STUDY

Requested by the LIBE committee

European Parliament

Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States

Policy Department for Citizens' Rights and Constitutional Affairs
Directorate General for Internal Policies of the Union
PE 608.864 - February 2019

EN

Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States

STUDY

Abstract

This study, commissioned by the European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs and requested by the European Parliament's Committee on Civil Liberties, Justice and Home Affairs, assesses the impact of disinformation and strategic political propaganda disseminated through online social media sites. It examines effects on the functioning of the rule of law, democracy and fundamental rights in the EU and its Member States.

The study formulates recommendations on how to tackle this threat to human rights, democracy and the rule of law. It specifically addresses the role of social media platform providers in this regard.

ABOUT THE PUBLICATION

This research paper was requested by the European Parliament's Committee on Civil Liberties, Justice and Home Affairs (LIBE) and was commissioned, overseen and published by the Policy Department for Citizens' Rights and Constitutional Affairs.

Policy Departments provide independent expertise, both in-house and externally, to support European Parliament committees and other parliamentary bodies in shaping legislation and exercising democratic scrutiny over EU external and internal policies.

To contact the Policy Department for Citizens' Rights and Constitutional Affairs or to subscribe to its newsletter please write to: poldep-citizens@europarl.europa.eu

RESPONSIBLE RESEARCH ADMINISTRATOR

Marion SCHMID-DRÜNER
Policy Department for Citizens' Rights and Constitutional Affairs
European Parliament
B-1047 Brussels
E-mail: poldep-citizens@europarl.europa.eu, marion.schmid-druener@ep.europa.eu

AUTHORS

Judit BAYER (scientific coordinator, editor), Budapest Business School
Natalija BITIUKOVA, Independent consultant
Petra BÁRD, Central European University
Judit SZAKÁCS, Center for Media, Data and Society at the Central European University
Alberto ALEMANNO, HEC Paris
Erik USZKIEWICZ, Hungarian Europe Society

CONTRIBUTORS

Sergio CARRERA, CEPS/ European University Institute/ Science Po/ Maastricht, contributed with comments and providing inputs to the Study.
Julia GUERIN, CEPS
Lina VOSYLIUTE, CEPS

ACKNOWLEDGEMENT

We are grateful to Kristina Irion and Éva Simon, as well as the other reviewers, for their valuable comments and ideas related to our paper.

LINGUISTIC VERSION

Original: EN

Manuscript completed in February 2019
© European Union, 2019

This document is available on the internet at:
http://www.europarl.europa.eu/supporting-analyses

DISCLAIMER

The opinions expressed in this document are the sole responsibility of the authors and do not necessarily represent the official position of the European Parliament.

CONTENTS

LIST OF ABBREVIATIONS

ACHPR	African Commission on Human and People's Rights
AI	Artificial intelligence
AVMS	Audiovisual and media services
ECHR	European Convention on Human Rights
ECJ	European Court of Justice
ECtHR	European Court of Human Rights
EDPS	European Data Protection Supervisor
EEAS	European External Action Service
EP	European Parliament
ERC	European Research Council
EU	European Union
GDPR	General Data Protection Regulation
GFCC	German Federal Constitutional Court
HLEG	High-Level Expert Group on Fake News and Online Disinformation
IATE	Inter-Active Terminology for Europe
IRA	Internet Research Agency
M5S	Movimento Cinque Stelle
NGO	Non-governmental organisation
OAS	Organization of American States
OLAF	European Anti-Fraud Office
OSCE	Organisation for Security and Co-operation in Europe
PACE	Parliamentary Assembly of the Council of Europe
ROG	Reporter ohne Grenzen
TEU	Treaty on European Union
TFEU	Treaty on the Functioning of the European Union
UK	United Kingdom
UN	United Nations
US	United States
VR	Virtual reality

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Scope

This study examines the causes and impact of disinformation and propaganda on democracy, human rights and the rule of law within the European Union. It analyses how new technology has transformed the operation and structure of the democratic public sphere in general, and in particular explores the recently experienced events of disinformation and propaganda campaigns in the light of interference with democratic processes through the manipulation of public opinion, as well as the international and national legislative and self-regulatory initiatives (discussed below). Finally, it recommends policy actions to correct those mechanisms that underlie the examined phenomena.

Academics, policy makers and journalists use a variety of terms to describe what is commonly called 'fake news'. Through the comparative analysis of an interdisciplinary pool of sources, the authors of this study use the terms 'disinformation' and 'propaganda' to designate content that is false, published with an intended strategic–political effect on an issue of public interest.

