International Humanitarian Legal Studies* 89, 89, 90, 92, 93.

1. Threat to Autonomy as Enabling Condition of Democracy

In this Section, I first discuss the new forms of manipulation enabled by AI and how such manipulation threatens citizens' autonomy (1.1). I then respond to some possible objections to my argumentation (1.2).

1.1 Unfolding the Threat

I develop the following thesis in this part: AI applications threaten democracy as they may be used to manipulate citizens' (i.e. both voters and representatives) decision-making processes, thereby threatening their autonomy as necessary precondition for participating in democratic processes. Manipulation is here understood as "imposing a hidden influence on someone's decision making".⁴ In my view, AI-enabled manipulation operates as an influence external to one's cognitive processes whose hiddenness is such that the influence is unconsciously internalised in individuals' decision-making processes.

More precisely, I believe that AI applications have the potential to manipulate citizens as they are able to shape in a personalised, dynamic and concealed manner the "choice architecture" of citizens, i.e. both the set of available choices and the way they are formulated. The choice architecture thus represents the context or environment in which individuals make decisions.⁵

AI applications are able to shape our choice architectures *in a hidden manner* since these technologies have become transparent to us literally, i.e. we experience the world *through* them. Such transparency has been made possible by our increasing use of technologies in our daily activities, with the result that technologies increasingly and pervasively mediate our experiences and perceptions of the world. In other words, because we increasingly use AI applications daily, we focus on the activities facilitated by these technologies

⁴ Susser (n 1) 405; Daniel Susser, Beate Roessler, Helen Nissenbaum, 'Online Manipulation: Hidden Influences in a Digital World' (2019) 4 Georgetown Law Technology Review 1, 26.

⁵ Susser (n 1) 404; Susser, Roessler, Nissenbaum (n 4) 39.

instead of focussing on the technologies themselves which influence the context in which we make decisions.⁶

The ubiquity of technology mediation gave rise to Big Data, consisting in (the collection of) vast amounts of data, including personal data⁷, and in the analysis thereof at high speed so as to make valuable inferences. Accordingly, Big Data subsequently led to the development of AI-enabled manipulation as the latter relies on the former for its operation. Indeed, the more that is known about each individual, the easier it is to shape her choice architecture so as to steer her decision making in the desired decision.⁸

Manipulation is in itself not a new phenomenon. It involves the exploitation of the manipulee's cognitive or affective weaknesses and vulnerabilities in order to steer her decisions towards the manipulator's ends. It does so without the manipulee's conscious awareness or in a way thwarting her capacity to become consciously aware thereof by undermining usually reliable assumptions. The exploited vulnerabilities can result from individual contingencies (e.g. habits, personality, personal history, etc.). Valuable vulnerabilities for manipulation purposes can also arise from trends emerging from the demographic groups to which each individual belongs, thereby potentially disclosing weaknesses that

⁶ Yoni Van Den Eede, 'In Between Us: On the Transparency and Opacity of Technological Mediation' (2011) 16 *Foundations of Science* 139, 141, 144; Susser (n 1) 404–05; Susser, Roessler, Nissenbaum (n 4) 33–34, 38.

⁷ "Personal data" is understood in this paper within the meaning of the General Data Protection Regulation: "any information relating to an identified or identifiable natural person ("data subject")" (Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1, article 4 (1)).

⁸ Zeynep Tufekci, 'Engineering the Public: Big Data, Surveillance and Computational Politics' (2014) <https://pdfs.semanticscholar.org/04e2/f184505a6b67c611bc57c05864385c024418.pdf?_ga=2.34130535.792266196.1589824767-2048156165.1589824767> accessed 18 May 2020, 20; Karen Yeung, 'Algorithmic Regulation: A Critical Interrogation' (2018) 12 *Regulation & Governance* 505, 514.

individuals themselves do not see.⁹ Accordingly, identifying the group vulnerabilities of individuals allows to refine manipulation.

In my view, the unprecedented threat of AI applications to democracy stems from the new forms of manipulation which they enable.

There is a twofold qualitative change brought about by AI-enabled manipulation in contrast to “analogue” manipulation (i.e. not involving AI). On the one hand, AI-enabled manipulation allows for *tailored* influences over decision-making processes due to technology’s pervasiveness in – or even surveillance of – daily lives. Such pervasiveness thus allows to more easily identify and subsequently exploit individuals’ weaknesses arising from both individual and group contingencies. On the other hand, AI-enabled manipulation is *dynamic* or *adaptive* in the sense that it can adapt and refine itself quickly in the light of individuals’ conducts notably on the internet, which can reveal changes of preferences.¹⁰

Furthermore, AI-enabled manipulation involves a quantitative change due to the unparalleled reach of these new forms of manipulation. Combined with the possibility to personalise and dynamically change each individual’s choice architecture, the wide reach of AI-enabled manipulation can enable its exploiter to more effectively reach her goal. Indeed, if a large number of voters manipulated by AI applications votes in the direction reflecting the manipulator’s interests, such interests would likely be achieved as they would represent “the will of the people”.¹¹

Overall, the rise of AI applications has thus led to more effective forms of manipulation.

