Global Campus Europe

UNIVERSITY OF COIMBRA

European Master's Programme in Human Rights and Democratisation
A.Y. 2021/2022

Social media and democratic elections: a dangerous cocktail?

Towards the prevention of undue influence on voters' political opinions in the EU.

Author: Anna Bisoffi
Supervisor: Jónatas Machado

ABSTRACT

The emergence of social media revolutionized the public sphere and the informational landscape, and despite their pervasive presence and connection to fundamental human rights, there are concerns about whether these platforms have been adequately regulated to reach conformity with democratic values. This work will investigate if and how social media may negatively impact the public political discourse, offering the premise to develop recommendations for European policies to avoid undue interference in the formation of voters' political opinions in relation to elections and voting processes.

After an introductory section, to contextualize the debate by providing an overview of how and why social media may negatively impact democracy, we will examine the regulation of political discourse in traditional media compared to online platforms. Then, we will bring our attention to the EU framework. In the absence of comprehensive legislation on political speech on social media, the discipline will be reconstructed through the analysis of different sectorial initiatives, including the recently approved Digital Markets Act and Digital Services Act. Subsequently, we will reconstruct the problems that need to be tackled to prevent interference with the electoral cycle, namely polarization, hate speech, targeted political advertising, foreign interference, and disinformation. The final part of this work will trace the areas of solutions in which the European Union could take action, including education and awareness raising campaigns, institutional strengthening policies, technological fixes, anti-trust law, and other regulatory measures, in order to reach conclusive recommendations.

TABLE OF CONTENT

INTRODUCTION

Social media are a relative novelty to our democracies, and their emergence revolutionized the public sphere and the informational landscape. Despite their pervasive presence in everyday life and their link to fundamental rights such as access to information and privacy, there are concerns about whether they are adequately regulated to reach conformity with democratic values.

The ultimate objective of this research is to reach evaluative conclusions and possible recommendations for European policies on making social media democracy-friendly, specifically to prevent undue interference from being exercised on the formation of voters' political opinions. This work will address the question: Do social media negatively impact the public political discourse, and, if so, how? Moreover, if that is the case, which laws or policies could be adopted at a European level to avoid this interference?

The question posed is not new, but most research has focused on younger democracies. However, as cases like the 2016 US elections and Brexit referendum have exposed, Western democratic societies are not immune from the threat social media pose to a fair political discourse around elections and voting processes. The digital nature of this issue makes it hard to tackle; from this stems the need for international action in which the European Union could lead the way, as it has done in the past on similar topics. In recent years, Ursula von der Leyen, President of the European Commission, has emphasized the necessity of regulating digital platforms in order to safeguard European democracies from outside interference and destabilisation.

The detrimental effects of social media hinder the right to participate in political affairs, which should be exercised "without undue influence or coercion of any kind which may distort or inhibit the free expression of the elector's will," and entails that "voters should be able to form opinions

independently, free of violence or threat of violence, compulsion, inducement or manipulative interference of any kind."[1]

The purpose of the first section is to contextualize the debate by underlining the link between social media platforms, elections, and public political discourse. A general overview of the ways in which social media may negatively impact democracy and human rights will be provided, along with a rundown of the distinctive features of these platforms that make them particularly dangerous and prone to misuse.

Following this, in the second section, we will examine the regulation applicable to political speech on social media. First, we will observe how political discourse was disciplined in traditional media compared to online platforms. Then, we will bring our attention to the EU framework. As there is no comprehensive legislation on social media on a European level, the discipline needs to be reconstructed through the analysis of different initiatives, from the e-commerce directive to the European Democracy Action Plan. The last part of this second section will be dedicated to draft legislation initiatives brought forward by the European Commission on topics ranging from competition law to political advertising rules, responding to the call for new social media regulation, including the recently approved Digital Markets Act and Digital Services Act.

After analysing the current European framework, we will look at potential future steps. In the third and last section, we will address the problems that need to be tackled regarding the research question, referring to both external and internal interference at every stage of the electoral cycle and seeing how these phenomena relate to social media and threaten to hinder democratic elections. The final part of this work will trace the areas of solutions in which the European Union could take action, including education and institutional strengthening policies, technological fixes, and regulatory

[1] UN Committee on Human Rights, General Comment 25, 'The Right to Participate in Public Affairs, Voting Rights and the Right to Equal Access to Public Service', 1510th meeting (fifty-seventh session), 12 July 1996.

measures, with the aim of reaching conclusive recommendations. We'll see how the issue's complex nature necessitates a multidisciplinary approach in order to achieve a satisfactory systematic strategy.

METHODOLOGY

The methodology employed has been primarily qualitative desk-research. For the purpose of framing the topic, the first section was based on the analysis of academic literature, journal articles, and institutional reports from different fields of study. In an everchanging sector, such as the one we are discussing, there is a need to refer to current sources, such as news articles, in order to keep the research on the topic up to date. The second section, on the regulatory framework, was built on the analysis of European relevant policy documents and other primary sources, with the aid of previous academic works. In order to propose recommendations, the third section was written with the support of journalistic investigations, institutional reports, and previous scholars' research.

SECTION I - SOCIAL MEDIA AND DEMOCRATIC ELECTIONS: A RISKY COMBINATION?

1. Introduction

This section has the aim of contextualizing the debate on social media for what relates to human rights and democracy. This will be achieved by underlining the link between these platforms, elections, and public political discourse, and by providing a general mapping of the ways in which social media may also negatively impact democracy and human rights. Lastly, the section will close with an overview of the features of these platforms that make them particularly dangerous and prone to misuse.

2. Social media, human rights and democratic elections: what is the connection?

Living in the wake of the fourth industrial revolution[2] means that barely any aspects of our lives are spared from the presence of technology, consequently, the human rights sector is not exempt from this influence. Specifically, social media use is linked, for better or worse, with many fundamental human rights. The right to privacy is often at the centre of public debate, however, freedom of expression and consequentially the right to access information, freedom of association, and many others are also strongly connected to these platforms.

Social media are, as expressed in the name itself, a tool that can be used in different ways; they can be a mean for self-expression, to promote human rights and democracy, to engage the public and contribute to informed participation. These platforms have been instrumental in social movements

[2] Klaus Schwab, *The Fourth Industrial Revolution* (Crown Publishing Group 2016) <https://www.penguin.co.uk/books/304971/the-fourth-industrial-revolution/9780241300756> accessed 21 June 2022.

such as Me Too and Occupy and the Arab Spring to spread unifying messages and facilitate logistics[3]. Subverting traditional protocols, social media create a direct line of contact between political representatives and their electorate favouring engagement even in sectors of the populations historically left out or less interested in politics.

Social media are great equalizers in the field of access to information and self-expression in relation to politics and elections, this can mean educated voters, better engagement of a wider portion of the electorate in the public political discourse, which in turn can result in bigger turn-out. They give the chance for minority voices to be heard and become part of the public dialogue through online mobilization. Studies have also shown how social media make it possible for candidates to run for office without substantial financial backing, making it easier for less established parties or people who belong to politically underrepresented categories to enter the political arena[4]. Broadly speaking, the Internet creates what have been called deliberative opportunities[5], a place to exchange views and take part in discussions, and that is the idea at the core of modern democracies, founded on the idea that the more freedom of expression, the more pluralism, the better quality of our democratic institutions. Or so we thought.

At the same time, social media have also been used with opposite goals: to track human rights defenders, to spread hate and organize human rights violations, as was the case of the Rohingya genocide, and to unduly influence people and elections. Many extremist politicians have been known to resort to social media as their main campaign instrument, from Brasil's Bolsonaro using WhatsApp

[3] Zeynep Tufekci, 'How Social Media Took Us from Tahrir Square to Donald Trump' [2018] MIT Technology Review <https://www.technologyreview.com/2018/08/14/240325/how-social-media-took-us-from-tahrir-square-to-donald-trump/> accessed 24 May 2022.

[4] Maria Petrova, Ananya Sen and Pinar Yildirim, 'Social Media and Political Contributions: The Impact of New Technology on Political Competition' (2021) 67 Management Science 2997.

[5] Holly Ann Garnett and Toby S James, 'Cyber Elections in the Digital Age: Threats and Opportunities of Technology for Electoral Integrity' (2020) 19 Election Law Journal: Rules, Politics, and Policy 111.

to reach voters[6], to Philippines' Duterte narrative on the war on drugs through Facebook[7], without forgetting Italy's Salvini that was quoted saying "Thank God we have the internet, thank God we have Facebook"[8].

As with any new phenomenon, sometimes the effects of novelty are overblown, and many myths on social media have already been debunked; therefore, this paper does not aim to give an apocalyptic view of these platforms. It looks likely that social media are here to stay, and the significant contributions that they have brought to society should not be ignored.

As we have already seen, several fundamental human rights are affected by social media, concerning our democratic institutions, the most relevant one being the right to participate in public affairs, as described in General Comment 25 of the Committee on Human Rights: "voters should be able to form opinions independently, free of violence or threat of violence, compulsion, inducement, or manipulative interference of any kind". The question posed by Carole Cadwalladr, an investigative journalist who played an essential role in uncovering the Cambridge Analytica data scandal, is simply: can we ever have a free election again?[9]

Genuine periodic elections are essential to ensure the accountability of representatives elected through the will of the people to exercise the legislative or executive powers, elections and voting broadly intended should also be free from undue influence on the citizen's will. Elections themselves are part of a cycle, an electoral process, of which voting is but one essential phase through which the

[6] Rafael Evangelista and Fernanda Bruno, 'WhatsApp and Political Instability in Brazil: Targeted Messages and Political Radicalisation' (2019) 8 Internet Policy Review; Caio Machado, 'WhatsApp's Influence in the Brazilian Election and How It Helped Jair Bolsonaro Win' [2018] Council on Foreign Relations <https://www.cfr.org/blog/whatsapps-influence-brazilian-election-and-how-it-helped-jair-bolsonaro-win> accessed 15 June 2022.

[7] Lauren Etter, 'Rodrigo Duterte Turned Facebook Into a Weapon, With a Little Help From Facebook' *Bloomberg News* (7 December 2017) <https://www.bloomberg.com/news/features/2017-12-07/how-rodrigo-duterte-turned-facebook-into-a-weapon-with-a-little-help-from-facebook> accessed 16 June 2022.

[8] 'League's Salvini: Center Right Ready to Run Italy' (*POLITICO*, 5 March 2018) <https://www.politico.eu/article/matteo-salvini-leagues-center-right-ready-to-run-italy/> accessed 25 May 2022.

[9] Carole Cadwalladr, *Facebook's Role in Brexit — and the Threat to Democracy* (2019) <https://www.youtube.com/watch?v=OQSMr-3GGvQ> accessed 16 May 2022.

electorate is able to influence government and policies. Within this cycle, there are other important elements such as access to information about the democratic process, the candidates and their programmes, as well as public political discourse. All of these elements are, nowadays, deeply connected with the Internet and social media.

Social media, as defined by the Merriam-Webster dictionary, are "forms of electronic communication (such as websites for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (such as videos)"[10]. The first recorded use of this expression with its current meaning dates back to 2004, which coincides with the birth of Facebook.

It is essential to note that social media were not originally meant to provide information or factor into public political discourse, at least not on the scale that they have. These platforms were born to connect people with friends and family, with the primary goal of cultivating relationships at a distance between people that knew each other. As we know, social media evolution brought them well beyond their original goal with an unfolding that was hardly imaginable at the time of their creation, and therefore, it should come as no surprise that they are mostly unfit to promote a fair public sphere[11].

Today we live in a world where 58.7% of the global population can be considered active social media users, and the percentage remains accurate even looking at Europe separately[12] . In 2021, 48% of U.S. adults stated they get news from social media 'often' or 'sometimes', with Facebook being the most popular platform for this activity, followed by Youtube and Twitter[13]. Statistics have also shown that it is primarily the younger generations of adults (age 18-29) to rely on such means in Europe, while

[10] 'Social Media' <https://www.merriam-webster.com/dictionary/social+media> accessed 15 May 2022.

[11] Facebook was induced to change its mission statement in 2017 from "Making the world more open and connected" to "Give people the power to build community and bring the world closer together" after critics noted the lack of positive proposition and the neutrality of the original motto towards societal gains.

[12] Mason Walker and Katerina Eva Matsa, 'News Consumption Across Social Media in 2021' [2021] *Pew Research Center's Journalism Project* <https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/> accessed 15 May 2022.

[13] ibid.

concurrently being less satisfied with news media and reporting scepticism towards their performance.

Some refer to the span of years from 2007 to 2012 as the golden age of social media; in these years, it was barely conceivable that they could ever become an instrument for anti-democratic action. Social media made the Arab spring possible and were "hailed as tools of freedom"[14]; they were seen as inherently democratizing instruments that could provide access to information and a place for citizens' mobilization. Those who spoke out at the time to express their fears and doubts faced backlash from the overwhelming majority[15]. In 2013, MIT Technology Review's front cover read 'Big Data will save politics'. In 2018, the same journal revisited the topic in a series of essays concluding with a letter by the editor-in-chief titled 'Why the pessimists are winning, for now' in which he affirmed that, at the time of the writing (just five years after the original articles), technology felt as likely to destroy politics as it did to save it[16].

Then came the algorithm, and the artificial intelligence behind it, which is uniquely devoted to pursue the profit agenda by guaranteeing the highest frequency of human interaction. Since social media replaced the chronological criteria for content presentation with a new structure based on recommended and sponsored posts, no home feed has looked the same. Some authors have referred to this moment as the 'tower of Babel' of the modern era. To have a dialogue with someone, there needs to be, as a premise, a common ground for discussion: this includes a common interpretation of reality or facts that both parties regard as true. With the introduction of personalized feeds, this became ever so difficult. Partisan news is not a novelty brought about by social media, but at times distinguished by traditional media, it was a more transparent and evident phenomenon. One could

[14] Tufekci (n 3).

[15] ibid.

[16] Gideon Lichfield, 'Why the Pessimists Are Winning, for Now' [2018] *MIT Technology Review* <https://www.technologyreview.com/2018/08/22/140680/why-the-pessimists-are-winning-for-now/> accessed 15 May 2022.

buy a variety of newspapers or could switch through different TV and radio channels; now, this partiality is covert and difficult to perceive as we have no access to another person's feed.

What happens now is that our perception of reality is influenced by the online world we navigate, and if that differs, our 'truths' may differ. We know this influence has had its effects even in Western democracies, as could be seen during the Brexit campaign[17], and we know how easily our home feeds affect our behaviour when it comes to elections and campaigns[18].

Social media is a relative novelty to our societies, and despite its pervasive presence in everyday life[19] and its importance to fundamental human rights such as self-expression, access to information, and privacy, there are doubts about them being adequately regulated to reach conformity with democratic values[20].

3. An overview of the social media phenomena that negatively impact democracy

A variety of occurrences on platforms such as Twitter and Facebook have the potential to be harmful towards democratic institutions and the free formation of political will. In this passage, we will look at the most spread and dangerous of them and how they impact human rights inherent to the participation in public affairs.

The algorithm learns from a user's habits, likes and dislikes, and it begins to cater their feed to their preferences[21]. Consequently, this creates an echo-chamber, or filter bubble, that personalizes our intake of news. Some argue that this mechanism re-enforces our beliefs and may lead to polarisation,

[17] Cadwalladr (n 9).

[18] As was discovered in an experiment run by Facebook in 2010, a slight change in people's feed increased voters' turn-out around by around 340,000 citizens in the 2010 US congressional elections. Zoe Corbyn, 'Facebook Experiment Boosts US Voter Turnout' [2012] Nature <https://www.nature.com/articles/nature.2012.11401> accessed 25 May 2022.