Starting from the research on 'post-truth' media in general, the authors have narrowed the focus by omitting diffuse misinformation, conspiracy theories and unpaid trolling, as well as offline propaganda and hate speech. The discussed disinformation could originate either from governmental forces or non-state actors, whether domestic or foreign. Elements of disinformation and propaganda are that such information (i) is designed to be wholly or partly false, manipulated or misleading, or uses unethical persuasion techniques; (ii) regards an issue of public interest; (iii) has the intention to generate insecurity, hostility or polarisation, or attempts to disrupt democratic processes; (iv) and is disseminated and/or amplified through automated and aggressive techniques, such as social bots, artificial intelligence (AI), micro-targeting or paid human 'trolls', often used to boost public visibility.

State of play

Major disinformation campaigns of the past four years illustrate the alleged interference with democratic processes, in particular elections and referenda. Drawing definitive conclusions from these would be beyond the limits of this study, although a few observations may be made by the reader. It should be noted that an exact causal correlation between disinformation and the political opinion and voting behaviour of individuals is not yet scientifically proven.[1] Nevertheless, the effect of media content on the audience has been greatly contested in relation to the traditional media, and various psychological experiments have supported contradicting theories,[2] although this ambiguity has not prevented legitimate regulation of mass media.

According to a recent study, a significant generational divide can be observed: people over 65 share seven times more fake news than young users do.[3] In addition, the sinking popularity of Facebook and the growing popularity of messaging services such as Snapchat[4] may also signal that this phenomenon, which has dominated public concerns for democracy in the past few years, may be taking a new direction.

The study explores the legal framework of social media platforms, including their place and assumed responsibility in the legal order, among various information-society service providers. It is found that social media

[1] Roozenbeek, Jon and van der Linden, Sander, The Fake News Game: Actively Inoculating Against the Risk of Misinformation, From: https://www.cam.ac.uk/sites/www.cam.ac.uk/files/fakenews_latest_jrr_aaas.pdf , pp. 3-4.

[2] See the contesting theories of Harold Lasswell (bullet, 1927), Paul Lazarsfeld (two-step influence, 1948), Joseph Klapper (selective perception, 1949), George Gerbner (cultivation, 1969), McCombs and Shaw (agenda-setting, 1972), Herman and Chomsky (framing, 1988), Dayan and Katz (performative effect, 1992) - to name a few.

[3] Andrew Guess, Jonathan Nagler and Joshua Tucker: Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Sci Adv* **5** (1), eaau4586. DOI: 10.1126/sciadv.aau4586

[4] Kantar Media: News in social media and messaging apps. Qualitative research report Prepared for the Reuters Institute for the Study of Journalism, University of Oxford with the support of the Google News Initiative. Sept. 2018.

service, which emerged after 2000, is not defined and its liability is not set out consistently by the relevant legal instruments. These include the E-Commerce Directive, the Audiovisual Media Services (AVMS) Directive, the ePrivacy Directive and the proposed ePrivacy Regulation, the Code of Conduct on countering illegal hate speech, the Commission Recommendation on measures to effectively tackle illegal content online, the Communication from the Commission on tackling online disinformation, the European Council decision of March 2018 and the Proposal for a Regulation on preventing the dissemination of terrorist content online. Based on these documents, the study uses the term 'platform providers' to designate those services that facilitate, organise and amplify the transmission of third-party content, through actions of their registered users. While platforms are ubiquitous also in other business sectors (for example, eBay), social media is a subcategory of theirs. **In accordance with many leading international actors, including the UN, Organisation for Security and Co-operation in Europe (OSCE) and Council of Europe, this study represents the position that platforms should not be made liable for third-party content.**

The human rights background for an envisaged policy framework has taken into account the following:

- international agreements and actions of the UN, such as the Joint Declaration of OSCE, the Organization of American States and the African Commission on Human and People's Rights (ACHPR);
- the Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, which emphasises that **the impact of these companies on the public sphere demands that they open themselves up to public accountability**;
- a Joint Letter under the Mandates of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression; and
- the **UN Guiding Principles on business and human rights**, which provides for the corporate responsibility of all business enterprises to respect human rights.[5]

Court cases such as *Delfi* v *Estonia*, *MTE and Index* v *Hungary* and *L'Oréal* v *eBay* have been compared with the conclusion that the legal situation of platform operators needs legislative intervention.