—

In my view, AI-enabled manipulation exacerbates the known threat of manipulation to citizens’ autonomy, which thereby also constitutes a threat to democracy (*infra* 5). Autonomy is here understood as the capacity to “rule oneself” or to act independently on the basis of one’s own reasons that one

⁹ Karen Yeung, ‘Hypernudge: Big Data as a Mode of Regulation by Design’ (2017) 20 Information, Communication & Society 118, 122; Susser, Roessler, Nissenbaum (n 4) 3, 26, 32.

¹⁰ Susser (n 1) 404; Susser, Roessler, Nissenbaum (n 4) 3.

¹¹ Susser, Roessler, Nissenbaum (n 4) 4, 29.

recognizes and endorses.¹² In fact, I believe that the presence of manipulation always implies a threat to one's autonomy as imposing a hidden influence on one's decision making leaves uncertain as to this person's (remaining) autonomy. Consequently, just as AI-enabled manipulation is the thread of my argumentation, so is the resulting threat to autonomy.

AI-enabled manipulation exacerbates the undermining of citizens' autonomy due to its quantitative and qualitative improvements (*supra*) as well as its hiddenness. More precisely, such manipulation challenges the following two features of autonomy generally recognized by autonomy theorists.

On the one hand, autonomous persons have the “cognitive, psychological, social, and emotional *competencies* [or capacities] to deliberate, to form intentions and to act on the basis of that process”.¹³

On the other hand, “autonomous persons can (at least in principle) critically reflect on their values, desires, and goals, and act for their own reasons, i.e. endorse them *authentically* as their own”. The latter condition amounts in fact to individuals' capacity for self-authorship over their actions.¹⁴

AI-enabled manipulation threatens both autonomy features. On the one hand, such manipulation impairs autonomous individuals' capacity to *competently* deliberate as it designs the features of individuals' choice architecture in a concealed, tailored and adaptive manner so as to influence them without their conscious awareness. On the other, such influence steers citizens to act, such as to vote for a candidate, for reasons they cannot understand, as they are not their own, and therefore cannot *authentically* endorse as their own.¹⁵

As a counter argument to the above, it may be contended that citizens do not always make decisions on the basis of reasons. Indeed, when citizens do not know what to decide, even for important decisions such as voting decisions, they may

¹² Andrew Sneddon, *Autonomy* (Bloomsbury UK 2013) 3; Susser (n 1) 406–07.

¹³ John Philip Christman, *The Politics of Persons: Individual Autonomy and Socio-Historical Selves* (CUP 2009) 154–55; Yeung, ‘Hypernudge ...’ (n 9) 124; Susser, Roessler, Nissenbaum (n 4) 36.

¹⁴ Sneddon (n 12) 7; Yeung, ‘Hypernudge ...’ (n 9) 124; Susser, Roessler, Nissenbaum (n 4) 18, 36.

¹⁵ Cass R Sunstein, *The Ethics of Influence: Government in the Age of Behavioral Science* (CUP 2016) 83; Susser (n 1) 406–07; Susser, Roessler, Nissenbaum (n 4) 38.

“pick” a decision simply because they “felt like it”.¹⁶ However, in that case, I believe that AI-enabled manipulation may still represent a threat to citizens’ autonomy and hence to democracy. Indeed, AI-enabled manipulation may perceive this lack of rational deliberation as a cognitive vulnerability which it can exploit in order to drive manipulees to cast the desired vote. This is because AI-enabled manipulation has the unparalleled potential to inculcate the manipulees reasons – the most appealing to them as made possible by targeted manipulation – to make the decision preferred by the manipulator.

By undermining citizens’ autonomy, I believe that AI-enabled manipulation endangers the foundations of liberal representative democracies.

The core idea of liberal representative democracies is that “political power derives its authority from the autonomous consent of the governed”.¹⁷ Accordingly, democracy presupposes that the governed are politically autonomous, which is the case when they are sufficiently able to participate in democratic processes in a way that reflects their capacity to self-rule.¹⁸

However, with AI-enabled manipulation, the manipulated citizens may have given a consent to political power which does not reflect their capacity to self-rule as they may not understand and hence endorse the reasons backing their consent. Therefore, they may not have given a consent reflecting their own will but rather that of the manipulators. In that case, votes do not represent the will of the people, which conflicts with the core of representative democracy as it requires Parliaments to express the will of the people and the public interest. Ultimately citizens’ acknowledgment of political institutions, especially the Parliament, as their own may be undermined.¹⁹ Moreover, the (input) legitimacy of the representatives, which is “based on the assumption that political choices are legitimate if and because they reflect the will of the people”, may be

¹⁶ Sunstein (n 15) 66.

¹⁷ Sneddon (n 12) 4.

¹⁸ Westholm, Montero, van Deth (n 2) 6–7; Sneddon (n 12) 3–4, 7.

¹⁹ Lawrence Pratchett, “The Core Principles of European Democracy” in *Reflections on the Future of Democracy in Europe* (Council of Europe Publishing 2005) 31, 33; Susser (n 1) 406.

impaired.²⁰ The same legitimacy problem occurs if representatives are manipulated when making decisions as in that case, their decisions do also not reflect the will of the people but that of the manipulator (*infra*).