[19] Luciano Floridi (ed), *The Onlife Manifesto* (Springer International Publishing 2015) <http://link.springer.com/10.1007/978-3-319-04093-6> accessed 16 May 2022.

[20] Thomas Wischmeyer, 'Making Social Media an Instrument of Democracy' (2019) 25 European Law Journal 169.

[21] Helen Margetts, 'Rethinking Democracy with Social Media' (2019) 90 The Political Quarterly 107.

but research shows that this effect may have been overestimated. It appears that while echo-chambers do indeed exist, social media, compared to other media, is still the source that provides for more differentiated voices and points of view. Nonetheless, research on this topic is still ongoing[22].

Micro-targeting is the practice that takes advantage of such algorithms and big data and translates it into an advanced state of digital advertising technology that produces political advertisements highly catered to the target thanks to personal data[23]. "The audiences that flow through social media platforms represent perhaps the most comprehensively measurable audiences in the history of media"[24] this was brought to the public's attention with the Cambridge Analytica scandal. The experts were not able to confirm or quantify the effect that this kind of political tactic had on voters, but it is hard to see how this would not qualify as undue influence needing better regulation.

Other infamous protagonists of the public debate are misinformation and disinformation, often called by the incorrect name of fake news. The Code of Practice on Disinformation by the European Commission defines them as "verifiably false or misleading information created, presented and disseminated for economic gain or to intentionally deceive the public" and it adds that "it may have far-reaching consequences, cause public harm, be a threat to democratic political and policy-making processes, and may even put the protection of EU citizens' health, security and their environment at risk"[25]. The spread of false information can undermine democratic governance that relies on citizens' access to quality information for decision-making since when misinformation is present, "no matter

[22] Adam Piore, 'No, Big Tech Didn't Make Us Polarized (but It Sure Helps): Social-Media Bubbles Tell Only Part of the Story of Why We're so Divided. The Rest Is in Our Heads.' [2018] MIT Technology Review <https://www.thefreelibrary.com/No%2C+big+tech+didn%27t+make+us+polarized+(but+it+sure+helps)%3A...-a0554051593> accessed 24 May 2022.

[23] Brad Parscale, who served as Donald Trump's campaign manager and digital advisor, said that in 2016 he typically ran 50,000 ad variations each day, micro-targeting different segments of the electorate. This is ended up amounting to 5.9 million compared to Clinton's 66,000 distinct ads tested. Alex Howard, 'US Election Campaign Technology from 2008 to 2018, and Beyond' [2018] MIT Technology Review <https://www.technologyreview.com/2018/08/22/140643/us-election-campaign-technology-from-2008-to-2018-and-beyond/> accessed 24 May 2022.

[24] Philip M Napoli, *Social Media and the Public Interest: Media Regulation in the Disinformation Age* (Columbia University Press 2019).

[25] Tackling online disinformation, https://ec.europa.eu/digital-single-market/en/tackling-online-disinformation

who's spreading it, people lose faith in the trustworthiness of their social institutions or perhaps even in the very idea of truth in politics[26]". This attitude proves to be an advantage for authoritarian political actors in elections and weakens democratic ones[27].

Hate speech represents expressions of incitement to hate and/or discriminate, directed towards certain people due to characteristics, real or perceived, connected to their identity, for instance, their nationality or their sexual orientation. Hate speech has become an instrument of political manipulation, used by politicians themselves or their supporters to reinforce division or harass the opposition[28] and they found a great tool in social media.

Lastly, astroturfing is defined as the "organized activity that is intended to create a false impression of a widespread, spontaneously arising, grassroots movement in support of or in opposition to something (such as a political policy) but that is in reality initiated and controlled by a concealed group or organization (such as a corporation)"[29]. This phenomenon was not born on the Internet, as it often happens, social media was the perfect ground for it. The false impression that is created by astroturfing distorts the perception of political actors: what feels like a majority and an existing consensus could very well be a relatively small amount of people, being particularly active online acting in unison or through sock-puppets[30], or even bots. Bots are used for various ends online[31], political bots are "automated social media accounts, often built to look and act like real people, in

[26] Zack Beauchamp, 'Social Media Is Rotting Democracy from Within' [2019] *Vox* <https://www.vox.com/policy-and-politics/2019/1/22/18177076/social-media-facebook-far-right-authoritarian-populism> accessed 15 May 2022.

[27] ibid; Ronald J Deibert, 'The Road to Digital Unfreedom: Three Painful Truths About Social Media' (2019) 30 Journal of Democracy 25.

[28] As Zack Beauchamp put it 'Pro-democracy politicians are much less likely to benefit from these kinds of illiberal messages. Their core supporters are much more likely to hold basic democratic commitments to equality and freedom, and thus they will lose rather than win support by stoking prejudice.' Beauchamp (n 26).

[29] 'Astroturfing' <https://www.merriam-webster.com/dictionary/astroturfing> accessed 25 May 2022.

[30] This is the practice of creating multiple users accounts for the same person to conduct disguised activity.

[31] For an analysis of possible use of bots with democratic aim Samuel C Woolley, 'Bots and Computational Propaganda: Automation for Communication and Control' in Nathaniel Persily and Joshua A Tucker (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press 2020) 96 <https://www.cambridge.org/core/books/social-media-and-democracy/E79E2BBF03C18C3A56A5CC393698F117> accessed 16 May 2022.

order to manipulate public opinion"[32] and to put in practice what is called computational propaganda. They can be used to boost the number of likes, shares and comments on specific content, faking momentum, or to spread information, real or not, and to harass the opposition, journalists, activists[33].

From everything seen so far, it is clear how social media nowadays have a significant role in campaigns and elections, which are crucial to a functioning democracy. As stated by Kofi Annan: "while democracy must be more than free elections, it is also true […] that it cannot be less"[34]. Especially since "trends in the public's media use become most apparent during periods of heightened political awareness, such as during political campaigns"[35].

4. Characteristics of social media that make them prone to anti-democratic use

As we have seen, the beginning of the era of social media was full of hope, and only few people could see beyond the circumstances at the time, but those who did realized that "power always learns"[36] and it didn't take long for powerful actors to take back control of social media and for them to be put to anti-democratic use. Joshua Tucker said it best in an interview: "ironically, the very same affordances that made social media useful for pro-democracy activists also make it a valuable tool for those who harbour anti-democratic sentiments in democratic societies"[37].

[32] Woolley (n 31).

[33] ibid.

[34] Former UN Secretary General Kofi Annan, Speech at the International Conference ''Towards a Community of Democracies'' (Warsaw, 27th June 2000). Press Release SG/SM/7467

[35] Diana Owen, 'The Past Decade and Future of Political Media: The Ascendance of Social Media', *Towards a New Enlightenment? A Transcendent Decade* (BBVA OpenMind 2019).

[36] Howard (n 23).

[37] Charles Koch Foundation, Interview with Joshua Tucker, 'How Does Social Media Impact Democracy?' (3 March 2021) <https://charleskochfoundation.org/stories/how-does-social-media-impact-democracy/> accessed 25 May 2022.

There are various features that make social media a good ground to foment antidemocratic behaviour, some are unique to them, and some are not, however, all of them are connected in a way that could be described as a perfect storm.

Social media are characterized by their low cost and open access; this means that they make it easier to communicate, collaborate and organize. This is not intrinsically a negative consequence, quite the contrary, grassroots movements came to light thanks to these platforms that offered great opportunities for mass political engagement[38], but this is true also for the planning of harassment, anti-democratic actions and human rights violations. They are a platform where content can be shared at great speed, if not immediacy, with a potentially endless number of recipients.

For these same reasons, the lack of barriers and readiness to create content and publish it at low costs, the Internet prompted a crisis in news media by removing its traditional gatekeepers: editors and publishers, governments, academic institutions, and professional journalists[39]. While democratizing the spread of news, this also meant a drop in quality of reporting and checks on content that prevented extreme anti-democratic voices from being heard on mainstream media[40]. This process was highly accelerated by social media, which are driven by virality.

The way the business model for these tech giants is built means that they make revenue by keeping people on the site, which incentivizes content that keeps attention. Unfortunately, the content that succeeds in this is the one that plays on people's emotions, and that appeals to their outrage. Dr. Brian Ott at Missouri State University and Dr. Greg Dickinson at Colorado State University have identified three characteristics of content on Twitter that has success in this: simplicity, impulsivity, and

[38] Owen (n 35).

[39] On the role of journalism in gatekeeping Andra Brichacek, 'Six Ways the Media Influence Elections' *School of Journalism and Communications-University of Oregon* (2016) <https://journalism.uoregon.edu/news/six-ways-media-influences-elections> accessed 11 May 2022.

[40] Charles Koch Foundation, Interview with Tucker (n 37).

incivility[41]. Social media platforms, some more than others, have always struggled to present complex ideas, and politics are anything but simple.

On the issue of hate speech, a study commissioned by UNESCO found the factors of anonymity, real or perceived, permanence and transnationality or cross judicial character, typical of the Internet, lead to a perception of immunity, lack of accountability, and therefore sanction, all of which exacerbates the damage inflicted on the victims and society[42].

When it comes to elections and democratic political discourse, it has been said that it is easier to spread misinformation on social media than to correct it, and easier to inflame social divisions than to mend them. Information overload intrinsically benefits anti-democratic authoritative politicians since it increases political apathy and undermines trust in established institutions[43].

Zack Beauchamp has made the divided a politician's use of social media into two types: normal use and abuse. "Normal use of social media is simply an online extension of typical democratic campaign tactics: paying for ads spreading your message or uploading a campaign video to YouTube. Abuse involves the deliberate spread of false information, attempts to undermine faith in established reality, trolling, and harassment" [44].

These phenomena and characteristics distinctive of social media might not be enough to hinder elections or put democratic values at risk by themselves, but when combined with a public sphere that's characterized by social distrust, weak institutions, and detached elites, like Western

[41] Brian L Ott and Greg Dickinson, 'Twittering Away Our Deliberative Capacity: Social Media and the Threat to Democracy' [2021] Research Outreach <https://researchoutreach.org/articles/twittering-deliberative-capacity-social-media-threat-democracy/> accessed 16 May 2022.

[42] Iginio Gagliardone and others, *Countering Online Hate Speech* (UNESCO 2015) 13.

[43] Beauchamp (n 26).

[44] ibid.

democracies have experienced in the last decade, it becomes dangerous, and measures should be taken to reduce the unintended consequences[45].

5. Conclusions

As seen in this section social media are closely linked to public political discourse and the right to participate in public affairs, along with many other important human rights. This connection is sometimes an opportunity and other times it becomes a threat or a challenge. Many phenomena connected to social media, namely disinformation and micro-targeting, put the democratic process of elections and the individual right to form an opinion without undue influence at risk. These negative consequences are accelerated by some characteristics of these platforms that may turn social media into a threat to our democracies if not regulated adequately.

[45] Tufekci (n 3).

SECTION II - THE EU FRAMEWORK ON SOCIAL MEDIA: A REGULATORY JIGSAW PUZZLE?

1. Introduction

In this section, we will look at the regulation that is applicable to social media. First, we will observe how political discourse was disciplined in traditional media compared to online platforms. Then, we will bring our attention to the EU framework. As there is no comprehensive legislation on social media on a European level, the discipline needs to be reconstructed through the analysis of different initiatives, from the e-commerce directive to the European Democracy Action Plan. Each one of these provisions takes a different approach and gives different insight to the problem at hand given its complexity. Finally, we will see that the call for new social media regulation has recently brought to a series of draft legislation initiatives by the European Commission, on topics ranging from competition law to political advertising rules.

2. Political discourse in traditional media

Legacy media, print or broadcast, dominated the information landscape before the advent of the Internet. Newspapers, television, and radio were the space for politicians to debate and present their electoral programme, and for citizens to access this information.

Television and radio were regulated initially on the basis of spectrum scarcity, the idea that since the bandwidth allowed only a finite number of transmissions, the State should intervene and allow licenses to broadcasters. When it came to political speech most States imposed some requirements to these platforms, such as equal access and fair distribution of time and space between the different political parties, and in some countries a cooling-off period before election day. It should be noted

that this applies to both the public and private sector[46]. This was done to enhance electoral integrity in consideration of the democratic function that media perform towards the public interest, but also considering the power of influence that they obviously held[47]. The widely accepted idea was that broadcasting was more than just a medium in the formation of the public's opinion, it was also a factor in that process[48].

Even with the advent of cable television, these set of rules did not falter. Political speech on these platforms is still regulated, the debates between candidates follow rules, they need to be balanced.

Despite social media having the same influence power as legacy media, or arguably greater, they are not equally regulated, there is a normative deficit on the subject. National courts have tried to apply the same laws by virtue of interpretation but that is hardly a sufficient solution due to the differences between these media. Social media do not have editorial oversight on news content since they function as 'mere' conduit, which makes it harder for them to supervise but also for third parties to hold them accountable. They have ways of personalizing the content for the users, who do not have control on the content they are shown; this contrasts with consumers of traditional media that have to proactively make a choice on what to see or read[49].

[46] Unlike in the U.S., public tv and radio is still present and powerful in Europe: Francis Fukuyama and Andrew Grotto, 'Comparative Media Regulation in the United States and Europe' in Joshua A Tucker and Nathaniel Persily (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press 2020) <https://www.cambridge.org/core/books/social-media-and-democracy/comparative-media-regulation-in-the-united-states-and-europe/0E4F255ADA3FC81BDC4365FF10DFDF3A> accessed 17 June 2022.

[47] David Yanagizawa-Drott, 'Propaganda and Conflict: Evidence from the Rwandan Genocide *' (2014) 129 The Quarterly Journal of Economics 1947; Maja Adena and others, 'Radio and the Rise of The Nazis in Prewar Germany' (2015) 130 The Quarterly Journal of Economics 1885; Alexandra A Siegel, 'Online Hate Speech' in Joshua A Tucker and Nathaniel Persily (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press 2020) <https://www.cambridge.org/core/books/social-media-and-democracy/online-hate-speech/28D1CF2E6D81712A6F1409ED32808BF1> accessed 17 June 2022.

[48] Wischmeyer (n 20).

[49] Gianmarco Gori, 'Social media ed elezioni. I limiti del diritto e il rischio di una modulated democracy' (2017) 26 Informatica e diritto; Dipayan Ghosh, 'Are We Entering a New Era of Social Media Regulation?' [2021] Harvard Business Review 5.

Given the peculiarities of social media, the regulation cannot be achieved by simply extending the sphere of application of laws on traditional media[50] to these platforms, this is why both scholars and civil society have been pushing for stronger and specific regulation.

3. The European Union Framework

The European Union, unlike some of its Member States[51], has not adopted any regulation specifically dedicated to social media, therefore the framework is fragmented and spread across a multitude of directives, regulations, and other initiatives. In this section, we will try to give a comprehensive recount of the various acts on the topic at hand. There are more specialized norms connected to social media, for example on terrorism, copyright and cyberbullying, which fall beyond the scope of this paper.

Given the complexity of the issues at hand, as presented in this paper, these policies are often aimed at tackling one of these problems, be it hate speech, political advertising online or disinformation. It is worthy of note that a package of proposals relevant to the topic might be on the verge of being adopted. There is an expectancy that it is Europe to lead the way in this matter, the same way it did with the GDPR and is expected to do with AI Regulation. In virtue of the so-called Brussels effect[52], and thanks to its market of around 450 million consumers, European regulation might have far-reaching influence[53].

[50] Wischmeyer (n 20).

[51] For instance, Germany, France, and Austria.

[52] Anu Bradford, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020) <https://oxford.universitypressscholarship.com/10.1093/oso/9780190088583.001.0001/oso-9780190088583> accessed 17 June 2022.

[53] Adam Satariano, 'E.U. Takes Aim at Big Tech's Power With Landmark Digital Act' *The New York Times* (24 March 2022) <https://www.nytimes.com/2022/03/24/technology/eu-regulation-apple-meta-google.html> accessed 6 June 2022.