The latest policies and legal measures developed at the Member State and the EU level to tackle disinformation and propaganda have been collected and analysed in a critical perspective, primarily including the German Network Enforcement Act, the French Act against Informational Manipulation and the Italian law against fake news, along with the co-regulatory initiative between the French government and Facebook, the Code of Practice to tackle online disinformation and the Commission Action Plan Against Disinformation. The analysis finds that the **legal restriction of content may pose a greater harm to democracy than disinformation itself.**

To provide a background for our assessment on how disinformation and propaganda may affect democratic elections, some national rules relating to election campaigns have been compared for identifying the anchors that may counter disinformation and propaganda. Three main threads could be identified: (i) the regulation of political advertising (e.g. Poland); (ii) the strict supervision of campaign finances (e.g. Portugal); and (iii) increasing awareness and media literacy (e.g. Sweden).

[5] UN Guiding Principles on Business and Human Rights Implementing the UN "Protect, Respect and Remedy Framework". 2011. OCHCHR. https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

Theoretical approach

The impact on democracy

Disinformation and propaganda events interfere with democracy in two ways: (i) they dominate and distort the public discourse and corrupt the process of democratic decision-making, and (ii) when this process leads to political success, the political force that won the elections through manipulation might capture the state and deconstruct the constitutional system. This process is very difficult to stop once the anti-democratic party is in power; but it could be prevented or slowed down with the tools of militant democracy – the self-defending constitutional state.

The emergence of social media marks the beginning of a new age of the public sphere (*Öffentlichkeit*).[6] This user-friendly communication interface allows for the publishing of content without the economic or educational entrance barriers; it facilitates the formation of groups and the creation of a "global village".[7] This decentralised and horizontal discussion cannot be supervised with the same instruments as the centrally organised, traditional mass media. This control vacuum has allowed rapid innovations in line with business interests, and become exploited by political opportunists. The authors argue that while the ubiquitous content itself can hardly be controlled, the architecture of this communication – algorithms and data flow – should.

The post-modern global world is characterised by political and existential uncertainties and threats that are immaterial and invisible (migration, terrorism, climate change, genetically modified food, etc.). The collapse of the hierarchical architectures of knowledge transmission (for instance through media, education and the church) has left behind a deficit of trust, a culture of relativism and what is called the 'post-truth era'. The culture of knowledge has been replaced by a culture of risks: a complex web of collective strategies exists through which fear, angst and anxiety are created and recycled. While social media theoretically has given voice to people who were underrepresented earlier by traditional media, their dissatisfaction has been exploited by political hijackers. Populist communication uses these people's likes to amplify their manipulative propaganda. The populist rhetoric pretends to represent the underprivileged, but in fact it supports the interests of another elite.

Human rights impact

The impact of disinformation and propaganda on human rights is divided in two main categories: (1) impact on data protection, privacy, human dignity and autonomy; and (2) violation of the rights of freedom of expression and the right to seek and receive information.

1) **Impact on privacy and data protection.** Personal data are the currency and the fuel that keep business and innovation moving. Data-driven business models seem to further expand on the supply of data ensured by the giant digital platforms, which experiment with the application of AI and machine learning based on the gigantic personal databases they control. Experimenting with psychological reactions of masses of people should be regulated or ruled out, similar to biological experimenting.

2) **Freedom of expression and freedom to receive information.** An open public discourse is one of the basic conditions of democracy, because this is how citizens can discuss their common matters, form political opinions and ultimately reach a political decision (e.g. voting in elections). To have a lively and rational discourse, media freedom, individual freedom of expression and the right to receive information are equally needed. Today's media environment gives individuals the chance to express their ideas at every possible instance – in this respect, the pluralism of ideas is overwhelming. This overwhelming volume of information makes navigation and access to trustworthy information a hard task. The (weakened) media system's earlier function of gatekeeping included filtering through professional

[6] Habermas, Jürgen: The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society. MIT Press, 1991 (Sixth edition).

[7] McLuhan, Marshall: Understanding Media. The extensions of Man. MIT Press, 1994 [1964].

editing, agenda defining and control by the political elite. These often-criticised checks also contributed to the stability of democratic systems.

Future technologies

Looking into the future of disruptive technologies, the authors find that the new services' reliance on personal data will only intensify. So too will the capacity to imitate reality, for example through augmented reality or virtual reality (VR), which enables falsification to become unrecognisable by both human and machine control (e.g. deep fakes).