In practice, AI-enabled manipulation means that AI applications can, for instance, influence in a tailored, dynamic and hidden way the electoral information each individual receives on social media or when making internet searches. In doing so, AI applications shape each individual's choice architecture in a unique way in the sense that no voter receives the same electoral information (*infra*).²¹ As a result, citizens may be led to vote for a candidate for whom they would not necessarily have voted in the absence of influence or to abstain from voting while they would have perhaps voted failing the interference (*infra*). Whether individuals decide to vote or abstain from voting while being influenced, the chosen decision may not always be the most beneficial to each citizens' interests.

Accordingly, besides what was stated above, AI-enabled manipulation further undermines individuals' autonomy as autonomous citizens usually act in their own interest for the sake of enhancing their own welfare. If citizens are not able to make voting choices reflecting their own welfare, their interests and will cannot be served by their elected representatives, as they are not made aware thereof. As a result, democracy is further undermined as elected representatives cannot express the will of the "people" in their legislative activities.²²

The same threat to representative democracy arises in relation to elected representatives' manipulation. AI applications may indeed shape representatives' decision-making processes so as to induce them to make legislative choices

²⁰ Magdalena Godowska, 'Democratic Dilemmas and the Regulation of Lobbying - the European Transparency Initiative and the Register for Lobbyists' (2011) 14 Yearbook of Polish European Studies 181, 183.

²¹ Jonathan Zittrain, 'Engineering an election' (2014) 127 Harvard Law Review 335, 336, 340; Tufekci (n 8) 26; Vyacheslav Polonski, 'How Artificial Intelligence Silently Took Over Democracy' (2017) World Economic Forum <www.weforum.org/agenda/2017/08/artificial-intelligence-can-save-democracy-unless-it-destroys-it-first/> accessed 19 May 2020; Karl M Manheim, Lyric Kaplan, 'Artificial Intelligence: Risks to Privacy and Democracy' (2019) 21 Yale Journal of Law and Technology 106, 147–50.

²² Susser (n 1) 406.

beneficial to the AI exploiter's interests. Therefore, the ultimate decision made by representatives may again reflect the manipulator's interests, and not citizens' interests.

The new forms of manipulation to which AI applications give rise thus threaten the way citizens participate in decision-making processes and the way public authority is exercised.²³

An objection to this argument could be that manipulation rarely *totally* deprives its target of autonomy because a manipulator never fully controls her. Indeed, in contrast to coercion, which is another way of influencing decision-making processes, manipulation does not imply the entire displacement of the target as the decision maker by way of compulsion. Manipulation involves in fact a more subtle insinuation into the target's decision-making processes as it impairs her capacity for conscious decision making. In other words, the manipulee makes a decision while not *fully* understanding why she took this decision or whether it served her own or someone else's interests.²⁴ In my view, such argument implies that if manipulated citizens cast a vote, this vote can still be deemed to reflect citizens' own will as it is "saved" by citizens' remaining autonomy. This remaining autonomy could indeed have steered the manipulees to cast a vote different than that preferred by the manipulator or could have deemed the manipulator's preferred decision to be in these manipulees' best interest, without being aware of the manipulation. Accordingly, the threats to democracy would disappear.

I submit the following counter arguments which qualify such objection.

On the one hand, as already stated (*supra* 5), it is common for people to lack a full understanding of the reasons for their choices, even without being manipulated, as many of them are based on unconscious processing (so-called "System 1"). Indeed, System 1 of processing information is intuitive and prone

²³ Snežana Samardžić-Marković, 'AI and Democracy' (High-level Conference on 'Governing the Game Changer – Impacts of Artificial Intelligence Development on Human Rights, Democracy and the Rule of Law', Helsinki, 26–27 February 2019) <<https://rm.coe.int/-artificial-intelligence-and-democracy-introductory-speech-by-snezana-/1680933353>> accessed 19 May 2020.

²⁴ Susser, Roessler, Nissenbaum (n 4) 17.

to bias. I have already argued before that such lack of rational deliberation can be exploited by manipulators. Nevertheless, it may be that manipulators do not perceive such lack of rational deliberation and that citizens actually act or vote intuitively. This would imply that regulatory responses to AI-enabled manipulation should distinguish between situations where our actions are determined by our own impulses and situations where they are determined by factors intentionally framed by others to influence our decision making.²⁵ This also begs the question, deserving further research, as to what legal value should be awarded to intuitive votes given that they are hardly identifiable but that, from a democratic viewpoint, they do not seem desirable. From the latter viewpoint, intuitive votes do indeed not ensure that voters truly express their will, which should become apparent after a rational deliberation involving a consideration of all candidates' positions.

On the other hand, I believe that distinguishing between full and partial undermining of autonomy leads to the following difficulties and hence does not rule out concerns about democracy.

Firstly, it is difficult, if not impossible, to quantify the extent to which manipulation impairs individuals' autonomy. Therefore, it seems more appropriate, from a regulatory perspective, to shift the focus from the extent of autonomy impairment to the presence of AI-enabled manipulation, which is thereby deemed reprehensible as such.²⁶ Accordingly, any interference with one's decision-making processes is problematic, regardless of the extent of autonomy impairment, precisely because it creates uncertainties as to the extent of individuals' remaining autonomy.