3.1. General Data Protection Regulation

The General Data Protection Regulation[54], known simply as GDPR, was a landmark act adopted in 2016 by the EU, which came into force in 2018. The GDPR represents a comprehensive privacy law which imposed stringent standard on the collection and use of personal data, in both the public and private sector. This act had the aim of bringing data law up to the digital age and to respond to growing concerns about the accumulation of data by big tech companies such as social media. Europe set the lead for others to follow and used its influence to impose privacy protection conditionality on trade agreements.

The GDPR made it possible for people to demand information on their online data, and created a legal mechanism to obtain, under certain conditions, the deletion of their personal data from online databases[55]. Moreover, it also put restriction on how companies obtain and manage this information. It made clear who should have access to which data and the lawful use that can be made of them (based on the purpose limitation principle), stating that the processing of data should be lawful, fair and transparent. The regulation focuses on the idea of free and informed consent, based on clear and affirmative action. With this regulation the EU took its first step and recognized the importance of human rights[56] over profit of big tech companies.

The regulation harmonized data protection practices at the EU level, and it has extra-territorial application as it applies to personal data of EU residents, regardless of where the organization is based or where the processing of data takes place. This means that digital companies need to comply if they

[54] 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) OJ L 119, 4.5.2016, p.1–88 ELI: http://data.europa.eu/eli/reg/2016/679/2016-05-04. The GDPR should be complemented by the ePrivacy Regulation (ePR) which, if approved, would replace the e-privacy directive (Directive 2002/58/EC).

[55] Implicitly recognizing the 'right to be forgotten'.

[56] The right to privacy (Article 8 of the EU Charter of Fundamental Rights).

want to offer services or products in Europe. Fines for failure to comply with these norms can go up to 4% of annual global turnover or 20 million euros, whichever is greater.

The European Commission also prepared specific guidance on the application of GDPR in the electoral context, stating that data protection "has become a key issue not only for individuals but also for the functioning of our democracies because it constitutes a serious threat to a fair, democratic electoral process and has the potential to undermine open debate, fairness and transparency which are essential in a democracy[57]."

According to this document, the infringement of these provisions in the electoral context could have severe effects on citizens' rights, both in gravity and in number, which could lead to high fines "in particular considering the importance of the issue of citizens' trust for the democratic process."[58]

The newly established European Data Protection Board, which brings together all national data protection authorities, as well as the European Data Protection Supervisor, plays a key role in the application of the GDPR by producing guidelines, recommendations, and best practices.

The GDPR applies to political parties (national and European), political foundations, data analytics companies and public authorities involved in the electoral cycle. Therefore, these actors need to process personal data in fair, transparent and lawful way, for specified purposes, and cannot use them for further ends incompatible with the original aim. Moreover, when obtaining data from third parties they will need to perform due diligence to check on the lawfulness of the collection before using it.

[57] (EC) European Commission. 2018. Commission Guidance on the Application of Union Data Protection Law in the Electoral Context. COM(2018)638 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0638

[58] Ibidem.

This provision has had an indirect effect on the present topic by prohibiting the processing of political opinions[59], since they are regarded as sensitive data, they benefit from a more rigorous regime[60]. The regulation does not, however, contain a specific reference to political micro-targeting, as the scope of application of the regulation is primarily commercial advertising. When it comes to profiling[61], as made evident by the Cambridge Analytica case, social media can be a useful source to create voters' profiles. These profiles can help determine which citizens are more easily influenced, either to vote in a certain way or even to not vote at all, and enable political organizations to impact the turnout and outcome of elections. The GDPR forces all data controllers, including political ones, to disclose whether they use such techniques and on their consequences. Data protection impact assessment is mandatory when the processing is likely to result in a substantial risk to the rights and freedoms of individuals[62], which includes the electoral process.

In relation to elections, the individual has the right to access their personal data, to request the deletion of their personal data if the processing is based on consent and that consent is withdrawn, if the data is no longer needed or if the processing is unlawful, as well as to have incorrect, inaccurate, or incomplete personal data corrected. The GDPR also allows individuals to give mandate to an NGO to complain on their behalf.

[59] Unless the individual has given explicit, specific, fully informed consent; such information is manifestly made public by them; when they are a current or former members of the organization or in a regular contact; and when processing is needed for reasons of "substantial public interest" (GDPR, Art. 9, para. 2).

[60] GDPR art.9 para. 1.

[61] Art. 4.4 GDPR: any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.

[62] Art. 35 and 36 GDPR.

3.2. E-commerce directive

The directive on electronic commerce, or e-commerce directive[63], was adopted in 2000 and came into force in 2002. It created an Internal market framework for online services, aiming to overcome obstacles to cross-border provision of services between EU member states, and to provide legal certainty for companies and consumers alike.

The directive applies to information society services[64] established in the EU. Art. 3 establishes the principle of country of origin, meaning that companies will be subject to the rules imposed by the Member State in which they are located and will be able to provide services in the whole EU without needing to comply to other countries' legislation. Derogations to this principle are possible under a series of conditions and with restricted justifications.

This provision detailed the rules on liability limitations for intermediary service providers when it comes to third-party content. The directive distinguishes between three set of activities: mere conduit, caching, and hosting. (art. 12-15). The directive establishes that intermediaries that perform these activities are not liable for illegal material uploaded to their sites. They become responsible in the moment they become aware of it, either independently or due to third-party notice[65]. Art. 15 also sets the prohibition for States to impose a general monitoring obligation on online intermediaries, meaning that they actively seek out facts or circumstances indicating illegal activity. This framework pushed platform providers to self-regulate and moderate content through blocking, filtering, or deleting it, without any form of oversight.[66]

---

[63] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce') https://eur-lex.europa.eu/eli/dir/2000/31/oj

[64] As defined in Article 1(2) of Directive 98/34/EC: any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services. That is not to say that free service is excluded since the European Court of Justice ruled that the directive applied also to companies that gained revenue indirectly through commercial advertising, see CJEU, 27 March 2013, Case C-291/13, Papasavvas case.

[65] Notice and take-down mechanism.

[66] Which 'may also result in further risks stemming from the lack of transparency of platform's algorithms and operations' Sorin Adam Matei, Franck Rebillard and Fabrice Rochelandet (eds), *Digital and Social Media Regulation: A Comparative*

According to the e-commerce directive, social media platforms fall under the category of 'hosting providers' even though their services are different from classic hosting services, doing much more than just 'store and disseminate' information[67]. It has been said that "this approach, which gives a certain immunity to platforms, has come under increasing strain in recent years as the power of social media companies has become apparent and as their impact on democracy has become more noticeable".[68] The existence of these provisions means that Member States cannot act to not incur in a violation of the distribution of competences with the EU[69]. The liability regime is especially important when it comes to hate speech, disinformation, dark ads, micro-targeting, and other phenomena typical of social media.

The directive has been in the centre of debate and criticism, as it is not considered in line with today's online environment, and the framework is set to be reviewed through the Digital Services Act to ensure transparency and accountability of online platforms.

3.3. Audiovisual Media Services Directive

The EU's Audiovisual Media Services Directive[70], or AVMSD, originally adopted in 2010 and reviewed in 2018, sets forth common rules for audiovisual media services in the single market. Among other areas, the directive is aimed at combating racial and religious hatred and safeguarding

Perspective of the US and Europe (Springer International Publishing 2021) 95 <https://link.springer.com/10.1007/978-3-030-66759-7> accessed 12 May 2022; Daniela Stockmann, 'Tech Companies and the Public Interest: The Role of the State in Governing Social Media Platforms' [2022] Information, Communication & Society 1.

[67] BAYER Judit and others, 'The Fight against Disinformation and the Right to Freedom of Expression' 49.

[68] Andrew Puddephatt, 'Social Media and Elections' (2019) 14 Cuadernos de discusión de comunicación e información <https://unesdoc.unesco.org/ark:/48223/pf0000370634> accessed 16 June 2022.

[69] Wischmeyer (n 20).

[70] Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities.

consumers; it contains the principle of separation, aimed at clearly distinguish organic content from ads, puts restrictions on political and public issue content, and limits targeting practices to protect vulnerabilities[71].

It was revised in 2018 to extend the scope of application of some of the provisions on TV broadcasting to video-sharing platforms[72]. The revised AVMSD is the first EU-level legislation that has addressed specific content regulation on any kind of digital platform, if only for audiovisual content[73].

There have been debates on whether this directive applies to social media. The definition in Art. 1 (1) (aa) establishes two alternative criteria for the recognition of VSPs: that the provision of audiovisual content is either the principal purpose of the service, or its essential functionality[74]. Based on this second definition the directive could be applied to social media.

Up to this point, from a legal perspective, social media have been considered intermediaries, lacking control and, therefore, not responsible for the hosted content (as decided by the E-commerce Directive). This idea has long been subject to debate[75]. Given the power these platforms hold, and the control that they have on content, even if only on its organisation instead of production, the AVSM directive is applicable to them. The recital 48 reads: "in light of the nature of the providers' involvement with the content provided on video-sharing platform services, the appropriate measures to protect minors and the general public should relate to the organisation of the content and not to the content as such." The regulation is directed towards the content distribution, its systemic treatment, unlike traditional broadcaster who are accountable for all of their content[76].

[71] Judit and others (n 67).

[72] It regulates, amongst others, YouTube's content.

[73] Lubos Kuklis, 'Video-Sharing Platforms In AVMSD – A New Kind Of Content Regulation', *Research Handbook on EU Media Law and Policy* (Elgar Publishing 2021) <https://papers.ssrn.com/abstract=3527512> accessed 18 June 2022.

[74] Principle of 'technological neutrality'.

[75] Kuklis (n 73).

[76] This interpretation is in line with national approaches such as Germany's NetzDG, and is compatible with the e-commerce directive.

Video sharing platforms (VSPs) have content management obligations to take measures to comply with the directive in, amongst others, the areas of commercial communication and hate speech[77]. The AVMSD establishes a systemic obligation to provide specific tools for labelling and reporting content and it legislates for transparency.

The revised AVMSD has introduced a system of provisional derogation from the principle of freedom of reception of services from other member states[78], and set out a complex procedure to request it. Among the possible justifications are the contravention of rules on incitement to violence or hatred, and "where an audiovisual media service provided by a media service provider under the jurisdiction of another Member State manifestly [...] prejudices or presents a serious and grave risk of prejudice to public security, including the safeguarding of national security and defence." This procedure could be used in the case of interference in the process of national elections by media from other countries, and disinformation.

It has been recommended that AVMSD be updated and adapted to the online environment, "to break the vicious cycle of the ever-increasing attention harvesting and the consequent data-harvesting"[79].

3.4. Code of conduct on hate speech online

In recent years, hate speech seems to have increased, in part this has been blamed on the widespread use of social media. Having acknowledged this change, in 2016 the European Union supported the establishment of a code of conduct, in collaboration with the most important IT companies: Facebook,

[77] This shifts the EU approach to tackling hate speech online from self-regulation of platforms (Code of Conduct) to more conventional regulation based on law.

[78] Art. 3.2 AVMSD.

[79] Judit and others (n 67) 11.

Microsoft, Twitter and YouTube[80]. This code has the objective to counteract the incitement to hatred committed through computerized means through a self-regulatory, non-legislative measure. The platforms have undertaken, by signing it, to prepare clear and effective procedures for examining the reports on content, so as to be able to remove such content or disable access to it, possibly within 24 hours of reception[81]. The application of self-regulation rules is assessed on an annual basis, which has shown a progressive improvement.

In 2017, the European Commission adopted a communication containing guidelines and principles, aimed at online platforms, to step up the fight against illegal content online[82]. This document provides guidelines for identifying and reporting such content, for collaboration with national authorities and for achieving greater transparency. As a follow-up to the Communication, a Commission Recommendation on Measures to Effectively Tackle Illicit Content Online was published in 2018[83], including a range of operational measures targeting states and businesses and a resolution by the European Parliament in 2020[84]. The European Union has again opted for a solution without binding effect, while reserving the right to resort to the legislative option should it prove useful.

[80] The EU Code of conduct on countering illegal hate speech online, June 2016. https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combatting-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en#theeucodeofconduct

[81] Instagram, Snapchat, Dailymotion and Google+ decided to adhere to the Code in 2018; the Jeuxvideo.com platform in 2019; TikTok in 2020; LinkedIn in 2021; and finally, Rakuten Viber in May 2022.

[82] European Commission. 2017. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Tackling Illegal Content Online: Towards an Enhanced Responsibility of Online Platforms. COM/2017/555 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A5 2017DC0555

[83] European Commission. 2018. Commission Recommendation on Measures to Effectively Tackle Illegal Content Online. C(2018) 1177 final. https://ec. europa.eu/digital-single-market/en/news/commission-recommendation-mea sures-effectively-tackle-illegal-content-online.

[84] European Parliament resolution of 25 November 2020 on strengthening media freedom: the protection of journalists in Europe, hate speech, disinformation and the role of platforms (2020/2009(INI)), 25 November 2020.

3.5. Code of practice on disinformation

In 2018, the European Commission adopted a Communication on tackling disinformation[85], which led to a code of practice on disinformation[86] to be signed by the most important tech companies later that year. This represented another instance of a self-regulatory instrument initiated by one of the bodies of the EU, and the first example of voluntary action taken by online platforms to contrast the phenomenon of disinformation and online political manipulation. According to the European Commission, disinformation "may have far-reaching consequences, cause public harm, be a threat to democratic political and policy-making processes, and may even put the protection of EU citizens' health, security and their environment at risk."

The Code listed 15 commitments organised under five pillars for action: scrutiny of advertisement placement, political advertising and issue-based advertising, integrity of services, empowering consumers, and empowering the research community.

During the period of the Parliamentary European elections in 2019, and during the pandemic in 2020 the European Commission set up a more intense monitoring mechanism, asking for monthly compliance reports by the companies, particularly from Twitter, Google, and Facebook[87].

In 2021, the Commission published guidelines following the assessment on the first period of implementation, it contained suggestions on how to improve the code and addressed shortcomings.

---

[85] European Commission. 2018. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Tackling Online Disinformation: A European Approach. COM(2018) 236 final. https://eur-lex.europa. eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0236&from=EN

[86] Born out of European Commission (2018) Final report of the High Level Expert Group on Fake News and Online Disinformation. The related Action Plan published in December 2018, requested the assistance of the European Regulators Group for Audio-visual Media Services (ERGA) in monitoring the implementation of this Code.

In 2020 the European Commission established the European Digital Media Observatory (EDMO)

The Code was signed in 2018 by Facebook, Google, Twitter, and Mozilla, and by advertisers and advertising industry, whilst other platforms followed.

[87] European Commission, Shaping Europe's digital future, Fourth set of reports – Fighting COVID-19 disinformation Monitoring Programme, December, 2020.

In June 2022, a renewed and strengthened version of the code of practice was presented, "signed by a broad range of actors, such as online platforms, players from the advertising ecosystem, fact-checkers, civil society, research, and other organisations, joining to fight disinformation by subscribing to commitments and measures relevant to their mission". Although not endorsed by the Commission, the latter has declared that the new code meets its expectations[88]. The new code was informed by the experience of the coronavirus pandemic[89] and the war in Ukraine, as well as by the results of practice over the last years.

The strengthened code contains measures on cutting financial incentives on the spread of disinformation, it covers new-found types of manipulation such as deepfakes and bots, it guarantees more transparency on political advertising by labelling efforts, and it allows more access to data for researchers. It also set up a Transparency Centre, to give citizens an overview of the implementation of the code, and a Task Force. The Task Force is composed by the signatories, representatives from the European Digital Media Observatory (EDMO), the European Regulators Group for Audiovisual Media (ERGA), the European External Action Service (EEAS), and is chaired by the European Commission. It will oversee on the implementation and provide a space to continuously review the measures contained in the code according to technological, societal, market and legislative developments.