Machine learning, advanced demographic analytics, the Internet of Things, voice and facial recognition appear to further increase the vulnerability of humans to erosion of privacy and having their personal data involuntarily exposed. **Without regulatory intervention, business services and technology developers will take exploitation of human psychological traits and social engineering to new heights.** Regulation should set the rules of the game, ideally with the cooperation of a wide range of stakeholders and global partners.

Recommendations

Disinformation and propaganda are symptoms of deeper structural problems in our societies and media environments. Rather than targeting the content itself, the vulnerabilities that these narratives exploit should be identified and addressed. The recommendations are divided into two sections: **strengthening democratic resilience and adapting media policy.** The first section includes imminent actions relating to the coming European Parliament (EP) elections, regulation of political and public issue advertising, data protection, civic education, mainstreaming science in policy-making and further research. The second section includes strengthening pillars of trust in the media and the obligations of platform providers.

Imminent actions relating to the 2019 EP elections

The existing European External Action Service (EEAS) election observation service or OSCE observation could be applied to monitor the EP elections in Member States similar to third-country national elections. For the future, however, a specific EU institutional capacity is recommended in the form of a supranational electoral authority with sufficient powers to monitor and undertake field visits to Member States, and to supervise political campaigns preceding the EP elections. The European Court of Auditors and the European Anti-Fraud Office (OLAF) should pursue the investigation of campaign finances, including sponsorship of social media advertisements.

Political and public issue advertising

The existing rules on commercial and political advertising should be applied to the online environment, including social media platforms, such as the identification and separation of advertisements clearly from all other content. With regard to the international media environment and with the objective of safeguarding democratic processes, it is recommended that the rules on all advertising, including political and public issue advertising, are harmonised in a directive or regulation. Besides the traditional principles that apply to commercial advertising, new rules are recommended. It is also suggested that the broader category of 'public issue advertising' is applied to cover political and other issue-based advertising, because some issues can become symbols of political views or action (for example migration or EU integration itself).

In the social media environment, labelling political ads may not be sufficient, because influencers' posts do not qualify as ads in the traditional sense. Therefore, besides the fact of sponsorship, also the source needs to be identified, and influencers should be identified as such. Greater transparency should be ensured by the force of law, in respect of the buyer of the advertisement, the publisher, in whose interest it is published, the targeted audience, targeting criteria and its reach. Users should have access to a repository providing information about what political and public issue ads they are targeted with. It is suggested that contracts between political parties and platforms are deposited and open for public scrutiny. Bots and automated accounts or AI should not be allowed to disseminate political and public issue ads.

Campaign financing

To regain voters' trust in the democratic process, campaign financing rules need to be revised profoundly. Should the EU be able to engage in this process, its impact on renewing democracy would have a **global historical significance**. The rules should include the requirement of more data on campaign spending, indicating the contracting media partner, the nature of the media content, the targeting criteria and supervision independent from the political parties. Until this happens, the new competence of the European Public Prosecutor's Office should be applied to monitor party expenditures.

Privacy and data protection

Tracking-based online advertising must be de-incentivised by rigorous enforcement of the existing ePrivacy Directive, the General Data Protection Regulation (GDPR) and by adopting, at the soonest, a robust ePrivacy Regulation, which outlaws 'tracking walls' and includes other safeguards as advocated by the regulators in the field like the European Data Protection Supervisor (EDPS).

Platform providers have to be responsible for protecting the personal data of their users, including the prevention of unlawful data mining on their respective platforms. Data protection authorities should proactively exercise their powers under Article 58 of the GDPR, including a power to carry out investigations into political micro-targeting practices, against (political) advertisers, digital platforms and intermediaries (data analytics companies, data brokers, etc.). The data protection authorities ought to proactively enforce the implementation of data protection by design and data protection by default obligations of the controllers, as well as the rules pertaining to free, unambiguous and informed consent. Political micro-targeting needs to be recognised as solely automated decision-making that produces significant effects on individuals under Article 22 of the GDPR.

Given the considerable civic powers of the social media platforms, their data protection-related obligations should exceed the minimum requirements enshrined in the EU and national data protection laws. For instance, the digital platforms would have to maintain a searchable repository of active and historical political and issue-based advertising targeting persons in the EU with detailed information about the criteria of targeting, buyers, etc.