Secondly, I believe that the view that the manipulee does not fully understand the reasons for her decisions or whether it served her interests presupposes that the manipulee suspects afterwards to have been subject to some kind of undue influence. I believe that AI-enabled manipulation may be so effective that it is difficult, if not impossible, for manipulees to even suspect such influence afterwards *by themselves*. In other words, AI-enabled manipulation may give the

²⁵ Sunstein (n 15) 89-90; Mark Egan, *An Analysis of Richard H. Thaler and Cass R. Sunstein's Nudge: Improving Decisions about Health, Wealth and Happiness* (Macat Library 2017) 35.

²⁶ In that sense, see Susser, Roessler, Nissenbaum (n 4) 41.

impression to the target that she votes autonomously. It does so by subtly influencing the target's decision making in a such a personalised and dynamic way that the target may unknowingly believe the manipulator's interests to be her interests and that she fully understands the reasons for her votes.²⁷

In that sense, I believe that manipulees cannot become aware of AI-enabled manipulation by themselves but through the help of external actors. Thus far, only the manipulator (or individuals from the group of persons behind the manipulation) or persons having been in touch with the manipulator (without participating in the manipulation) can disclose AI-enabled manipulation.²⁸ However, the questions remain on how legislation could be drafted so as to incentivise such persons to disclose AI-enabled manipulative practices and on what other steps or actors States could take or involve to facilitate such disclosure, especially in the electoral context. Such questions fall out of the scope of this paper but warrant further research which I strongly encourage.

The Cambridge Analytica scandal is a good illustration of my thought. Until the disclosure of the hidden AI-enabled manipulation to which millions of Facebook users have been subject, I do not think that many users questioned their voting decisions. Instead, several actors, including researchers who had been contacted by Cambridge Analytica but did not participate therein, have disclosed the practices of this firm to newspapers. According to one of these researchers, it was strong ethical opposition that led him to disclose the AI-enabled manipulation.²⁹ Nevertheless, in line with my above suggestion, further

²⁷ In that sense, see Manheim, Kaplan (n 21) 109, 150.

²⁸ I however do not exclude the possibility that in future more and more journalists will be able to reveal AI-enabled manipulation, notably through the command of AI applications (although such tools currently have inherent limitations (*infra*)).

²⁹ See statement of one of the researchers having contributed to the disclosure of the scandal: Michal Kosinski, 'Statement on Cambridge Analytica' (2018) <<https://drive.google.com/file/d/1zRaTAx0mpRC0m7-3wQRaDPYTOGMdvNBt/edit>> accessed 19 May 2020. See one of the first press articles disclosing the scandal: Harry Davies, 'Ted Cruz using firm that harvested data on millions of unwitting Facebook users' *The Guardian* (11 December 2015) <<https://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data>> accessed 19 May 2020.

research could focus on whether other, legally enforceable and less subjective, incentives to disclose AI-enabled manipulation could be sustained by law.

In the light of the above, I conclude that by undermining citizens' autonomy, AI-enabled manipulation has the potential to transform citizens into passive actors (or "puppets").³⁰ Consequently, citizens are no longer democratic agents as undermining their autonomy implies that they can no longer participate in democratic processes so as to express their will. As a result, democracy as the rule by the people is undermined. Another related consequence in my view is the erosion of the notion of the law, as the incarnation of the will of the people, since the latter's existence is undermined.³¹ This begs the question, deserving further research, as to whether there are other reasons why AI applications disrupt our understanding of the law.

1.2 Possible Objections

Besides the previous objections stated above, I here respond to further potential and more general objections to my above argumentation.

Firstly, some may contend that it is difficult, if not impossible, to assess if citizens are being manipulated. I agree with this objection and I have already explained above why I believe that regulatory efforts should focus on the presence of AI-enabled manipulation as such. A further suggestion building thereupon is for regulatory efforts to take a general instead of an individual approach in the face of AI-enabled manipulation. Hence, instead of scrutinizing individual cases, focussing on and examining the manipulation strategy at a general level could be more effective and less time-consuming as it could allow to reveal individual instances of manipulation.³² This would notably require from regulators to qualify a given attempt of influencing citizens as AI-enabled manipulation. Such qualification exercise however requires the existence of a

³⁰ Susser, Roessler, Nissenbaum (n 4) 17; Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (PublicAffairs 2019) 21 ff.

³¹ In that sense, see Anthony J Casey, Anthony Niblett, 'The Death of Rules and Standards' (2017) 92 *Indiana Law Journal* 1401.

³² Sunstein (n 15) 86; Susser, Roessler, Nissenbaum (n 4) 41.

legal definition of AI-enabled manipulation which could include its constitutive elements (e.g. dynamic, tailored, hidden, large scale, etc.).

Secondly, one may question whether AI-enabled manipulation really threatens democracy when a manipulee would have casted the same vote without manipulation. In my opinion, AI-enabled manipulation remains problematic for democracy in that case. Indeed, I believe that it is only a coincidence if the manipulee's autonomous vote equates the vote she casts when manipulated. Hence, it is not the outcome of the manipulation that counts but rather the interference into one's decision making process as it leaves uncertain as to the manipulee's autonomy (*supra* 7).³³ Since AI-enabled manipulation can change the set of available electoral options and their understanding, it may shape the reasons why manipulees vote for a certain candidate, hence affecting the authenticity component of autonomy. In other words, the reasons why manipulees cast the same vote without and under manipulation may differ and the manipulees may thus not authentically endorse the reasons for their manipulated vote. Therefore, concerns for democracy do not disappear.