The first report on implementation is set to come out in early 2023. The DSA, when approved, is expected to make mandatory some of the provisions, turning the code into a co-regulation effort. ERGA's report states that the 2022 code is seen "as an opportunity to test some of the proposals in

[88] List of signatories https://digital-strategy.ec.europa.eu/en/library/signatories-2022-strengthened-code-practice-disinformation#:~:text=16%20June%202022-,Signatories%20of%20the%202022%20Strengthened%20Code%20of%20Practice%20on%20Disinformation,measures%20relevant%20to%20their%20mission.

[89] Council conclusions on strengthening resilience and countering hybrid threats, including disinformation in the context of the COVID-19 pandemic, 15 December 2020. European Parliament, 'Disinformation: how to recognise and tackle COVID-19 myths', 30 March 2021. European Commission, Joint Communication to the European Parliament, the European Council, the Council, The European Economic and Social Committee and the Committee of the Regions on Tackling COVID-19 disinformation - Getting the facts right, JOIN(2020) 8 final, 2020. European Commission, Fourth set of reports – Fighting COVID-19 disinformation Monitoring Programme, 2020

the DSA related to access to data, audits, external oversight, or risk-mitigating measures[90]". It will also interlink with the upcoming measures on political advertising.

3.6. The example of Germany

Due to lack of action by the EU through mandatory means some Member States have acted on their own to contrast hate speech and disinformation. Particularly Germany led the change with the so-called NetzDG law[91], a legislation that entered into force in Germany in 2017, designed to combat abuses conducted through social networks and specifically hate speech and 'fake news'. This provision establishes the obligation for IT platforms with more than two million registered users in Germany to have an effective and easily accessible, as well as transparent, mechanism for users to report illegal content. The provider is required to eliminate the manifestly illegal content within 24 hours, in more complicated cases the time is extended to seven days, and there is the possibility of deferring to accredited self-regulation bodies. The person who considers themself harmed by contents that have not been removed or disabled can lodge a complaint with a judge, who will rule with a binding decision. Any violation of this legislation corresponds to a pecuniary administrative sanction which can amount to between five hundred thousand euros and five million euros. In addition, companies that fit the category have biannual reporting obligations to the State.

3.7. European Democracy Action Plan

Nearing the end of 2020, the European Commission issued the European Democracy Action Plan (or EDAP), following public consultations. This document focused on three main points for future

[90] ERGA, Statement about the proposed Regulation on political advertising, 2022.

[91] Netzwerkdurchsetzungsgesetz (German Network Enforcement Act) 2017

development to strengthen the resilience of European democracies and to approach the challenges that the new information landscape presents.

Firstly, it addresses the need to promote free and fair elections, this goal is to be achieved, among other means, through the regulation of political advertising and funding of political parties. The European Cooperation Network on Elections is set to enhance cooperation among Member States, including through information exchanges on elections' integrity related issues such as the cyber-security of elections. The second item on the agenda of the EDAP is to strengthen media freedom and pluralism, by supporting media through the digital transition. Finally, the commission draws attention to the problem of disinformation, presenting various practical solutions to tackle the problem, including media literacy, which led to the adoption of the renewed code of practice.

The Action Plan underlined the opportunity for legislative action in certain fields to ensure the exercise of human rights and the operation of the democratic process. It represents a shift in EU strategy, which now moves from soft instruments towards legally binding tools in the area of digital platforms' regulation.

3.8. A look into the (near?) future

Under the premise of the EDAP, the Commission proposed legal action in various areas, among them: the proposal on greater transparency in paid political advertising, the Digital Markets Act (DMA) and the Digital Services Act (DSA). This 'package' of legislations is supposed to interact with one another and complement previous provisions.

3.8.1. The Digital Services Act and the Digital Markets Act

The Digital Services Package, composed of the Digital Services Act[92] and the Digital Markets Act[93], is a set of legislations aimed to harmonize the rules on digital services at the European level, to create a safer digital space where the fundamental rights of users are protected as well as to establish a level playing field for businesses. The two initiatives were proposed in December 2020 after a wide range of consultations with different stakeholders[94], political agreements were reached in March (DMA) and April 2022 (DSA).

This ensemble of measures represents the effort of the Commission to strengthen and update a framework that remains unchanged since the early 2000s despite technological and societal evolution. This new regulations for the digital space are intended to discipline social networks, as well as online marketplaces, content-sharing platforms, and online travel and accommodation platforms.

The DSA's goal is to build a safer online environment and safeguard fundamental human rights in the digital sphere. This objective is to be achieved through transparency measures directed at online platforms, tackling illegal online content and other risks, establishing responsibilities, and a clear accountability and transparency framework for providers of intermediary services.

Following the approach of the GDPR, the provisions will apply to any intermediary offering its services in the EU, regardless of their place of establishment.

The obligations put forward in the DSA will be applied to companies proportionally to their size, nature of services, and impact. To this end the Act distinguishes Very Large Online Platforms

[92] European Commission. (2020). Proposal for a regulation of the European Parliament and of the council on a single market for digital services (Digital Services Act) and amending Directive 2000/31/EC. European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/ HTML/?uri = CELEX:52020PC0825&from = en

[93] Proposal for a regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act) COM/2020/842 of 15 December 2020.

[94] Consultations were carried out with the private sector, users of digital services, CSOs, national authorities, academia, the technical community, international organisations and the general public.

(VLOPs)[95] and Very Large Online Search Engines (VLOSEs), these companies considering their power will have to follow stricter rules and requirements.

VLOPs will have reporting obligations towards the Commission, they will have to put in place easy to access, user-friendly mechanisms to report illegal content[96], and they will have to provide transparency over the main parameters of the decision-making algorithm that is used to offer content[97]. Companies are also required to grant access to information to 'vetted researchers'.

The Digital Services Act requires digital platforms to address systemic risks by performing due diligence and publishing annual risk assessments[98]. These companies are then expected to implement measures to mitigate the identified risks, related for instance to illegal content, human rights violations, or election interference. The Commission can invite VLOPs and other stakeholders to adopt code of conducts when systemic risks are found, these codes should be the result of co-regulation efforts in collaboration with the European Commission and the European Board for Digital Services. Platforms are obligated to publish the main parameters of their recommender systems in a manner accessible for the users and to provide options, at least one of which should be something other than profiling, to empower users and fulfil transparency requirements. At the same time, the DSA does not forbid or imposes restrictions on any type of micro-targeting, only aiming for advertising transparency. The DSA sets up mechanisms for monitoring by the supervisory public authorities, as well as audits carried out by independent bodies to oversee compliance. The conclusions get published; however, they are not enforceable and do not have legal consequences.

[95] Platforms with over 45 million users.

[96] When coming from a 'trusted flagger' the report will have to be prioritised based on an internal complaint handling system.

[97] For an in-depth critical analisys of the DSA in connection to disinformation see Judit and others (n 67).

[98] Adam Satariano, 'Facebook Hearing Strengthens Calls for Regulation in Europe' *The New York Times* (6 October 2021) <https://www.nytimes.com/2021/10/06/technology/facebook-european-union-regulation.html> accessed 6 June 2022.

In response to the war in Ukraine, a new article has been inserted to introduce a crisis response mechanism, in light of the wave of disinformation brought about by the Russian invasion. On the basis of this article, it will be possible to analyse the effect on the crisis of VLOPs' and VLOSEs' operations, to proceed with steps to ensure that fundamental human rights are respected.

For what concerns content liability, the Act reiterates the provisions of the e-commerce directive, notice and action mechanism and three levels of liability[99]. Platform providers are regarded as a subcategory of hosting providers even though their services differ from classic hosting services. A typical hosting service is one that consists of the storage of information provided by and at the request of a recipient of the service, whereas 'online platform' is defined as a subcategory of hosting service that stores and disseminates to the public information. However, some argue that the expression 'store and disseminate' does not correctly convey the full length of these platform's activities. While hosting providers do not alter the content or its presentation, platforms "disseminate third-party content, but by setting the rules of dissemination, including the rules of ranking, ordering, prioritising and deprioritising, they shape the informational experience of users and thereby have a formative impact on public discourse. They are governors of the informational landscape."[100] Therefore, this classification is determined to be insufficient and unjustified by some[101].

On the other hand, the DMA tackles the problems of consumers' rights and competition laws, to ensure a competitive and fair digital sector by setting clear rules on unfair market strategies, imposing obligations, and promoting good practices.

[99] Art. 3-7 DSA.

[100] Judit and others (n 67).

[101] ibid; Yaël Eisenstat, 'How to Hold Social Media Accountable for Undermining Democracy' [2021] Harvard Business Review 5.

The Act is also directed at big platforms that act as gatekeepers in the digital economy, meaning those few that hold enough power to act as private rule-makers and control at least one of the so-called 'core platform services', these include online search engines and social networking.

To be classified as a gatekeeper a platform needs to fit one of two alternative criteria: either have had an annual turnover of at least €7.5 billion within the European Union in the past three years or have a market valuation of at least €75 billion, and secondly it must have at least 45 million monthly end users and at least 10.000 business users established in the EU. In addition, the platform must also control one or more core platform services in at least three member states.

If one of these companies contravenes the provisions of the DMA it can incur in fines that go up to 10% of its total worldwide turnover, 20% for a repeated offense. If the violations are of a systematic nature the Commission reserves the right to open a market investigation impose behavioural or structural changes if it sees fit. Unlike the system set up by the GDPR, which has shown some limitations, the regulation will be enforced by the Commission itself.

3.8.2. Proposal for regulation on the transparency and targeting of political advertising

In November 2021, on the basis of the European Democracy Action Plan the Commission proposed a draft legislation on the regulation of political advertising in its various forms, including online targeting[102]. The proposal, which is being discussed in national Parliament and at the EU level, has the objective of harmonizing the rules on transparency of political advertising and protect personal data by regulating targeting and amplification techniques. The rules, if approved, would apply not only to political advertising companies, but to any data controller using those techniques.

[102] Proposal for a regulation of the European Parliament and of the Council on the transparency and targeting of political advertising COM/2021/731 final.

Firstly, the act would give a common definition, lacking until now, of political advertising. In the proposal it is defined, in broad terms, as "the preparation, placement, promotion, publication or dissemination, by any means, of a message: (a) by, for or on behalf of a political actor, unless it is of a purely private or a purely commercial nature; or (b) which is liable to influence the outcome of an election or referendum, a legislative or regulatory process or voting behaviour."[103] To determine if a message fits into the category the Commission states that content, language employed, context, objective and means of publishment or dissemination, would be taken into consideration.

According to some scholars, this definition appears to set too broad a scope, paragraph (b) especially could be in contrast with freedom of expression as "any kind of reflection or criticism of a political system, which is typical in open democratic debate, could be captured by Article 2(2)(b)."[104] Content on controversial issue and 'hot topics' at the centre of public debate could be sufficient to influence elections' outcome or voting behaviour.

A political advertising publisher is defined as "a natural or legal person that broadcasts, makes available through an interface or otherwise brings to the public domain[105]." The provisions on transparency would, therefore, apply to all media, including newspapers, television, radio, social media, websites, mobile applications, computer games and other digital interfaces [106]. These online means are of particular importance in the context of political advertising given the cross-border nature of the services and the challenges they provide in the enforcement. As seen, the scope of application of the proposed regulation is much broader than that of the DSA.

[103] Art. 2(2) of the proposal.

[104] Susanna Lindroos-Hovinheimo, 'The Proposed EU Regulation on Political Advertising Has Good Intentions, But Too Wide a Scope' (*European Law Blog*, 23 February 2022) <https://europeanlawblog.eu/2022/02/23/the-proposed-eu-regulation-on-political-advertising-has-good-intentions-but-too-wide-a-scope/> accessed 21 June 2022.

[105] Art. 2(11) of the proposal.

[106] Recital 2 of the proposal.

It is important to note that the draft legislation, if approved, would prevent Member States from maintaining or introducing rules or measures that derogate from the provisions of the regulation, in pursuit of harmonization[107].

The ERGA put out a statement on the proposal in which it explicitly appreciates the approach of imposing obligations on the entire value chain of political advertising, and not only on online platforms, the consistency with the DSA, and the reference to national regulatory authorities, and ERGA itself, to guarantee cooperation in the oversight of the regulation[108]. At the same time, the document underlines some space for improvement, among the suggestions, are the creation of comprehensive and easily accessible political ads repositories, a common sanctioning regime, a clearer definition of political advertising, and the designation of a reference authority body at the EU level to enhance cooperation.

4. Conclusions

After having illustrated the principles behind the regulation of political speech in legacy media, it is clear that a simple expansion of the scope of application of those rules is not sufficient to adequately control public discourse on social media.

The EU framework on the matter is the result of a regulatory patchwork made up of different initiatives with various degrees of enforceability and aimed at different sectorial problems: disinformation, hate speech, anti-trust… As it is described by the European Commission itself, despite a range of targeted, sector-specific interventions at EU level, there are still significant gaps

[107] Art. 3(1) of the proposal.

[108] ERGA statement about the proposed Regulation on political advertising 2022.

and legal burdens to address[109]. While some of these voids will be filled by the upcoming legislation (DSA, DMA and more) the discipline may remain fragmented and less than coherent.

It is easy to see how electoral laws and political speech regulation were disrupted by the spread of new technologies. If democracy is 'the only game in town' we could say that the rules have changed and will likely continue to do so, along with the continuous development of technological tools. The European legislator should make sure that laws are up to the challenge and able to evolve with the changing game field and maybe, in the close future, regulate social media with specific measures.

[109] https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

SECTION III – A SYSTEMIC RESPONSE TO A SYSTEMIC CHALLENGE: WHAT IS NEXT FOR THE EU?

1. Introduction

After having discussed the topic of the European regulatory framework, it is time to take a look towards the future. In this last section, we will first address the problems that need to be tackled regarding the research question, referring to both external and internal interference at every stage of the electoral cycle, and seeing how these phenomena relate to social media and threaten to hinder democratic elections. Secondly, we will trace the areas of solutions in which the EU could and should act, including education and institutional strengthening policies, technological fixes, and regulatory measures, with the aim of reaching conclusive recommendations. As the problem is multifaceted, we will see how it needs to be approached from different angles and fields of study to achieve a satisfactory systematic strategy.

2. Problem areas

The concept of influencing the electorate's decision is not *per se* frowned upon; the existence of procedures that have the potential to impact citizens' preferences and the autonomy of the electorate are not mutually exclusive. On the contrary, this is precisely what campaigning and propaganda are about and the cornerstone of a pluralist democracy. Parties and politicians have a right to disseminate their ideas and programmes to the public to inform them and convince them which way to vote. However, democratic systems have regulated the relationship between the media system and the democratic process, seeking, ultimately, to guarantee transparency of influence operations and equal opportunities for political actors to participate in such processes. This is done in order to prevent these influence mechanisms from degenerating to the point of jeopardizing the possibility of forming a free

public opinion. The core of the issue is that undue influence should not be exercised, which could refer to the content of a message, but also to the manner in which said message is distributed.

Under the neoclassical paradigm of pluralist democracy, the citizen is reduced to a political consumer, a subject who is, arguably, in an even worse condition than the economic consumer: in fact, while the economic consumer enjoys a certain control over the goods and services he buys, the products that are offered to the political consumers are complex and beyond the possibility of their control, making it hard for the citizen to rationally choose the product that, on the basis of a cost-benefit analysis, turns out to be the most suitable for his preferences[110].

Election interference can occur at different levels: it can affect voter preferences, voter turnout, and the voting process. It can also manifest at different times, although it becomes more harmful leading up to the actual voting day. Due to social media, voters' capacity to make informed choices relying on fair and balanced information has been severely compromised. In this part, we will enumerate the most common dangers that social media poses to democracy and fair elections.