Media policy

Communication policy may have two vectors: supply and demand. Regulatory intervention on the supply side could diminish the amount of disinformation and propaganda that are disseminated, and dilute it with trustworthy information. The means to achieve this include regulation of the social media environment. Platforms should not be liable for third-party content, but need to be responsible for expediently administering their platforms: to protect the personal data and privacy of their users (not only the registered); to ensure that their algorithms discriminate neither among content nor users; to separate sponsored content and advertisements from other content; to create algorithms fostering and promoting diversity of content; to ensure transparency of their algorithms and offer options to users in selecting their settings for content, including diversity and the option to identify and disable fake accounts. Users who regularly reach large audiences with public issue content have to be distinguished as 'influencers' (including for example, politicians and non-governmental organisations (NGOs)). Instead of the notice-and-takedown system, the notice-and-notice system is recommended, except for cases of manifestly and dangerously illegal content. With a dominant market share, the responsibilities should be higher. Self-regulation should not replace enforceable regulation, in particular in the areas of political and issue-based advertising, and privacy protection. Any other self-regulation efforts by the digital platforms (e.g. content moderation, ad transparency initiatives) ought to be subject to external scrutiny and impact assessment in order to determine their efficiency and compliance with the fundamental rights.

The means to intervene on the demand side aim at raising awareness and improving media literacy. Programmes on media literacy need to be complemented by civic education in all Member States on EU values of democracy and human rights at all levels, which should start with the appropriate training of teaching staff. Where necessary, the organisation of local programmes are recommended, for example utilising travelling buses to include those parts of society otherwise difficult to reach.

To create pillars of trust, the Commission should initiate and continue to support programmes for investigative journalism and for improving fact-checking services and credibility indices. It should also promote the creation of a common, European, high-quality media service transmitted by contemporary technology**,** with a view to enhancing EU cohesion, offering a common European perspective and developing the European narrative. Special attention should be paid to the ethical operation of public service media, and accusations of engaging in dissemination of disinformation and propaganda have to be investigated by the Commission.

INTRODUCTION

Description of the problem

In recent years, the world has been witnessing a revival of populism. What could be viewed as an exceptional undercurrent in 2000, has tripled its number of votes by today.[8] Since 2016, the year of Brexit and the election of Donald Trump, it has become clear that the global political trend has taken a new turn, and this direction has only been reinforced by some of the recent elections in Europe and in Brazil. The elections to the European Parliament are anticipated with great excitement on both sides of the political spectrum.[9]

The phenomenon is examined and analysed by several excellent research institutions, political and media analysts. What has caused this to happen, and where does it lead us? While the thriving of populism can have diverse causes, one element stands out: the sweeping transformation of the global public sphere. Public discourse has gone global, and been completely restructured as regards its actors, gatekeepers, influencers and audience. Social media created a new public space that can be flexibly utilised to get any message to selected parts of the audience. It has no national boundaries, a feature that can be used for good or bad purposes.

The results of the Brexit referendum came as a shock to many in the European Union and beyond, but a bigger shock came when reports about the falsity of the political arguments, and later about the manipulative elements of the campaign, were published. Meanwhile, watching the US political campaign that led to the victory of Donald Trump, and the following revelations about the data abuse by Cambridge Analytica, shattered what we had thought about representational democracy and election campaigns.

Understanding the patterns may be a complex task in the overabundant media environment. This study undertakes to set an order in this chaos and break down disinformation and information manipulation into types and categories, to provide a clear picture of the actors, events and processes, to analyse the phenomena in the light of existing and anticipated legislative instruments, and to propose new directions for policy action.

Scope of the problem

In 2016, 'post-truth' was designated as the word of the year, signalling the crisis of journalism and the growing presence of misinformation and disinformation (on definitions, see chapter 1). 'Fake news' was designated as the term of the year by the *Collins English Dictionary* in 2017, in the same year when the World Economic Forum addressed the problem of misinformation, disinformation and propaganda in its risk assessment, also pointing at the risk caused by the decreasing trust in institutions.

Manipulation and propaganda are as old as the hills – so what has changed in recent years? The manipulative campaigns of 2016 were organised strategically, as well-financed, concerted actions of a professional team, with the intent to influence – domestic or foreign – political processes. New technology made organising such campaigns significantly more accessible, promising a higher likelihood of success, both faster and with practically no risk. On the other hand, the same technology leaves traces, and enables investigation and revelation of the malicious actions.

The relationship between cause and effect still needs to be scientifically proven, but in at least the few cases mentioned above, the disinformation and propaganda actions appeared to have made a significant impact on public discourse and on voting behaviour. Some of the studies in the field suggest that manipulation of people's newsfeed or search results could influence their voting behaviour. Ensuring fair elections after all these events requires additional efforts: investment into monitoring, policy initiatives and regulatory actions.