Thirdly, one may argue that as choices in general, including votes, are always influenced and conditioned by social, cultural, economic and political contexts, it is difficult to distinguish these contextual influences from the manipulator's influences.³⁴ Nevertheless, such contextual influences can be exploited by AI-enabled manipulation as group vulnerabilities of each citizen, thereby allowing to better tailor the manipulation (*supra* 4). As a result, there is perhaps no need to distinguish between these contextual influences and AI-enabled manipulation as the latter may incorporate the former in its process of shaping individuals' decisions.

Fourthly, another objection could be that it is difficult to distinguish the effect, on individuals' decision making, of AI-enabled manipulation from that of analogue manipulation. I agree with that objection but perhaps in certain cases, the exploiters of AI applications are the same persons who seek to manipulate voters via analogue means. By combining different forms of manipulation, manipulation is indeed rendered more effective. Therefore, in

³³ Susser, Roessler, Nissenbaum (n 4) 42.

³⁴ Yeung, 'Hypernudge ...' (n 9) 129; Susser, Roessler, Nissenbaum (n 4) 42–43.

these cases, there is perhaps no need to distinguish between analogue and AI-enabled manipulation. Furthermore, such difficulty could perhaps be overcome if the focus is again placed on the presence of analogue and AI-enabled manipulation, which would allow to distinguish them. If that is the case, regulatory responses should in my view take account of the fact that AI-enabled manipulation is more effective in achieving the desired outcome than analogue manipulation notably due to its ability to constantly personalise manipulation (*supra* 4). Therefore, manipulators using AI applications to achieve their ends should bear a larger or special legal responsibility for the more effective threat to democracy they pose than manipulators using analogue means.³⁵

2. Threat to Equal Participation

Apparent from the previous Section is that AI-enabled manipulation disrupts liberal representative democracy as it has the potential to eliminate any form of democratic participation by citizens. Such elimination would in fact be concealed as citizens would still physically vote without having the cognitive capacities or autonomy to do so. Another – less radical – challenge of AI-enabled manipulation to democracy is the potential to disrupt the democratic principle of equal participation. According to this principle, each citizen has the same ability to express her will to representatives.³⁶

In my opinion, both the elimination of participation and unequal participation as a result of AI-enabled manipulation can occur at the same time. As the elimination of participation cannot currently reach all members of an electorate (as it notably depends on *all members* accessing the internet, which is not the case of e.g. some elderly), if some members are not cognitively able to

³⁵ In that sense, see Susser (n 1) 407.

³⁶ Sydney Verba, 'Political Equality. What is It? Why Do We Want It?' (Review Paper for Russell Sage Foundation 2001) <<https://www.russellsage.org/sites/all/files/u4/Verba.pdf>> accessed 19 May 2020, 2; Westholm, Montero, van Deth (n 2) 3; Jan Teorell, Paul Sum, Mette Tobiasen, 'Participation and Political Equality: an Assessment of Large-Scale Democracy' in Jan W van Deth, José Ramón Montero, Anders Westholm (eds), *Citizenship and Involvement in European Democracies: A Comparative Analysis* (Routledge 2007) 384, 385.

participate anymore in democratic processes, it is likely that equality in participation is impaired. Indeed, other members not reached by AI-enabled manipulation retain their cognitive capacities to participate in such processes.

There are in my view at least two ways in which AI-enabled manipulation may threaten equality in participation.

On the one hand, as mentioned, AI-enabled manipulation relies on large collection of personal data found notably but not exclusively on the internet. However, not all individuals leave the same amount of personal data on the internet as such amount notably depends on membership to (several) social networks and the number of internet searches initiated by each individual. Therefore, I believe that there are different degrees of manipulation to which AI applications may lead depending on the data that the algorithm finds on each individual. As a result of these different degrees of manipulation, there are different extents to which the manipulee's autonomy are impaired so that some individuals may be more vulnerable to AI-enabled manipulation than others. This means that AI-enabled manipulation's effect on some individuals' decision making may not be such as to hinder their capacity to express a will or vote that they endorse (*supra*). Accordingly, equality in participation is undermined due to these different degrees of autonomy impairments. In that relation, such different degrees of AI-enabled manipulation may also reflect discrimination against protected classes, thereby further leading to unequal participation. Since there is more data produced on such classes than on non-protected classes, the former are more vulnerable to AI-enabled manipulation, hence less likely to express autonomous choices.³⁷ However, I am aware that it may be impossible to quantify the interference of AI-enabled manipulation on citizens' autonomy and there will therefore always be an uncertainty in that regard. To overcome

³⁷ For instance, in the US, many communities of colour are more thoroughly surveilled than white communities (Alvaro M Bedoya, 'The Color of Surveillance: What an infamous abuse of power teaches us about the modern spy era' *Slate* (2016) <<https://slate.com/technology/2016/01/what-the-fbis-surveillance-of-martin-luther-king-says-about-modern-spying.html>> accessed 19 May 2020; Susser, Roessler, Nissenbaum (n 4) 40–41). Tufekci (n 8) 3; Anja Bechmann, 'Data as Humans: Representation, Accountability, and Equality in Big Data' in Rikke Frank Jørgensen (ed), *Human Rights in the Age of Platforms* (MIT Press 2019) 73, 77–78.

such difficulty, I reiterate that regulatory efforts should focus on the presence of manipulation, which should be deemed reprehensive as such.