2.1. Polarization

Political, or ideological, polarization measures how much people agree or disagree on sensitive issues[111]. It seems that social media, by design, exacerbate political division and radicalization of political thought. This polarization is then magnified by the algorithms that will feed you 'more of the same' to keep you on the platform, in an ever-slightly increasing degree of partisan and biased content. You will end up seeing what your community and your peers are seeing, in a like-minded

[110] Gori (n 49).

[111] Ally Daskalopoulos and others, 'Thinking Outside the Bubble: Addressing Polarization and Disinformation on Social Media' *Center for Strategic and International Studies Journalism Bootcamp* (28 October 2021) <https://journalism.csis.org/thinking-outside-the-bubble-addressing-polarization-and-disinformation-on-social-media/> accessed 14 July 2022.

circle, isolating from other issues that easily become invisible and only exist in a parallel bubble[112]. The growing extremism makes it harder to reach compromises, which are essential to a functioning and healthy democracy, and it also contributes to the disruption of the public sphere. Polarization threatens social cohesion and public discourse, both fundamental in a functioning democracy, especially during the electoral cycle. Polarization is particularly insidious in countries that are already highly politically divided. Social media can be weaponized to increase polarization, which in turn can cause instability and even violence around elections[113].

One particular form of polarization, affective polarization, meaning the tendency of those belonging to one party or political belief to view people that belong to the opposite party negatively and hold biases against them[114], is regarded as particularly dangerous as it leads to hostility[115]. This process leads to citizens perceiving politics and society as an 'us' versus 'them', eroding at the mutual respect and the legitimacy of the political opponents[116].

Lastly, it is worth mentioning that the effect that social media have on polarization is still debated and uncertain. This is mainly due to the fact that it is difficult to prove who ends up in a like-minded community on these platforms because of the algorithmic design, as opposed to those who are doing so of their own will, and secondly, since researchers lack data to reach scientific conclusions.

---

[112] Ethan Zuckerman, 'Self-Segregation on Social Networks and the Implications for the Ferguson, MO Story' *EthanZuckerman.com* (27 August 2014) <https://ethanzuckerman.com/2014/08/27/self-segregation-on-social-networks-and-the-implications-for-the-ferguson-mo-story/> accessed 6 July 2022.

[113] Andrew Deck, 'Hate Speech and Disinformation Spike on TikTok in Run-up to Kenya's Elections' *Rest of World* (8 June 2022) <https://restofworld.org/2022/hate-speech-and-disinformation-spike-on-tiktok-in-kenya-election-run-up/> accessed 9 June 2022.

[114] Daskalopoulos and others (n 111).

[115] James N Druckman and others, 'Affective Polarization, Local Contexts and Public Opinion in America' (2021) 5 Nature Human Behaviour 28.

[116] Murat Somer and Jennifer McCoy, 'Déjà vu? Polarization and Endangered Democracies in the 21st Century' (2018) 62 American Behavioral Scientist 3.

2.2. Hate speech

Despite how widespread and discussed hate speech is, there is no shared legal definition of it at an international level. At the outset, it can be said that hate speech are forms of incitement to hate and/or discrimination directed against certain people due to characteristics related to their identity[117]. Hate speech has been used in politics to intimidate the 'enemy', and at the same time to win over sympathizers.

Social media have given hate speech a platform and often unaccountability, due to anonymity and their cross-national nature. The implications for election and democracy of hate speech are numerous, but mainly it represents a threat to a tolerant society as it is inherently anti-democratic. The mutual security required for peaceful contestation is undermined by hate speech, which raises concerns for electoral integrity[118]. In extreme cases, hate speech has even brought about electoral episodes of violence[119].

Another aspect of hate speech that is underlined by the criminological doctrine is the so-called silencing effect. This is a psycho-social phenomenon that brings to the inhibition of the legitimate request to participate as equal and reliable partners in the processes of democratic deliberation, made by members of the group subject to derogatory messages.[120]

As the regulation moves forward, it is important to highlight how leaving the decision about taking down potentially harmful content to social media companies does not seem to be the best choice, as

[117] Irene Spigno, *Discorsi d'odio. Modelli Costituzionali a Confronto* (Giuffrè 2018) 17 <http://www.libreriadelgiurista.it/libri/9788814228322/discorsi-dodio-modelli-costituzionali-a-confronto.html> accessed 14 July 2022.

[118] 'Safeguarding the Legitimacy of Elections: The Kofi Annan Commission Launches Final Report' (*Kofi Annan Foundation*, 23 January 2020) <https://www.kofiannanfoundation.org/supporting-democracy-and-elections-with-integrity/safeguarding-the-legitimacy-of-elections-the-kofi-annan-commission-launches-final-report/> accessed 7 June 2022.

[119] ibid; Deck (n 113).

[120] Filippo Bellagamba, 'Dalla criminalizzazione dei discorsi d'odio all'aggravante del negazionismo: nient'altro che un prodotto della legislazione penale "simbolica"?' [2018] Criminalia 28, 272.

it entails a hard balance with freedom of expression that deserves accountability and transparency[121]. It is hard to distinguish hate speech from legitimate political speech, especially since a clear definition is still lacking, and politicians are usually awarded a higher degree of freedom of expression given their position. As has been noted, "governments have yet to find a way for platforms to address this trend without interfering in electoral campaigns or creating partisan rulings about what political speech should be censored"[122].

2.3. Political advertising and micro-targeting

Political advertising shares the same issue of hate speech, as there is no common definition of it; the strategy for better regulation should start there[123]. The development of this definition is important because propaganda can be subtle and indirect, operating through dark ads[124], and covert political content. It is equally important to define the scope and purpose of such advertising, which includes those pertaining to political organizations, those that seek to influence the results of elections or referendums, those that seek to influence policies or decisions, and those that seek to influence legislative processes.

When political advertising is combined with micro-targeting, and your newsfeed, recommended content, and search results are customized based on the data previously collected on you, which is then used to create a psychographic or behavioural profile on social media, the question of undue influence becomes evident. The combination of political advertisement and targeting can be used to nudge people in one direction or another, and even to demobilize a section of the electorate to keep them from voting altogether.

[121] That also represented a problem with the code of conduct that social media adopted with the support of the EU.

[122] 'Safeguarding the Legitimacy of Elections: The Kofi Annan Commission Launches Final Report' (n 118).

[123] If the proposal of the European Commission on regulation is approved, it will give a new, very broad definition of political advertising, as we have seen in section 2.

[124] Ads that are only visible to the targeted group.

Some social media platforms, like Twitter, have banned political advertising altogether, while institutions have tried to put restrictions on the use of personal data for targeting, like the GDPR, although more systematic and strong measures such as a total ban on the use of targeting or on dark ads have yet to be implemented. Political advertising is at the moment subject to the same rules and regulations as other forms of advertising on a European level, but it is important for it to be always identified as such and should never be mistaken for news or editorial content. As we have seen in the second section, potential EU regulation on political advertising at large is forthcoming[125].

There is also a need for a systematic and comprehensive, user-friendly catalogue with open access containing archived information on ads run on social media with political aims. This archive should indicate the amount of money spent on political advertising and by whom, as well as targeting criteria, to stop widespread direct foreign interference and to make it clear who is behind the message. This would also prove to be a strong obstacle to running 'parallel campaigns'[126].

2.4. Foreign interference

Foreign interference operations are carried out by both state and non-state actors with political, economic, or military aims. There is widespread consensus that Russian interference, for instance, is aimed at undermining confidence in democratic institutions and processes at large[127].

Foreign interference in internal politics is not a novelty by any means; what is new is the insidiousness that can be achieved through the use of social media with these ends, through bots, disinformation, hate speech, or political advertising. As has been said, "by exploiting the open, anonymous, and

[125] Proposal for a regulation of the European Parliament and of the Council on the transparency and targeting of political advertising COM/2021/731 final.

[126] By parallel campaigns we mean the activity of political actors that target different audiences with different content and messages in order to personalize and convince them.

[127] Tim Maurer and Erik Brattberg, 'Russian Election Interference: Europe's Counter to Fake News and Cyber Attacks' [2018] Carnegie Endowment for International Peace <https://carnegieendowment.org/2018/05/23/russian-election-interference-europe-s-counter-to-fake-news-and-cyber-attacks-pub-76435> accessed 13 July 2022.

borderless nature of digital technologies, social media have provided novel opportunities for bad actors to meddle transnationally"[128].

National sovereignty on elections is recognized and guaranteed by international conventions; these malign foreign actors should not be allowed to influence the outcome of voting operations[129]. Other than the actual influence that these campaigns have on votes, they also add on to the lack of trust that citizens have in the electoral integrity and democratic processes. The opportunity for an international convention on this topic has been underlined[130].

One of the problems of foreign interference is that it is hard to distinguish it from legitimate campaign activity; this is especially due to the professionalization of election manipulation. A number of actors are involved, with dubious ties and relationships, hard to detangle and regulate: commercial entities, political consultants, national groups, communication firms, and foreign government[131].

2.5. Disinformation

Disinformation is false or misleading information that is shared in an attempt to hurt or profit from others[132], whereas misinformation is inaccurate information that is disseminated with no malicious intent[133], this distinction is not always clear in practice as it is based on the subjective intention of the author.

[128] 'Safeguarding the Legitimacy of Elections: The Kofi Annan Commission Launches Final Report' (n 118).

[129] Article 2 of the United Nations Charter guarantees the territorial and political integrity of states.

[130] 'Safeguarding the Legitimacy of Elections: The Kofi Annan Commission Launches Final Report' (n 118).

[131] ibid.

[132] The harm intended may be directed toward specific citizens, groups, institutions, or processes. Claire Wardle and Hossein Derakhshan, 'Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making' (Council of Europe 2017) <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html> accessed 14 July 2022.

[133] ibid.

The insidiousness of disinformation, with regards to elections and democracy, is in the fragmentation of the public arena, as it leads to a lack of agreement on truth, facts, and realities, which is fundamental for political debate. The effects of these phenomena are heightened in periods leading up to elections and voting, when informed choices by citizens are essential. It also further fuels the already serious crisis of trust in media, contaminating public discourse and acting as a pollutant in the information ecosystem[134]. False information is a challenge to democratic governance, which depends on citizens having access to reliable information for making decisions[135]. Elections are based on informed decisions made by voters; a polluted information environment can push citizens in the wrong direction, leading them to vote against their own preferences. It can also undermine the public's trust in free and fair elections "by sowing doubt about the integrity of the ballot box and the professional, impartial behavior by election management bodies (EMBs), and spreading rumors that call into question the legitimacy and accuracy of an election."[136]

Disinformation clashes with the right to participate in public affairs "without undue influence or coercion of any kind which may distort or inhibit the free expression of the elector's will," which also entails that "voters should be able to form opinions independently, free of violence or threat of violence, compulsion, inducement or manipulative interference of any kind."[137]

One of the problems with disinformation is that it comes in uncountable forms[138]. From satire and parody to fabricated content designed to deceive and harm, not all types of disinformation obviously

[134] ibid.

[135] In general, incorrect information in the news tends to attract readers more than real information. A 2018 study by the MIT Media Lab on the circulation of news on the Twitter platform, found that false information disseminates "significantly farther, faster, deeper, and more broadly than the truth in all categories of information."

[136] 'Safeguarding the Legitimacy of Elections: The Kofi Annan Commission Launches Final Report' (n 118).

[137] UN Human Rights Committee (HRC), CCPR General Comment No. 25: Article 25 (Participation in Public Affairs and the Right to Vote), The Right to Participate in Public Affairs, Voting Rights and the Right of Equal Access to Public Service, 12 July 1996, CCPR/C/21/Rev.1/Add.7.

[138] For an accurate and detailed distinction of the different kinds of disinformation see Aimee Rinehart, 'Fake News. It's Complicated.' [2017] *First Draft* <https://firstdraftnews.org:443/articles/fake-news-complicated/> accessed 16 July 2022.

have the same consequences and potential to harm democratic discourse[139]. Arguably the most dangerous is not the blatantly false information but the more subtle, misleading content, which is hard to debunk or rebut, but mostly detect. For this reason, it is hard to treat it as a single phenomenon. In fact, all the topics discussed so far are somehow related to disinformation: polarization is also intensified by echo-chambers in which non-verified information is spread but contained and, therefore, not rebutted, some fake news have been categorized as hate speech, disinformation is ever so efficient when combined with micro-targeting, and foreign interference has often been exercised through the dissemination of misleading or false information.

Despite disinformation having been around for longer than anyone can remember, it seems to be regarded as a key issue in recent years; this is mostly due to the fact that social media functioned as an accelerant for it.

The authors and actors involved in misinformation can be public or private, and often utilizes bots and artificial intelligence to amplify their reach. It could come as a surprise that disinformation connected to elections and other topical political phases is sometimes done for profit and not necessarily with political ends. So-called farms have been discovered creating astounding quantities of fake news content to generate traffic and consequently revenue through ads[140]. These misleading but attention-grabbing articles help maximize ad-click revenue by keeping users on the platform. Even more surprisingly, these farms are often not even located in the country about which they write, and don't employ people who speak the same language of the content produced[141].

---

[139] On the most recent wave of disinformation connected to elections on social media see Deck (n 113); Tim Culpan, 'TikTok Is the New Front in Election Misinformation' *Washington Post* (29 June 2022) <https://www.washingtonpost.com/business/tiktok-is-thenew-front-in-election-misinformation/2022/06/28/0aedd53e-f73b-11ec-81db-ac07a394a86b_story.html> accessed 6 July 2022; Alistair Coleman, 'French Election: Misinformation Targets Candidates and Voting System' *BBC News* (22 April 2022) <https://www.bbc.com/news/61179620> accessed 6 July 2022.

[140] Karen Hao, 'How Facebook and Google Fund Global Misinformation' [2021] MIT Technology Review <https://www.technologyreview.com/2021/11/20/1039076/facebook-google-disinformation-clickbait/> accessed 17 May 2022.

[141] One of Meta's initiatives has been to require that political advertisers prove to live in the country that they are targeting.

Lastly, it has unfortunately been proven that once exposed to disinformation sources, it is hard to counter its consequences due to backfire effect; for this reason a better solution needs to be found to counter it at its source.

3. Possible ways forward: a multidimensional approach

To work towards a vote free from undue influence and to affect electoral manipulation at every stage, different interventions in various phases are needed, namely: in the production of information (journalists and politicians), on its distribution (social media), up to its consumption and understanding (voters/citizens/consumers). Every potential answer brings some difficulties and limits with it, as they all have consequences for other important human rights, such as freedom of speech or the right to privacy. The following part presents a critical analysis of different types of solutions that have been proposed by legal scholars, technology experts, and the social science academia. This examination will be carried out in order to reach recommendations on future actions, to harness the opportunities that these technologies offer while mitigating the harm that they might cause.

Possible policies measures by the EU to address these challenges should be targeted at voters, social media platforms, and other actors. A long-term, holistic approach to developing resilience unquestionably calls for comprehensive plans and tools, but for those to be successful, they must be guided by regular cross-country sharing of best practices and lessons learned[142].

[142] Richard Youngs and others, 'European Democracy Support Annual Review 2021' [2021] Carnegie Endowment for International Peace 58.

3.1. Education and awareness raising

In this chaotic informational landscape, it is paramount to have an electorate with acute capabilities in critical thinking, research skills, and discerning capabilities. This would help the general population to detect manipulation, disregard hate speech, and discredit unreliable sources of information.