[8] 'Revealed: one in four Europeans vote populist'. 20. Nov, 2018. https://www.theguardian.com/world/ng-interactive/2018/nov/20/revealed-one-in-four-europeans-vote-populist

[9] See for example: 'Hungary's Orban eyes EU takeover by anti-immigration parties'. Jan 10, 2019. https://www.irishtimes.com/news/world/europe/hungary-s-orban-eyes-eu-takeover-by-anti-immigration-parties-1.3753892

Democracies presuppose a lively public political discussion, which should not be stifled by over-harsh regulation of expression. The discussion of political issues enjoys the highest protection in liberal democracies. Is there a way to minimise the effects of disinformation and propaganda, while preserving the openness of the public discourse? Early internet optimists envisaged a direct democratic participation and a global democracy enabled by the online network. Indeed, online communication allows views that used to be suppressed in traditional media to emerge and be heard, and even amplified through echo chambers. This feature – typically represented by social media platforms – enables the overthrow of established political systems. What was celebrated as the Arab Spring, now threatens European and other democracies, even breaking down the myth of American exceptionalism (the belief that the United States is a uniquely liberal nation based on wide popular support of democratic ideals and personal liberty). The question is whether the views that suddenly emerge and gain political support are really those of the people? Or do they represent only a political group hunting for votes through spreading populistic propaganda and disinformation? Social media provides politicians direct access to the audience without their messages being reframed by professional journalists. The features of populist communication strategies (e.g. people-centrism, anti-elitism, promoting direct democracy) perfectly align with social media characteristics (see more in chapter 2, section 2.1).[10]

Legitimate criticism of the ruling government, and dissent over the status quo gets mixed with manipulation and propaganda by foreign governmental forces. Russia is suspected of being behind many of the disinformation actions, but they deny it and evidence is insufficient to form a basis for international legal consequences. And while the informational spaces of liberal democracies are open and accessible to anyone around the globe, Russian and Chinese counterparts are controlled and protected. Some of the disinformation and propaganda actions relate to raising hostility against 'outgroups', such as migrants or national minorities. Increasing polarisation is an outspoken purpose of the Kremlin's information war "to destabilise a society and a state through massive psychological conditioning of the population, and also to pressure a state to make decisions that are in the interest of the opponent".[11]

European integration, as a sui generis formation of sovereign states, is particularly vulnerable as regards its fragile cohesion and slow reaction time. The Union is based on the common values of the Member States and the consequential mutual confidence between them: if Member States disagree in their values such as the respect for the rule of law, democracy and fundamental rights, then mutual confidence is shaken. European democracy is rooted in the democratic legitimacy of the representatives of Member States. If this legitimacy can be questioned, then the democratic legitimacy of EU institutions and their actions becomes questionable as well.

Concerted propaganda campaigns can have the largest impact in societies where media freedom and pluralism have already been limited, and people are deprived of the possibility to check the information against independent sources. If a populistic party is in government, there is a clear risk of a democratic backslide. Governmental investment in the dissemination of propaganda is by definition a violation of media freedom and pluralism, and paves the way for other rule of law violations. In correlation with the populist justification of acting on behalf of the 'ordinary' people and ignoring the minority and critical voices, populist governments are prone to deconstruct democratic institutions.

Populistic political rhetoric in itself would not amount to a threat to democracy. However, when user databases are processed in order to find the vulnerabilities of citizens, profiles are created and political messages are targeted at people who are most likely to be susceptible, whereas other messages are shared with different people, such micro-targeting splinters public discourse, depriving the citizenry of the right to informed political decision-making. These aggressive dissemination practices were uncovered in the political campaigns of 2016

[10] Ernst, Nicole - Sven Engesser, Florin Büchel, Sina Blassnig and Frank Esser (2017) Extreme parties and populism: an analysis of Facebook and Twitter across six countries. Information Communication and Society. 20:9, 1347-1364. at 1350.

[11] Russell, Martin: Russia's information war: Propaganda or counter-propaganda? EPRS European Parliamentary Research Service. Members' Research Service PE 589.810. at p.2. http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/589810/EPRS_BRI(2016)589810_EN.pdf

and suspected in some other campaigns. Social media companies collect an enormous amount of personal data about their users, and make them available for commercial and political actors, to enable targeted advertising. The Big Data industry can create profiles, patterns and predictions related to internet users, sometimes knowing more about them than the users themselves. Data protection laws are limited to the EU, and their effect can be assessed only after years of consequent enforcement. In the current online environment, monetisation and abuse of personal data is among the biggest threats, especially with regard to future technologies and to the lobbying power of the interested stakeholders.