On the other hand, in countries where voting is not compulsory, AI-enabled manipulation may be used to deter voters from certain (minority or vulnerable) groups to vote, i.e. suppress their votes. Such practice creates an inequality in the possibility of diverse voter groups to express their will in the form of votes as groups subject to AI-enabled manipulation are deprived from such possibility.³⁸ Accordingly, equality in participation is disrupted.

Attempts from a foreign government to suppress votes of certain (protected) groups using AI applications have actually already occurred in the 2016 US presidential elections.³⁹

A problem for democracy arising from AI applications disrupting equality in participation is the undermining of the legitimacy of parliamentarians. Indeed, if not all voters had an equal chance to vote for their representatives, the latter cannot be said to represent and act in accordance with the “will of the people”. A further problem is that such disruption is likely to lead to the “tyranny of the majority”, whereby the majority, possibly representing the interests of the manipulator (*supra* 6), uses its political power to serve its own interests at the expense of the rights of others and of the public good.⁴⁰

³⁸ Elaine Kamarck, ‘Malevolent Soft Power, AI, and the Threat to Democracy’ (Brookings Report 2018) <www.brookings.edu/research/malevolent-soft-power-ai-and-the-threat-to-democracy/> accessed 19 May 2020.

³⁹ The Russian government has been found to have interfered in the 2016 US Presidential election. More precisely, Russia, supportive of republican candidates, used AI applications (e.g. fake accounts, fake news, etc.) to suppress Afro-Americans’ votes, which have historically leaned towards democratic candidates (Alec Tyson, Shiva Maniam, ‘Behind Trump’s victory: Divisions by race, gender, education’ (Pew Research Center 2016) <www.pewresearch.org/fact-tank/2016/11/09/behind-trumps-victory-divisions-by-race-gender-education/> accessed 19 May 2020; Kamarck (n 38); S Mueller, ‘Report on the Investigation into Russian Interference in the 2016 Presidential Election’ (US Department of Justice 2019) <www.justice.gov/storage/report.pdf> accessed 19 May 2020, 14, 25).

⁴⁰ *Osmanoğlu and Kocabaş v Switzerland* App no 29086/12 (ECtHR, 10 January 2017) para 84; Ronald J Terchek, Thomas C Conte, *Theories of Democracy: a Reader* (Ringgold Inc 2002) 5; Verba (n 36) 3-4; David Held, *Models of Democracy* (Polity Press 2006) 72; Van Den Eede (n 6) 153.

3. Threat to Public Forum

Today, much political and civic speech takes place online, where AI applications such as fake news and Deepfakes flourish.⁴¹ These applications constitute disinformation strategies as well as concrete examples of AI-enabled manipulation and are in my view the most illustrative of the nature of the challenges of AI applications to the public forum. These disinformation strategies consist in disseminating deceptive information and/or showing recipients the information that they are mostly likely to endorse.

Such disinformation strategies have always existed in the electoral context but AI allows them to more effectively deceive their recipients. This is especially the case of Deepfakes, which consist in altering audio and video messages so as to deceive their recipients' senses. In doing so, the latter are induced to believe that the information conveyed is accurate. Such AI applications can thus be used in order to discredit electoral candidates, so as to incite voters to cast their votes in favour of the manipulator's preferred candidate, or to discredit elected representatives. Other AI applications contributing to disinformation are the algorithms operating many online platforms and search engines which can determine the visibility of political content. Accordingly, groups which cannot afford to rely on such algorithms will be more and more hidden from public view, or there will be changes in their reach that are beyond their control.⁴²

These disinformation strategies are problematic for democracy mainly because they disrupt the public space for deliberation. The unparalleled disruption of AI applications to the public forum stems in my view from the targeted disinformation deployed on each individual, whereby all citizens receive tailored and adaptive, hence necessarily different, electoral information. As a result, such strategies alter the set of information available in the public forum and required for citizens to vote autonomously, i.e. to cast enlightened votes, and to participate in public debates. Indeed, such disinformation strategies make it difficult, if not impossible, for citizens to distinguish official and accurate

⁴¹ Tufekci (n 8) 25.