Fact-checking operations have shown some limits, mainly due to the fact that they can backfire and end up bringing more visibility to false information, providing it with a larger audience[143]. Secondly, when carried on by people it is a costly and lengthy procedure, and does not extend to private message social platforms such as Whatsapp and Telegram. Furthermore, it has been demonstrated that providing the correct facts once someone has been exposed to disinformation does not undo the damage[144].

While fact-checking and other strategies focus on situations in which the harm has already been done, digital media literacy education aims at preventing it from occurring. A well-educated and informed population contributes to a stronger democracy, it helps to build resilience. The idea is to construct a cognitive firewall, or cognitive resilience, in order to instil a healthy scepticism for misinformation[145]. This is even more urgent at a time in which advanced technologies such as deepfakes are becoming more common and therefore dangerous. Voter education is essential to boosting societal resilience. For instance, the government of Sweden started a national campaign to educate high school students about Russian propaganda[146].

Also crucial is the public disclosure by government and intelligence officials of pertinent information on cyber activities targeting democratic institutions. Governments and the EU should engage in

[143] Craig Silverman, 'The Backfire Effect' [2011] Columbia Journalism Review <https://www.cjr.org/behind_the_news/the_backfire_effect.php> accessed 13 July 2022.

[144] Elizabeth Kolbert, 'Why Facts Don't Change Our Minds' [2017] *The New Yorker* <https://www.newyorker.com/magazine/2017/02/27/why-facts-dont-change-our-minds> accessed 13 July 2022.

[145] Flemming Splidsboel Hansen, 'Russian Hybrid Warfare: A Study of Disinformation' (Danish Institute for International Studies 2017).

[146] Maurer and Brattberg (n 127).

awareness raising campaigns, both general and specific to elections or times of heightened political tensions, about the means and motivations behind these political tactics. Additionally, the data should, whenever possible, be supported by forensic proof to raise public awareness and aid in future planning of action.

Most didactic initiatives focus on younger people, both for reasons of opportunity, as it is easier to insert digital literacy in the traditional education course, and because younger generations are more likely to spend time online and defer to social media to look for information[147]. Nonetheless, it is the author's opinion that it is crucial to focus more on the older section of the electorate for two reasons. Firstly, people that grew up when social media were already present know how to better navigate it, compared to non-digital natives[148]. Secondly, because this constitutes the generation of parents, teachers, and leaders, and we cannot skip this demographic in the training process; it is essential that the whole of society has a better understanding of these phenomena and their consequences. Nevertheless, digital literacy should become a compulsory part of the education curriculum in the entire European Union[149].

Initiatives to provide the general population with digital verification skills and to empower it against manipulation in its many forms should be put in place, encouraged, and supported by the EU. The public should be educated on how to find resources and how to evaluate their content without accepting them at face value[150]. Media literacy is also about recognizing both cognitive and

[147] Katerina Eva Matsa and others, 'Younger Europeans Are Far More Likely to Get News from Social Media' (*Pew Research Center's Journalism Project*, 30 October 2018) <https://www.pewresearch.org/journalism/2018/10/30/younger-europeans-are-far-more-likely-to-get-news-from-social-media/> accessed 15 May 2022.

[148] According to some studies, people over the age of 65 are up to seven times more likely to share disinformation than people age 18-29 Alex Hern, 'Older People More Likely to Share Fake News on Facebook, Study Finds' *The Guardian* (10 January 2019) <https://www.theguardian.com/technology/2019/jan/10/older-people-more-likely-to-share-fake-news-on-facebook> accessed 11 July 2022.

[149] This could also help with other issues caused by social media, such as mental health problems.

[150] The Center for Media Literacy produced a document containing 'Five Key Questions of Media Literacy': Who created this message? What creative techniques are used to attract my attention? How might different people understand this message differently than me? What values, lifestyles and points of view are represented in, or omitted from, this message?

confirmation biases, our own or the authors', to critically think about the information presented to us and push ourselves to look up different versions of the story, perhaps explained from a different perspective[151].

3.2. Policies to strengthen the institutions

Weak institutions and public distrust towards the leading class make democracy vulnerable to foreign interference and a number of internal threats. This is why strategies to tackle phenomena such as hate speech and political manipulation cannot prescind from taking the broader picture into account.

A lot of the challenges that social media bring to democracy are connected to the concomitant crisis of traditional media. Public trust in journalism is low and declining[152], and in addition to this, the business model has changed: to make a profit news organization need to generate more content with the highest chances of engagement. This has led to an increase in quantity and a decrease in quality. It is important that the European Union supports professional journalism, encouraging media organizations to reinforce and update already existing journalistic quality and ethical standards and elaborate principles for the coverage of the election period news to protect against disinformation and voters' manipulation. As expressed in a recommendation by the Committee of Ministers of the Council of Europe: "these standards should reflect their particular roles and responsibilities in democratic processes"[153]. Some authors contend that reviving professional standards in the mainstream media, which can then verify the claims made in the digital sphere, is one solution to the

Why is this message being sent? As most media messages are organized to gain profit and/or power. The document can be found at https://www.medialit.org/sites/default/files/14B_CCKQPoster+5essays.pdf.

[151] Belinha De Abreu, 'Gatekeeping Misinformation with Media Literacy Education.' (2021) 50 Knowledge Quest 26.

[152] Rasmus Kleis Nielsen and Richard Fletcher, 'Democratic Creative Destruction? The Effect of a Changing Media Landscape on Democracy' in Joshua A Tucker and Nathaniel Persily (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press 2020) <https://www.cambridge.org/core/books/social-media-and-democracy/democratic-creative-destruction-the-effect-of-a-changing-media-landscape-on-democracy/8C6548E16FA63289FC4C731AC512B075> accessed 14 July 2022.

[153] Recommendation CM/Rec(2007)15 of the Committee of Ministers to member states on measures concerning media coverage of election campaigns.

current issues of polarization and misinformation[154]. Working with trustworthy legacy media to develop a counter-narrative to election-related misinformation is another potential area for intervention.

It is necessary to establish a communication channel between government and the media. Government representatives and media outlets interacting actively helps to guard against the intentional spread of false information. For instance, nations may want to create a permanent media council like the one in Sweden, which frequently brings together members from the government and the media[155].

The European Union could also fund and encourage the institution of fact-checking organizations dedicated to debunking fake news, possibly through local civil society organizations that are able to understand the language and the context of the particular situation.

A code of conduct, or simply an agreement between political parties and other actors to behave in a certain way during elections and refrain from unfair and unethical practices, should also be pursued, as it seems to have given good results in other realities[156]. For instance, in the future, political actors could agree to not use deepfakes and chatbots, or other innovative nudging techniques[157].

One recurring recommendation, and its importance cannot be overstated, is for authorities to require social media platforms to provide more data on their practices. The European Union should increase the obligations of these companies to release access to data to independent researchers in order for them to study specific areas. For instance, it would allow to audit algorithms for bias towards

---

[154] Yochai Benkler, Robert Farris and Hal Roberts, *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (Oxford University Press 2018) <https://library.oapen.org/handle/20.500.12657/28351> accessed 15 June 2022; Jacob Rowbottom, 'Campaign Finance and Electoral Speech in the Media' in James A Gardner (ed), *Comparative Election Law* (Edward Elgar Publishing 2022) <https://www.elgaronline.com/view/edcoll/9781788119016/9781788119016.xml>.

[155] Maurer and Brattberg (n 127).

[156] Nigeria's Abuja Accord on Election Conduct was first adopted in 2015, with the collaboration of Kofi Annan, and it has been renewed ever since 'Protecting Electoral Integrity In The Digital Age' (Kofi Annan Commission on Elections and Democracy in the Digital Age 2020).

[157] This is what happened in Germany in 2017, when according to reports, political parties reached a 'gentlemen's agreement' not to use leaked material for political gain. Additionally, they promised not to deploy social media bots. Youngs and others (n 142).

extremism, comprehend the impact of social media on political polarization, and investigate the links between online hate speech and offline violence.[158]. As has been said, "although the platforms themselves generate new sources of real-time transactional data that might be used to understand this changed environment, most of this data is proprietary and inaccessible to researchers, meaning that the revolution in big data and data science has passed by democracy research"[159]. Disclosure of more data would bring not only scientific studies on problems and solutions but also stronger oversight of social media's business operations.

The European Union should sponsor and support steps towards international cooperation in the sector, with the aim of sharing information and good practices. This coordination is all the more important given the cross-national nature of social media.

One innovative solution that has been suggested is the establishment of an internet ombudsman at the European level. A motion for a recommendation was put forward in the Parliamentary Assembly of the Council of Europe, and a resolution was adopted in 2020[160]. This ombudsman institution would be responsible for the hard decisions that need to be made when restricting freedom of expression. It would also facilitate the job of social media companies that could defer to it in difficult cases and, therefore, avoid responsibility for making the wrong choice. As stated in the mentioned resolution: "in order to avoid freedom of expression being limited in a discriminatory manner while at the same time making efforts to fight against illegal content on the internet, the Parliamentary Assembly is proposing that consideration be given to establishing an ombudsman institution (or equivalent) with the requisite independence, powers and authority to assess whether internet content is legal or illegal. Internet intermediaries could submit questionable cases to the institution for its recommendations on

[158] The DSA contains such provisions demanding data access be provided to 'vetted researchers'.

[159] Margetts (n 21).

[160] Parliamentary Assembly of the Council of Europe, Resolution 2334 (2020), Towards an internet ombudsman institution.

how to deal with them."[161] The main problem with this idea is the plausible length of the procedure, in fact a response to instances of foreign interference, hate speech, or disinformation, could become pointless by the time it's addressed. Another possible obstacle could be the volume of requests that could bring the system to a standstill.

3.3. Technological fixes

Even though it's not possible to agree with those that believe that a technological problem can be easily solved with a technological solution[162], we also need to recognize that technology will be involved in solving some of the aforementioned issues.

Artificial intelligence is being used to detect hate speech, identify fake news, discover bots, and delete political manipulation. There are many reasons why this solution has sometimes been deemed inappropriate. First, the way AI functions is inscrutable even to its own designers. AIs are black boxes when it comes to their reasoning, therefore, when they are being used to censor speech, this turns out to be highly problematic as it is hard to contest the motivation behind these decisions. Furthermore, AIs have historically struggled with language content, as opposed to quantitative information. These instruments are, at the moment, unable to interpret discourse in context the way a person does, they lose nuances in speech, irony, and much more[163]. Nonetheless, this seems to be the road that platforms such as Twitter and Facebook have chosen in the last years due to its inexpensiveness, particularly

[161] Ibid.

[162] Especially since some would argue that it is not a technological problem at all, just a societal one that has been exacerbated by social media.

[163] Recent studies have revealed that these machines are still unable to understand context or discern the speaker's motivation or intent, failing to recognize particular elements as socially useful. In this study the researchers focused on the communication style embraced by drag queens, analysed with Perspective, an algorithm used to identify the level of speech toxicity in order to inform decisions on content removal Alessandra Gomes, Denys Antonialli and Thiago Dias Oliva, 'Drag Queens and Artificial Intelligence: Should Computers Decide What Is "Toxic" on the Internet?' [2019] *InternetLab* <https://internetlab.org.br/en/news/drag-queens-and-artificial-intelligence-should-computers-decide-what-is-toxic-on-the-internet/> accessed 11 July 2022.

for fact checking[164]. One of the main strategies for this operation focuses on finding basic statistical claims and evaluating their accuracy against a structured database. These techniques are becoming more accurate as a result of the enormous datasets.

Another way to deal with the problems that were previously illustrated is, instead of looking at the content, to focus on the way it gets disseminated. This means to take a look at the way in which algorithms prioritize, highlight, and amplify certain messages over others. Any attempt to control the influence of these intermediaries in a media system where platforms with algorithmic operations control the flow of information, requires an understanding of how these algorithms take advantage of the processing of user input to model the platform's architecture[165]. Algorithm changes could be used to demonetize ads and bot-like activity, depersonalize the feed, and to provide alternative sources of information to users.

In an internal Meta report, Allen suggested adopting what is known as a graph-based authority measure to rank content as one strategy to combat misinformation, instead of prioritizing engagement. This would reverse the present trend by amplifying higher-quality pages like news and media outlets and undermining lower-quality pages like clickbait and trolls[166]. In order to degrade bad actors in its search rankings, this technology evaluates a web page's quality based on how frequently it references and is cited by other quality web pages.

Extensions and apps have also been created to label content, identifying it as being viral, or flagging it as not credible[167], or to tweak one's newsfeed on social media. For instance, following the 2016 US presidential election, Zuckerman and his co-workers created a tool called Gobo that allows users

[164] Working with tools such as FullFact.

[165] Gori (n 49).

[166] Karen Hao, 'Troll Farms Reached 140 Million Americans a Month on Facebook before 2020 Election, Internal Report Shows' [2021] MIT Technology Review <https://www.technologyreview.com/2021/09/16/1035851/facebook-troll-farms-report-us-2020-election/> accessed 13 July 2022.

[167] In general, it seems like the better solution is to flag content instead of taking it down.

to customize their own bubbles using sliders that manage content filters[168]. The 'politics' slider, for instance, has options ranging from 'my perspective' to 'lots of perspectives'. People who choose the latter are exposed to media sources they might not typically view. Such apps, however, demand that users of the news take the proactive steps of acknowledging a partisan bias in their news intake and choosing to take action to address it[169].

Technological solutions cannot solve these problems by themselves, but they can be used to aid human resources in their tasks by bringing content to their attention due to its manipulative, toxic, or false nature, while leaving the ultimate decisions in the hands of a person who is better equipped at making choices that relate to such essential human rights.

3.4. Anti-trust law and business model regulation

All the alternatives that we have seen so far belong to what can be called a reconfiguration pattern of dealing with social media, meaning to adapt the laws and guidelines created for the offline world to the online realm[170]. Another option is, instead, to aim for a structural reform of social media as they are conceived. On this point, Bennett contends that the focus of the current platform regulation discussion is too narrowly focused on individual issues, such as hate speech and disinformation, instead of trying to address the root causes, namely the business model, or 'techno-capitalism'[171].

While social media companies' representatives have spoken in the past about the opportunity for legislative intervention that would take liability away from them[172], they have always been contrary

[168] Piore (n 22).

[169] Other initiatives of this kind include InVid, which was developed with EU funding, Social Mirror, and Flip Feed.

[170] Wischmeyer (n 20).

[171] W Lance Bennett, 'Killing the Golden Goose? A Framework for Regulating Disruptive Technologies' [2021] Information, Communication & Society 1.

[172] Mark Zuckerberg, 'The Internet Needs New Rules. Let's Start in These Four Areas.' *Washington Post* (30 March 2019) <https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html> accessed 16 July 2022.

to the establishment of competition rules that challenge their market dominance. Market power concentration combined with media power concentration can supress innovation and reduce consumers' choice, for this reason the US House Subcommittee on Antitrust has underlined the opportunity to tackle the monopoly powers of Meta, Alphabet, Amazon, and Apple[173]. It is generally accepted that dismantling tech monopolies could be a crucial step in fostering alternatives that would enable future advancements and more user choice[174]. According to this perspective, as users switch to other platforms, the market will ideally resolve dominance issues driving the established platforms to change their business practices. What this point of view does not take into consideration is the peculiarity of social media platforms. Firstly, these operations do not directly address free speech or privacy concerns; secondly, these networks are most successful when they are able to corner the market, as they will have more users and larger data sets[175]. Basically, social media are designed to work at their best as monopolies, and increased competition would not change the economy's growing reliance on personal data or alleviate the ethical issue of platforms' covert behavioural manipulation.

The targeted advertising-based business model that is ingrained in social media and other digital technology, according to critics of the industry[176], is the root of the issue. As users' data can be sold for profit to target advertising or sell services to them, the majority of profit-making technology companies seek to generate income by attracting and gathering information about their active and

[173] Lauren Feiner, 'House Democrats Say Facebook, Amazon, Alphabet, Apple Enjoy "monopoly Power" and Recommend Big Changes' *CNBC* (6 October 2020) <https://www.cnbc.com/2020/10/06/house-democrats-say-facebook-amazon-alphabet-apple-enjoy-monopoly-power.html> accessed 13 July 2022.