Scope of the study

The study focuses on the intersection of those types of information that are published with an intended strategic–political effect on a topic of public interest, with the hypothesis that these may have the most effect on democratic processes and society. To fixate on what is absolutely relevant in the mentioned context, we have narrowed the focus from a broad concept of 'post-truth' media – which would include the diffuse misinformation, conspiracy theories and unpaid trolling, as well as offline propaganda and hate speech – to strategically disseminated political content that aims to mislead the audience.

Accepting the limited correlation between cause and effect, it has been suspected that those campaigns supposedly generated by foreign governmental actors and targeted across borders at a foreign population might have stirred the highest level of controversies (see more in chapter 2, section 2.3).

Table 1: Types of false information, categorised by intention and subject

Subject matter	**Intended strategic effect**	**No evidence of intended strategic effect**
Political/matters of public interest	***Disinformation and propaganda***	Rumours, flat-Earth, vaccination theories – harmful to society
Private interest	E.g. misleading advertisement	Gossip, celebrity rumours

Source: Authors.

Table 2: Types of disinformation and propaganda with their assessed impact on democratic values, rule of law and fundamental rights

Disinformation and propaganda	**Targeted at domestic population**	**Targeted at foreign population**
A. Source is a non-state actor, e.g. political party, or unidentified person	A1. Unethical political campaign, misleads society; if the political party is successful in the elections B1 may follow	A2. Citizens' actions, as well as disguised or unattributed attacks: similar to election hacking by 'patriotic citizens'; states may be responsible under international law for the aggressive actions of non-state actors acting on their territory against another state
B. Source is a state (governmental) actor	B1. Governmental political propaganda; clear transgression of democratic values, rule of law and human rights; captured state tries to strangle democracy; within the EU, a cause for the Article 7 mechanism	B2. Information warfare, interference with sovereignty; global threat against democracy – threatens geopolitical stability

Note: Yellow (A1) = harmful; orange (A2 and B1) = very harmful; red (B2) = critical threat.

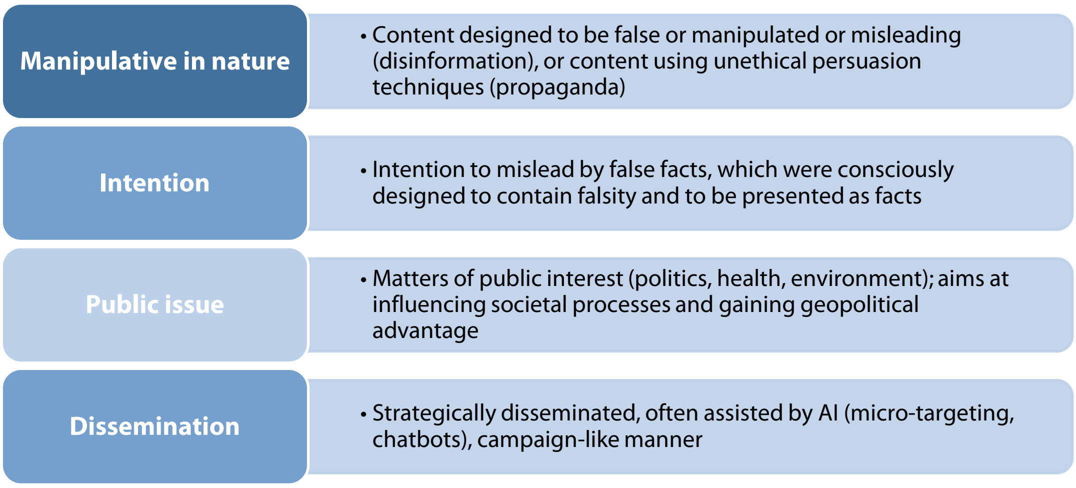
Source: Authors.

To develop an appropriate definition and terminology for what is often called 'fake news', resources in the field of communication studies, political science, internet policy, law and behavioural studies have been critically assessed (chapter 1, section 1.1). The study uses the terms 'disinformation' and 'propaganda' to describe the phenomena of information characterised by the following elements:

- is designed to be false or manipulated or misleading (disinformation), or is content using unethical persuasion techniques (propaganda);
- has the **intention** of generating insecurity, tearing cohesion or inciting hostility, or directly to disrupt democratic processes;
- is on a topic of public interest; and
- often uses automated dissemination techniques to amplify the effect of the communication.