⁴² Tufekci (n 8) 26; Kamarck (n 38); Manheim, Kaplan (n 21) 137, 142, 146-48; Birgit Schippers, 'Artificial Intelligence and Democratic Politics' (2020) 11(1) *Political Insight* 32-33.

information from fake news or to grasp *all* information which are part of the public forum. This implies that such disinformation strategies do not comply with democracy's requirement that, during election times, the different sides have an adequate opportunity to present their respective cases, and the electorates have the freedom to obtain news and to consider the views of *all* competing parties. Moreover, such practices may cause social polarisation within the public forum, with the impact on democracy being the formation of distinct groups that can no longer understand each other and hence making the reaching of political compromises almost impossible.⁴³

Further AI applications altering the democratic public debate and amounting to AI-enabled manipulation are fake accounts or bots. These applications indeed instigate the belief that citizens are engaging in constructive dialogues with fellow citizens whereas in reality they do not discuss with "real" people and the genuine purpose of such dialogues is to manipulate citizens. Such dialogues may be used in order to change citizens' opinion on electoral candidates with the ultimate purpose being to steer citizens to vote for the candidate preferred by the bots' exploiters. Because fake accounts alter the public forum in the same manner as disinformation strategies, they are collectively referred to in this paper as "AI-enabled alteration" strategies.⁴⁴

In the face of new disruptive technologies, regulatory responses could be to apply by analogy laws regulating allegedly similar phenomena. Hence, the question arises whether AI-enabled alteration strategies can be equalized with the (not necessarily condemnable) practice of propaganda. In my view, they cannot.

Propaganda refers to "an organised effort to spread a particular doctrine or belief" on a large scale.⁴⁵ Both propaganda and AI-enabled alteration can be deployed not only by political parties to promote their candidates but also by

⁴³ Sen (n 2) 9-10; Tufekci (n 8) 26; Sunstein (n 15) 45, 65; Dirk Helbing and others, 'Will Democracy Survive Big Data and Artificial Intelligence?: Essays on the Dark and Light Sides of the Digital Revolution' (2019) <www.researchgate.net/publication/327271384_Will_Democracy_Survive_Big_Data_and_Artificial_Intelligence_Essays_on_the_Dark_and_Light_Sides_of_the_Digital_Revolution> accessed 19 May 2020, 5; Manheim, Kaplan (n 21) 150; Schippers (n 42) 33.

⁴⁴ Polonski (n 21); Kamarck (n 38); Manheim, Kaplan (n 21) 151.

⁴⁵ Edward L Bernays, *Propaganda* (Horace Liveright 1928) 20.

other actors willing to interfere with elections such as foreign states (*supra* 11).⁴⁶ AI-enabled alteration raises however more profound concerns for democracy than propaganda, which would justify a distinct regulatory effort. A first set of concerns relates to the fact that contrary to propaganda, AI-enabled alteration is personalised and adaptive to each recipient. Another set of concerns regards the fact that, in contrast to propaganda, AI-enabled alteration and more generally AI-enabled manipulation rely on widespread surveillance of citizens for their operation. As a result of such widespread surveillance, firms developing AI-enabled manipulation possess unparalleled power of behavioural modification.⁴⁷

In my view, such unparalleled power which these firms possess *risk* in the long term leading to the defeat of democracy over “technocracy”. Indeed, these firms’ expertise in surveillance technologies could allow them to manipulate citizens and their representatives to such an extent that they or their clients would *de facto* be running States.⁴⁸ Only a perception of democracy would remain as citizens would continue to vote whereas their voting decisions would not be autonomous and fully informed as a result of AI-enabled manipulation and alteration.

My reasoning however begs questions as to whether such a society would be achievable and sustainable notably given the conflicting interests of the various firms concerned and/or of their clients.

Furthermore, for the sake of clarity, my view is not that the displacement of democracy by technocracy *will* occur in future. My point seeks rather to open

⁴⁶ Karl E Ettinger, ‘Foreign Propaganda in America’ (1946) 10(3) *The Public Opinion Quarterly* 329, 329; Kamarck (n 38).

⁴⁷ Yeung, ‘Hypernudge ...’ (n 9) 130. See also Zuboff, *The Age of Surveillance Capitalism ...* (n 30); Shoshana Zuboff, ‘“We Make Them Dance”: Surveillance Capitalism, the Rise of Instrumentarian Power, and the Threat to Human Rights’ in Rikke Frank Jørgensen (ed), *Human Rights in the Age of Platforms* (MIT Press 2019) 3.

⁴⁸ Marc Hudson, ‘Ending Technocracy with a Neologism? Avivocracy as a Conceptual Tool’ (2018) 55 *Technology in Society* 136, 136-37; Yeung, ‘Algorithmic Regulation ...’ (n 8) 518; Julie E Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism* (OUP 2019) 3. In that sense, see J Benjamin Hurlbut, ‘Laws of Containment: Control Without Limits in the New Biology’ in Irus Braverman (ed), *Gene Editing, Law, and the Environment: Life Beyond the Human* (Taylor & Francis 2017) 77, 86–91.

up to a potential challenge to democracy to be taken into account by regulatory responses to AI-enabled manipulation and alteration.

4. A Necessary Shift of Focus?

In the light of the above, perhaps the challenges posed by AI applications to representative democracy revolve around the use made of such technology. Indeed, it may be because humans use AI applications for illegitimate ends from a democratic perspective, i.e. manipulation, that democracy and its core principles are disrupted. This would imply that the regulatory target should shift from AI as a technology to human use of such technology.