[174] Damian Tambini, 'Social Media Power and Election Legitimacy' in Damian Tambini and Martin Moore (eds) (Oxford University Press 2018) <https://global.oup.com/academic/product/digital-dominance-9780190845124?cc=gb&lang=en&> accessed 12 July 2022.

[175] Major businesses in the technology sector appear to be trying to persuade regulators not to break up big tech by arguing that competition regulation will actually hurt users who profit from the status quo and enjoy free services and the advantages of network effects. Stockmann (n 66).

[176] Shoshana Zuboff, *The Age of Surveillance Capitalism* (PublicAffairs 2019) <https://www.publicaffairsbooks.com/titles/shoshana-zuboff/the-age-of-surveillance-capitalism/9781610395694/> accessed 13 July 2022; Ethan Zuckerman, 'The Internet's Original Sin' *The Atlantic* (14 August 2014) <https://www.theatlantic.com/technology/archive/2014/08/advertising-is-the-internets-original-sin/376041/> accessed 16 July 2022.

growing user base[177], as it directly translates into commercial income for the company. Zuboff refers to this economic model as 'surveillance capitalism', in contrast to conventional advertising utilized in traditional media, highlighting its manipulative elements that allow for influence over individuals' political and consumer behaviour[178].

A systematic overview of alternative funding and management options for IT infrastructure is provided by Landwehr, Borning, and Wulf[179]. These options include alternative business models for for-profit corporations, public funding programs, non-governmental organizations, cooperatives, and models with no or minimal funding. One, perhaps radical, alternative is to regulate social media as public utilities. Zuckerman has proposed the idea of a tax-payer funded platform[180] built to perform a public mission to deliver a "diverse and global view of the world."[181] This solution brings with it the problems that State-owned media always do, among them censorship and narrative control. Another option to influence the business model is to offer a subscription option; by paying a price, people could refuse to sell their own data and perhaps tweak their feed to avoid personalization and ads altogether. This would build an opt-out solution, unfortunately limited to those that can afford to pay this price and have the awareness to do so; therefore, this mechanism would probably result in protecting individuals that already have strong critical thinking skills and discerning capabilities.

[177] 'They make their money by keeping people on their sites and apps; that aligns their incentives closely with those who stoke outrage, spread misinformation, and appeal to people's existing biases and preferences.' Tufekci (n 3).

[178] Zuboff (n 176).

[179] Marvin Landwehr, Alan Borning and Volker Wulf, 'Problems with Surveillance Capitalism and Possible Alternatives for IT Infrastructure' [2021] Information, Communication & Society 1.

[180] There have been countries, like Papua New Guinea and Egypt, that have gone in this direction. David Martin, 'Papua New Guinea to Ban Facebook for One Month, Broaches State-Run Social Network | DW | 29.05.2018' *DW.COM* (29 May 2018) <https://www.dw.com/en/papua-new-guinea-to-ban-facebook-for-one-month-broaches-state-run-social-network/a-43982064> accessed 13 July 2022; Mildred Europa Taylor, 'Egyptian Government Launches New Social Media Platform to Rival Facebook' *Face2Face Africa* (26 March 2018) <https://face2faceafrica.com/article/egyptian-government-launches-new-social-media-platform-rival-facebook> accessed 13 July 2022.

[181] Piore (n 22).

3.5. More regulation

Social media's worldwide reach and private nature undermined media governance frameworks, which used to rely predominantly on the state as the primary actor. This event generated a shift away from national regulation and toward supranational and international regulation, as well as a trend toward self-regulation as the industry took a more active role in regulation[182]. For what pertains to social media, the trend has so far been to leave these platforms to regulate themselves through voluntary measures. This has been justified with the idea that these companies are in a better position to address these challenges, being the best informed. Unfortunately, social media are not neutral, and they don't have incentives to remove content that generates traffic, which hate speech and disinformation may represent. At the same time, it has been noted that to leave these decisions on essential rights, such as freedom of speech and participation in public affairs, in the hands of private economically driven actors is not ideal[183].

The opposite solution, hard and unidirectional regulation by the institutions, also presents some downsides as norms that restrict freedom of speech always do. Legislation on the topic of hate speech and disinformation has often been criticised and labelled as censoring[184]. The best approach seems to be co-regulation, where norms are created as the result of a dialogue between businesses and authorities, and the latter are able to monitor implementation and intervene in cases of inaction. Given the speed of technological advancements opposed to the length of legislative processes, co-regulation seems to be the most apt choice for the opportunity of quick policy upgrades. Co-regulation allows for a multilateral consultation with different stakeholders and expertise; it is fundamental that any policy in this area is created through an inclusive process in which social media companies, civil society organizations, and traditional media representatives are heard from. This type of regulation

[182] Stockmann (n 66).

[183] The main issue is that these platforms tend to overcompensate in the removal of content to avoid fines and responsibility.

[184] For instance, Germany NetzDG law has been heavily criticised Wischmeyer (n 20).

also has the advantage of shifting expenses for standard-setting and monitoring on businesses while also encouraging the industry to voluntarily follow public interest[185]. The EU has to create incentives for social media businesses to take a proactive role in mitigating potential hazards by detecting deception tactics, sharing knowledge, and taking action to identify and remove fraudulent accounts.

It is also recommendable to establish a silence period of at least 24 hours on social media preceding important voting processes, in the same way it is set up in certain European countries for legacy media. These countries prohibit political advertising, and sometimes opinion polls or exit polls for a timeframe that ranges from 24 to 48 hours before voting day[186].

Additionally, it has been suggested that regulation activity could intervene on implicit personalization, forcing platforms to design the algorithms to ensure that users interact with multiple sources of information and with content of a different orientation than that produced by explicit personalization, which could go to tackle disinformation, micro-targeting, and foreign interference[187]. Ads can be attached to content, instead of people, and the feed could be algorithmically diversified. In addition to the already discussed open-access catalogue of political advertising, authorities could enforce a ban on dark ads or restrict even further the use of personal data in targeting.

Finally, we have to stress that social media platforms' responsibility for content shared by third parties needs to be redefined, given the role that they have come to play in recent years. Although these companies probably should not be equated to traditional media that have strong editorial power, such as newspapers and TV channels, they also cannot continue to be defined as simple intermediaries[188],

[185] Stockmann (n 66).

[186] In Mexico a political party was found in breach of electoral silence for posting tweets through intermediaries during the silence period. European Commission for Democracy Through Law (Venice Commission), 'Study on the role of social media and the Internet in democratic development', [21 novembre 2018], (CDL-LA(2018)001).

[187] Gori (n 49).

[188] This concept is strongly expressed by Eisenstat, stating: 'To still categorize social media companies — who curate content, whose algorithms decide what speech to amplify, who nudge users towards the content that will keep them engaged, who connect users to hate groups, who recommend conspiracy theorists — as "internet intermediaries" who should enjoy immunity from the consequences of all this is beyond absurd. The notion that the few tech companies who steer how more than 2 billion people communicate, find information, and consume media enjoy the same blanket

particularly since they do exercise monitoring and moderation of content to an extent[189]. An *ad hoc* category between mere conduit and publisher should be created to discipline the role that these platforms have ended up performing[190]. As has been said, "the privileges of social media are increasingly difficult to justify in light of their political impact and of the ongoing process of media convergence"[191]. This demand for liability could be met by requiring more transparency, imposing reporting or disclosing duties[192], or demanding due diligence with specific focus on social media's effect on democracy and elections[193], taking notes from regulation on other specific issues in the sector of Human Rights and Businesses[194]. Another option is to demand oversight or auditing on topical features such as recommendation engines, targeting tools, and algorithmic amplification[195].

All things considered, it's the author's opinion that there is space for the three forms of regulation; the more technical matters should be left to social media companies to decide on, while the most topical and delicate matters, such as election interference and democratic discourse, have to be subject to co-regulation while authorities, and namely the EU, should reserve the right to intervene with harder discipline in the case on particularly delicate topics, or in the occurrence of the ineffectiveness of these platforms' actions.

immunity as a truly neutral internet company makes it clear that it is time for an upgrade to the rules. They are not just a neutral intermediary.' Eisenstat (n 101).

[189] They have been described as proxy censors, social media block, filter and delete content under certain circumstances Stockmann (n 66).

[190] Rowbottom (n 154).

[191] Wischmeyer (n 20).

[192] Following Germany's example.

[193] Companies should be required by law to assess the possibility of disguised propaganda efforts in any country in which they have business operations. Johan Farkas, 'Disguised Propaganda on Social Media: Addressing Democratic Dangers and Solutions' (2018) XXV The Brown Journal of World Affairs <https://bjwa.brown.edu/25-1/disguised-propaganda-on-social-media-addressing-democratic-dangers-and-solutions/> accessed 13 July 2022.

[194] Social media companies could have specific obligations of due diligence efforts on democratic interference the same way these requirements have been implemented regarding children's rights.

[195] Eisenstat (n 101).

4. Conclusions

In this section, we have first seen the main challenges to democracy brought about or exacerbated by social media. Polarization, hate speech, targeted political advertising, foreign interference, and disinformation are not a novelty, they have existed long before Twitter or Facebook, but the way they evolved with technology and social networks have made them even more insidious for the public sphere and democratic discourse.

They represent a multifaceted problem that requires a multifaceted solution. Each of the possible paths we have seen has its merits and limits, for this reason the only possible way forward is to embrace a multidisciplinary and holistic approach to the situation.

As we have seen in this section, while digital literacy and policies to strengthen the institutions are the best long-term strategy, technology can aid us in identifying the problems and adjusting the algorithms. More drastic solutions, such as implementing anti-trust norms that would transform social media's business model, have also been touched on. Finally, the European Union needs to continue, and in fact finally is, moving away from self-regulation to a more co-regulatory approach.

We should note that certain topics, such as echo chambers and bots, and new technologies that will undoubtedly prove to be disruptive, like deepfakes and chatbots, deserve a deeper understanding and more research going forward.

As we have seen, there is no quick fix to the threats that social media pose to democratic discourse and fair and free elections. Going forward, strategies should be informed by data disclosed by social media companies, and results need to be shared on an international level to allow cooperation and shape best practices. Since research on actual and potential effects of social media on democratic processes is still in its infancy, solutions may evolve based on empirical evidence derived from future experiences.

CONCLUSION

The objective of this research has been to determine whether and how social media negatively affect the public political discourse and to develop policy recommendations for the European Union to avoid undue interference in the formation of personal political opinions concerning elections and voting processes.

As seen in the first section, social media are closely linked to public political discourse and the right to participate in public affairs. This connection can prove to be both an opportunity and a challenge. Many phenomena connected to social media put the democratic process of elections and the individual right to form an opinion without undue influence at risk. The distinctive features of these platforms further accelerate these negative consequences. After illustrating the principles behind the regulation of political speech in legacy media, it became clear that a simple expansion of the scope of application of existing rules is insufficient to adequately control public discourse on social media and specific measures are necessary.

The EU framework on the topic, discussed in the second section, is the result of a regulatory patchwork comprised of numerous initiatives with various degrees of enforceability and aimed at different sectorial problems. As declared by the European Commission itself, there are still significant gaps and legal burdens to address, despite a range of targeted and sector-specific interventions. While some of these voids will be filled by the upcoming legislation (DSA, DMA, and more), the discipline may remain fragmented and less than coherent. The spread of new technologies disrupted electoral laws and political speech regulation; the European legislator should ensure that these laws are evolving along with the continuous development of technological tools and should consider, in the near future, regulating social media with specific measures.

In the last section, we addressed the main challenges to democracy brought on or exacerbated by social media. Polarization, hate speech, targeted political advertising, foreign interference, and

disinformation are not a novelty to our societies; nevertheless, technology and social networks have made them even more insidious for the public sphere and democratic discourse. These phenomena represent a complex and multi-layered problem that requires an equally sophisticated answer. As each possible remedy has its merits and limits, the only way forward is to address these challenges from a multidisciplinary and holistic standpoint. While digital literacy and institutional strengthening policies are the best long-term strategy, technology can have a role in aiding with identifying the problems and reshaping the algorithms. Other, more drastic, solutions that aim at transforming social media's business model, such as anti-trust norms and the reconfiguration of these platforms as public utilities, are also available. Finally, we have underlined the appropriateness for the EU to leave the era of social media platforms' self-regulation behind and move towards a co-regulatory approach while reserving itself the right to legislate on the most salient and delicate situations.

Certain topics deserve to be objects of more research, among them echo chambers, bots, and new technologies that will undoubtedly prove to be disruptive, such as deepfakes and realistic chatbots.

As we have seen, there is no quick fix to social media's threats to democratic discourse and fair and free elections. Going forward, EU strategies should be rooted and derived from data disclosed by social media companies, and the outcomes of these approaches need to be shared on an international level to foster cooperation and establish best practices. Since research on social media's actual and potential effects on democratic processes is still in its infancy, policy tactics may change based on empirical evidence derived from future experiences.

The interests at stake are essential to our democracies, and the repercussions of inaction or policy mistakes can be severe. This fact should not refrain us from using social media to our advantage and from harnessing the benefits they offer; it should only mean that there is a need to stay vigilant and that, as technology continues in its inevitable development, laws and policies need to evolve

accordingly at the same rate of speed and sophistication. To conclude, in the words of Kofi Annan: "technology does not stand still; neither can democracy"[196].

[196] 'Safeguarding the Legitimacy of Elections: The Kofi Annan Commission Launches Final Report' (n 118) 1.