An illustration of my argument is the fact that the threats posed by AI to democracy can possibly be solved *in future* by using AI applications themselves since several of them are already being used to detect and remove undesirable content online. However, as mentioned, such use of AI applications to remedy their threats to democracy will possibly only be achievable in future due to the current limitations of these technologies, such as intrinsic biases or the risks of unduly blocking desirable content.⁴⁹

A likely objection to my proposal of targeting human use of AI could be that it is not because of human use of technologies that AI poses challenges to legal institutions in general, and democracy in particular, but because AI develops in an unforeseeable manner. Nevertheless, in my view, the focus should still be placed on human use of AI since humans may have intended the unforeseeable developments of AI. Indeed, it is likely that because AI applications have the capacity to generate unique solutions not considered by humans and potentially improving humans' lives, these technologies' exploiters have an incentive to develop AI systems which can generate unexpected solutions.⁵⁰

⁴⁹ See Katarina Kertysova, 'Artificial Intelligence and Disinformation: How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Countered' (2018) 29 *Security and Human Rights* 55, 59–61.

⁵⁰ Matthew U Scherer, 'Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies' (2016) 2 *Harvard Journal of Law & Technology* 353, 364–66.

Regulatory efforts could accordingly place responsibility on AI applications' exploiters. But even in that case, such exploiters may not be concerned about their responsibility if the costs for breaching it would not override the potential profits thereby gained. This could be the case given notably the potential large profits made from surveillance of citizens.⁵¹

The core regulatory question that then arises is how to ensure that individuals use technologies for the common good, including enhancing democracy? As will be shown, current proposals to tackle this question seem limited and therefore this topic warrants further research.

Some proposals made in the electoral context relate to placing responsibility on the campaigns themselves, which should thus monitor AI-enabled manipulation strategies. However, campaigns may for instance have the incentive to block information which would tarnish the image of the campaigns' candidates whereas such information would actually be accurate. Therefore, I believe that perhaps a more neutral actor could be involved in such monitoring, and further research could focus on this topic. Avoiding to unduly block information should also be a central concern for regulatory responses to AI-enabled manipulation in order to avoid censorship which could lead to and perpetuate authoritarian regimes.⁵²

Further proposals relate to educating citizens about AI-enabled manipulation and alteration by teaching them how to distinguish real from fake news, real from bot accounts and how to identify manipulation.⁵³ Nevertheless, perhaps educating citizens will not suffice to overcome AI-enabled manipulation's threats

⁵¹ Zuboff, 'We Make Them Dance ...' (n 47) 15.

⁵² Kamarck (n 38). See also discussions regarding the EU Directive on Copyright in the Digital Single Market (Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L130/92): Martin Senftleben and others, 'The Recommendation on Measures to Safeguard Fundamental Rights and the Open Internet in the Framework of the EU Copyright Reform' (2018) 40 *European Intellectual Property Review* 149; Sophie Stalla-Bourdillon and others, 'A Brief Exegesis of the Proposed Copyright Directive' (2017) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2875296> accessed 19 May 2020.

⁵³ Kamarck (n 38).

to democracy as such manipulation affects individuals' choice architecture. Accordingly, AI-enabled manipulation may affect decision making at such an unconscious level of human thinking that one may still be manipulated even if one is aware of the manipulation to which one is subject.

Conclusion

This paper sought to critically analyse my thesis according to which the reason why AI applications disrupt democracy lies in the new forms of manipulation enabled by AI applications.

I showed that AI-enabled manipulation is especially disruptive due to its ability to tailor manipulation to each individual, to adapt to each individual's change of conduct or thinking, its transparency (or improved hiddenness) and its potential wide reach.

I then analysed my thesis that AI-enabled manipulation challenges democracy because it disrupts three of its main principles or assumptions: citizens' autonomy, equal participation in democratic processes such as elections and that the public forum disseminates all the information required to cast enlightened votes. I conducted this analysis in the light of potential objections which allowed to refine my thesis.

I moreover made throughout my argumentation several proposals for further research, notably on how to ensure that individuals use technologies to enhance democracy.

However, I am aware of the possible limitations of my argumentation.

Firstly, my argumentation seems more theoretical than concrete. This is notably due to the nature of the topic, democracy, which is in itself an ideal type. This was a conscious choice that I made as it allowed my paper to explore why AI applications pose systemic challenges to democracy, i.e. why they challenge the very assumptions of democracy. Perhaps future research could thus investigate why the concrete electoral laws that protect citizens' autonomy, equality in participation and the public forum such as laws regulating campaign advocacy are currently unable to respond to the challenges posed by AI-enabled manipulation.

Secondly, my argumentation presumes that all democracies are similar, and hence that democratic principles and presumptions operate similarly everywhere. This allowed my argumentation to be relevant to democratic systems in general, and not to specific democratic systems. However, I am aware that democracy is always dependent on context and develops differently in dissimilar countries.⁵⁴ Therefore, I believe that regulatory responses to AI-enabled manipulation should be tailored to each democracy or at least leave a margin of appreciation to each democracy in case of international responses.

Thirdly, I am aware that AI-enabled manipulation is only one of the many ways allowing to influence citizens' decision making, and will continue to develop as AI develops. I also acknowledge that AI applications disrupt other foundations of democracy, such as fundamental rights, than that explored in this paper. My thesis does thus not provide a definitive or complete answer as to why AI disrupts democracy.

Lastly, as AI-enabled manipulation is posing unprecedented threats to the foundation of our legal systems, democracy, it is up to us to seize such threats as opportunities for democracy to develop and strengthen its values and principles.⁵⁵

⁵⁴ Pratchett (n 19) 32.

⁵⁵ Pratchett (n 19) 33.