BIBLIOGRAPHY

Abreu BD, 'Gatekeeping Misinformation with Media Literacy Education.' (2021) 50 Knowledge Quest 26

Adena M and others, 'Radio and the Rise of The Nazis in Prewar Germany' (2015) 130 The Quarterly Journal of Economics 1885

'Astroturfing' <https://www.merriam-webster.com/dictionary/astroturfing> accessed 25 May 2022

Beauchamp Z, 'Social Media Is Rotting Democracy from Within' [2019] *Vox* <https://www.vox.com/policy-and-politics/2019/1/22/18177076/social-media-facebook-far-right-authoritarian-populism> accessed 15 May 2022

Bellagamba F, 'Dalla criminalizzazione dei discorsi d'odio all'aggravante del negazionismo: nient'altro che un prodotto della legislazione penale "simbolica"?' [2018] Criminalia 28

Benkler Y, Farris R and Roberts H, *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (Oxford University Press 2018) <https://library.oapen.org/handle/20.500.12657/28351> accessed 15 June 2022

Bennett WL, 'Killing the Golden Goose? A Framework for Regulating Disruptive Technologies' [2021] Information, Communication & Society 1

Bradford A, *The Brussels Effect: How the European Union Rules the World* (Oxford University Press 2020) <https://oxford.universitypressscholarship.com/10.1093/oso/9780190088583.001.0001/oso-9780190088583> accessed 17 June 2022

Brichacek A, 'Six Ways the Media Influence Elections' *School of Journalism and Communications-University of Oregon* (2016) <https://journalism.uoregon.edu/news/six-ways-media-influences-elections> accessed 11 May 2022

Cadwalladr C, *Facebook's Role in Brexit — and the Threat to Democracy* (2019) <https://www.youtube.com/watch?v=OQSMr-3GGvQ> accessed 16 May 2022

Charles Koch Foundation, Interview with Joshua Tucker, 'How Does Social Media Impact Democracy?' (3 March 2021) <https://charleskochfoundation.org/stories/how-does-social-media-impact-democracy/> accessed 25 May 2022

Coleman A, 'French Election: Misinformation Targets Candidates and Voting System' *BBC News* (22 April 2022) <https://www.bbc.com/news/61179620> accessed 6 July 2022

Corbyn Z, 'Facebook Experiment Boosts US Voter Turnout' [2012] Nature <https://www.nature.com/articles/nature.2012.11401> accessed 25 May 2022

Culpan T, 'TikTok Is the New Front in Election Misinformation' *Washington Post* (29 June 2022) <https://www.washingtonpost.com/business/tiktok-is-thenew-front-in-election-misinformation/2022/06/28/0aedd53e-f73b-11ec-81db-ac07a394a86b_story.html> accessed 6 July 2022

Daskalopoulos A and others, 'Thinking Outside the Bubble: Addressing Polarization and Disinformation on Social Media' *Center for Strategic and International Studies Journalism Bootcamp* (28 October 2021) <https://journalism.csis.org/thinking-outside-the-bubble-addressing-polarization-and-disinformation-on-social-media/> accessed 14 July 2022

Deck A, 'Hate Speech and Disinformation Spike on TikTok in Run-up to Kenya's Elections' *Rest of World* (8 June 2022) <https://restofworld.org/2022/hate-speech-and-disinformation-spike-on-tiktok-in-kenya-election-run-up/> accessed 9 June 2022

Deibert RJ, 'The Road to Digital Unfreedom: Three Painful Truths About Social Media' (2019) 30 Journal of Democracy 25

Druckman JN and others, 'Affective Polarization, Local Contexts and Public Opinion in America' (2021) 5 Nature Human Behaviour 28

Eisenstat Y, 'How to Hold Social Media Accountable for Undermining Democracy' [2021] Harvard Business Review 5

Etter L, 'Rodrigo Duterte Turned Facebook Into a Weapon, With a Little Help From Facebook' *Bloomberg News* (7 December 2017) <https://www.bloomberg.com/news/features/2017-12-07/how-rodrigo-duterte-turned-facebook-into-a-weapon-with-a-little-help-from-facebook> accessed 16 June 2022

Evangelista R and Bruno F, 'WhatsApp and Political Instability in Brazil: Targeted Messages and Political Radicalisation' (2019) 8 Internet Policy Review

Farkas J, 'Disguised Propaganda on Social Media: Addressing Democratic Dangers and Solutions' (2018) XXV The Brown Journal of World Affairs <https://bjwa.brown.edu/25-1/disguised-propaganda-on-social-media-addressing-democratic-dangers-and-solutions/> accessed 13 July 2022

Feiner L, 'House Democrats Say Facebook, Amazon, Alphabet, Apple Enjoy "monopoly Power" and Recommend Big Changes' *CNBC* (6 October 2020) <https://www.cnbc.com/2020/10/06/house-democrats-say-facebook-amazon-alphabet-apple-enjoy-monopoly-power.html> accessed 13 July 2022

Floridi L (ed), *The Onlife Manifesto* (Springer International Publishing 2015) <http://link.springer.com/10.1007/978-3-319-04093-6> accessed 16 May 2022

Fukuyama F and Grotto A, 'Comparative Media Regulation in the United States and Europe' in Joshua A Tucker and Nathaniel Persily (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press 2020) <https://www.cambridge.org/core/books/social-media-and-democracy/comparative-media-regulation-in-the-united-states-and-europe/0E4F255ADA3FC81BDC4365FF10DFDF3A> accessed 17 June 2022

Gagliardone I and others, *Countering Online Hate Speech* (UNESCO 2015)

Garnett HA and James TS, 'Cyber Elections in the Digital Age: Threats and Opportunities of Technology for Electoral Integrity' (2020) 19 Election Law Journal: Rules, Politics, and Policy 111

Ghosh D, 'Are We Entering a New Era of Social Media Regulation?' [2021] Harvard Business Review 5

Gomes A, Antonialli D and Dias Oliva T, 'Drag Queens and Artificial Intelligence: Should Computers Decide What Is "Toxic" on the Internet?' [2019] *InternetLab* <https://internetlab.org.br/en/news/drag-queens-and-artificial-intelligence-should-computers-decide-what-is-toxic-on-the-internet/> accessed 11 July 2022

Gori G, 'Social media ed elezioni. I limiti del diritto e il rischio di una modulated democracy' (2017) 26 Informatica e diritto

Hao K, 'Troll Farms Reached 140 Million Americans a Month on Facebook before 2020 Election, Internal Report Shows' [2021] MIT Technology Review <https://www.technologyreview.com/2021/09/16/1035851/facebook-troll-farms-report-us-2020-election/> accessed 13 July 2022

Hao K, 'How Facebook and Google Fund Global Misinformation' [2021] MIT Technology Review <https://www.technologyreview.com/2021/11/20/1039076/facebook-google-disinformation-clickbait/> accessed 17 May 2022

Hern A, 'Older People More Likely to Share Fake News on Facebook, Study Finds' *The Guardian* (10 January 2019) <https://www.theguardian.com/technology/2019/jan/10/older-people-more-likely-to-share-fake-news-on-facebook> accessed 11 July 2022

Howard A, 'US Election Campaign Technology from 2008 to 2018, and Beyond' [2018] MIT Technology Review <https://www.technologyreview.com/2018/08/22/140643/us-election-campaign-technology-from-2008-to-2018-and-beyond/> accessed 24 May 2022

Judit B and others, 'The Fight against Disinformation and the Right to Freedom of Expression'

Kolbert E, 'Why Facts Don't Change Our Minds' [2017] *The New Yorker* <https://www.newyorker.com/magazine/2017/02/27/why-facts-dont-change-our-minds> accessed 13 July 2022

Kuklis L, 'Video-Sharing Platforms In AVMSD – A New Kind Of Content Regulation', *Research Handbook on EU Media Law and Policy* (Elgar Publishing 2021) <https://papers.ssrn.com/abstract=3527512> accessed 18 June 2022

Landwehr M, Borning A and Wulf V, 'Problems with Surveillance Capitalism and Possible Alternatives for IT Infrastructure' [2021] Information, Communication & Society 1

'League's Salvini: Center Right Ready to Run Italy' (*POLITICO*, 5 March 2018) <https://www.politico.eu/article/matteo-salvini-leagues-center-right-ready-to-run-italy/> accessed 25 May 2022

Lichfield G, 'Why the Pessimists Are Winning, for Now' [2018] *MIT Technology Review* <https://www.technologyreview.com/2018/08/22/140680/why-the-pessimists-are-winning-for-now/> accessed 15 May 2022

Lindroos-Hovinheimo S, 'The Proposed EU Regulation on Political Advertising Has Good Intentions, But Too Wide a Scope' (*European Law Blog*, 23 February 2022) <https://europeanlawblog.eu/2022/02/23/the-proposed-eu-regulation-on-political-advertising-has-good-intentions-but-too-wide-a-scope/> accessed 21 June 2022

Machado C, 'WhatsApp's Influence in the Brazilian Election and How It Helped Jair Bolsonaro Win' [2018] Council on Foreign Relations <https://www.cfr.org/blog/whatsapps-influence-brazilian-election-and-how-it-helped-jair-bolsonaro-win> accessed 15 June 2022

Margetts H, 'Rethinking Democracy with Social Media' (2019) 90 The Political Quarterly 107

Martin D, 'Papua New Guinea to Ban Facebook for One Month, Broaches State-Run Social Network | DW | 29.05.2018' *DW.COM* (29 May 2018) <https://www.dw.com/en/papua-new-guinea-to-ban-facebook-for-one-month-broaches-state-run-social-network/a-43982064> accessed 13 July 2022

Matei SA, Rebillard F and Rochelandet F (eds), *Digital and Social Media Regulation: A Comparative Perspective of the US and Europe* (Springer International Publishing 2021) <https://link.springer.com/10.1007/978-3-030-66759-7> accessed 12 May 2022

Matsa KE and others, 'Younger Europeans Are Far More Likely to Get News from Social Media' (*Pew Research Center's Journalism Project*, 30 October 2018) <https://www.pewresearch.org/journalism/2018/10/30/younger-europeans-are-far-more-likely-to-get-news-from-social-media/> accessed 15 May 2022

Maurer T and Brattberg E, 'Russian Election Interference: Europe's Counter to Fake News and Cyber Attacks' [2018] Carnegie Endowment for International Peace <https://carnegieendowment.org/2018/05/23/russian-election-interference-europe-s-counter-to-fake-news-and-cyber-attacks-pub-76435> accessed 13 July 2022

Napoli PM, *Social Media and the Public Interest: Media Regulation in the Disinformation Age* (Columbia University Press 2019)

Nielsen RK and Fletcher R, 'Democratic Creative Destruction? The Effect of a Changing Media Landscape on Democracy' in Joshua A Tucker and Nathaniel Persily (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press 2020) <https://www.cambridge.org/core/books/social-media-and-democracy/democratic-creative-destruction-the-effect-of-a-changing-media-landscape-on-democracy/8C6548E16FA63289FC4C731AC512B075> accessed 14 July 2022

Ott BL and Dickinson G, 'Twittering Away Our Deliberative Capacity: Social Media and the Threat to Democracy' [2021] Research Outreach <https://researchoutreach.org/articles/twittering-deliberative-capacity-social-media-threat-democracy/> accessed 16 May 2022

Owen D, 'The Past Decade and Future of Political Media: The Ascendance of Social Media', *Towards a New Enlightenment? A Transcendent Decade* (BBVA OpenMind 2019)

Petrova M, Sen A and Yildirim P, 'Social Media and Political Contributions: The Impact of New Technology on Political Competition' (2021) 67 Management Science 2997

Piore A, 'No, Big Tech Didn't Make Us Polarized (but It Sure Helps): Social-Media Bubbles Tell Only Part of the Story of Why We're so Divided. The Rest Is in Our Heads.' [2018] MIT Technology Review <https://www.thefreelibrary.com/No%2C+big+tech+didn%27t+make+us+polarized+(but+it+sure+helps)%3A...-a0554051593> accessed 24 May 2022

'Protecting Electoral Integrity In The Digital Age' (Kofi Annan Commission on Elections and Democracy in the Digital Age 2020)

Puddephatt A, 'Social Media and Elections' (2019) 14 Cuadernos de discusión de comunicación e información <https://unesdoc.unesco.org/ark:/48223/pf0000370634> accessed 16 June 2022

Rinehart A, 'Fake News. It's Complicated.' [2017] *First Draft* <https://firstdraftnews.org:443/articles/fake-news-complicated/> accessed 16 July 2022

Rowbottom J, 'Campaign Finance and Electoral Speech in the Media' in James A Gardner (ed), *Comparative Election Law* (Edward Elgar Publishing 2022) <https://www.elgaronline.com/view/edcoll/9781788119016/9781788119016.xml>

'Safeguarding the Legitimacy of Elections: The Kofi Annan Commission Launches Final Report' (*Kofi Annan Foundation*, 23 January 2020) <https://www.kofiannanfoundation.org/supporting-democracy-and-elections-with-integrity/safeguarding-the-legitimacy-of-elections-the-kofi-annan-commission-launches-final-report/> accessed 7 June 2022

Satariano A, 'Facebook Hearing Strengthens Calls for Regulation in Europe' *The New York Times* (6 October 2021) <https://www.nytimes.com/2021/10/06/technology/facebook-european-union-regulation.html> accessed 6 June 2022

Satariano A, 'E.U. Takes Aim at Big Tech's Power With Landmark Digital Act' *The New York Times* (24 March 2022) <https://www.nytimes.com/2022/03/24/technology/eu-regulation-apple-meta-google.html> accessed 6 June 2022

Schwab K, *The Fourth Industrial Revolution* (Crown Publishing Group 2016) <https://www.penguin.co.uk/books/304971/the-fourth-industrial-revolution/9780241300756> accessed 21 June 2022

Siegel AA, 'Online Hate Speech' in Joshua A Tucker and Nathaniel Persily (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press 2020) <https://www.cambridge.org/core/books/social-media-and-democracy/online-hate-speech/28D1CF2E6D81712A6F1409ED32808BF1> accessed 17 June 2022

Silverman C, 'The Backfire Effect' [2011] Columbia Journalism Review <https://www.cjr.org/behind_the_news/the_backfire_effect.php> accessed 13 July 2022

'Social Media' <https://www.merriam-webster.com/dictionary/social+media> accessed 15 May 2022

Somer M and McCoy J, 'Déjà vu? Polarization and Endangered Democracies in the 21st Century' (2018) 62 American Behavioral Scientist 3

Spigno I, *Discorsi d'odio. Modelli Costituzionali a Confronto* (Giuffrè 2018) <http://www.libreriadelgiurista.it/libri/9788814228322/discorsi-dodio-modelli-costituzionali-a-confronto.html> accessed 14 July 2022

Splidsboel Hansen F, 'Russian Hybrid Warfare: A Study of Disinformation' (Danish Institute for International Studies 2017)

Stockmann D, 'Tech Companies and the Public Interest: The Role of the State in Governing Social Media Platforms' [2022] Information, Communication & Society 1

Tambini D, 'Social Media Power and Election Legitimacy' in Damian Tambini and Martin Moore (eds) (Oxford University Press 2018) <https://global.oup.com/academic/product/digital-dominance-9780190845124?cc=gb&lang=en&> accessed 12 July 2022

Taylor ME, 'Egyptian Government Launches New Social Media Platform to Rival Facebook' *Face2Face Africa* (26 March 2018) <https://face2faceafrica.com/article/egyptian-government-launches-new-social-media-platform-rival-facebook> accessed 13 July 2022

Tufekci Z, 'How Social Media Took Us from Tahrir Square to Donald Trump' [2018] MIT Technology Review <https://www.technologyreview.com/2018/08/14/240325/how-social-media-took-us-from-tahrir-square-to-donald-trump/> accessed 24 May 2022

Walker M and Matsa KE, 'News Consumption Across Social Media in 2021' [2021] *Pew Research Center's Journalism Project* <https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/> accessed 15 May 2022

Wardle C and Derakhshan H, 'Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making' (Council of Europe 2017) <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html> accessed 14 July 2022

Wischmeyer T, 'Making Social Media an Instrument of Democracy' (2019) 25 European Law Journal 169

Woolley SC, 'Bots and Computational Propaganda: Automation for Communication and Control' in Nathaniel Persily and Joshua A Tucker (eds), *Social Media and Democracy: The State of the Field, Prospects for Reform* (Cambridge University Press 2020) <https://www.cambridge.org/core/books/social-media-and-democracy/E79E2BBF03C18C3A56A5CC393698F117> accessed 16 May 2022

Yanagizawa-Drott D, 'Propaganda and Conflict: Evidence from the Rwandan Genocide *' (2014) 129 The Quarterly Journal of Economics 1947

Youngs R and others, 'European Democracy Support Annual Review 2021' [2021] Carnegie Endowment for International Peace 58

Zuboff S, *The Age of Surveillance Capitalism* (PublicAffairs 2019) <https://www.publicaffairsbooks.com/titles/shoshana-zuboff/the-age-of-surveillance-capitalism/9781610395694/> accessed 13 July 2022

Zuckerberg M, 'The Internet Needs New Rules. Let's Start in These Four Areas.' *Washington Post* (30 March 2019) <https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html> accessed 16 July 2022

Zuckerman E, 'The Internet's Original Sin' *The Atlantic* (14 August 2014) <https://www.theatlantic.com/technology/archive/2014/08/advertising-is-the-internets-original-sin/376041/> accessed 16 July 2022

Zuckerman E, 'Self-Segregation on Social Networks and the Implications for the Ferguson, MO Story' *EthanZuckerman.com* (27 August 2014) <https://ethanzuckerman.com/2014/08/27/self-

segregation-on-social-networks-and-the-implications-for-the-ferguson-mo-story/> accessed 6 July 2022