We find that the falsity of information is not necessarily the most decisive factor among the examined phenomena. Messages intending to manipulate or divide the audience, or incite hostility against social groups, are mingled among this harmful type of political communication. Aggressive dissemination practices, however, have been identified as a distinctive factor: micro-targeted political advertising, paid trolls and political bots (see chapter 1, section 1.3.).

Figure 1: Common elements in definitions of disinformation and propaganda



Manipulative in nature
- Content designed to be false or manipulated or misleading (disinformation), or content using unethical persuasion techniques (propaganda)

Intention
- Intention to mislead by false facts, which were consciously designed to contain falsity and to be presented as facts

Public issue
- Matters of public interest (politics, health, environment); aims at influencing societal processes and gaining geopolitical advantage

Dissemination
- Strategically disseminated, often assisted by AI (micro-targeting, chatbots), campaign-like manner

The role and responsibility of social media

While social media platforms themselves do not generate content, they transmit, organise and amplify it. Their terms of service enable the massive profiling and micro-targeting that have been exploited by political campaigns. Users have been deluded into believing that the encountered information is spontaneous, citizen-generated, objective and universally encountered by other users, while in fact it may have been strategic, political and micro-targeted. The classical theorist John Milton said this about the truth: "Let her and Falsehood grapple; who ever knew Truth put to the worse, in a free and open encounter?"[12] But truth appears to play on uneven ground with falsehood, which is sometimes supported by strong financial means and by aggressive dissemination techniques.

The gravity of the situation has induced several research projects and policies as well as legislative initiatives. The study has assessed them and builds on these documents, among them the Communication of the Commission on Tackling Online Disinformation, the Council of Europe Study on the Internet and Electoral Campaigns, the European Parliament Resolution on media pluralism and media freedom in the European Union and the Resolution on EU strategic communication to counteract anti-EU propaganda by third parties. Legislative initiatives against disinformation, such as the French law as well as the Italian and Czech attempts to regulate, have been analysed, while the German Network Enforcement Act has been carefully examined to see whether it could counteract disinformation and propaganda. The European Commission has urged the creation of a code

[12] John Milton: Aeropagitica, 1644.

of practice against disinformation, which raised mixed reactions. The study includes the most recent materials available, such as the Action Plan against Disinformation, the Report from the Commission on the implementation of the Communication "Tackling online disinformation" and the Facebook Baseline Report on Implementation of the Code of Practice on Disinformation.

As a result of these instruments, significant changes have occurred during recent years in the practices of some of the most influential platforms, especially Facebook and Twitter. Although adequate regulation is missing, and the roles and responsibilities of social media platforms are still not defined, self-regulation and policy actions have made some difference in the online environment, which hosts the public discourse. On legal policy, one significant improvement is that a content-oriented approach is being replaced by a platform administration-oriented and content-neutral (or: content-agnostic) approach. Importantly, content-oriented regulation that would oblige social media companies to erase questionable content would put pressure on them to exercise censorship and is likely to sacrifice a large volume of content otherwise protected by the right to freedom of expression. At the same time, those people who have uploaded the problematic content could get away and upload the content again and again – without adequate rules on platform administration.

The positive effect of the well-intended changes is yet to be seen, while many other things are still to be done – for example, the passing and enforcement of adequate data protection rules along with considering the issue of market concentration, given that some social media companies reach more people than any traditional mass media company ever has.

Outline of the study

Chapter 1 overviews the usage and definitions of the various terms referring to what is often called ‘fake news’. To develop an appropriate definition and terminology, resources in the field of communication studies, political science, internet policy, law and behavioural studies have been critically assessed. Section 1.3 describes and assesses the most important disinformation actions that have happened in the past four years. The richly documented collection offers an informative reading of both foreign and domestic manipulation campaigns (the methodology of data collection is set out below). Even more data on the disinformation campaigns has been added in Annex 2.

Chapter 2 scrutinises the impact of disinformation and propaganda on democracy, fundamental rights, the rule of law and EU integration. Section 2.1 explores which features of the new media environment are relevant in the context of disinformation and propaganda and how these affect the public discourse. The flaws in the democratic public discourse are causing distortions in the democratic process (2.2), passing on to the democratic legitimacy of the EU (2.3). Section 2.4 studies the way in which the malpractices directly impact human rights.

Chapter 3 explores the limits of existing legal regulation, first of all the extent and content of the responsibility of social media from a normative perspective. Legal experts are still divided on the issue of whether social media platforms should be liable for third-party content. This study takes the progressive view of content-neutral policy in this respect, which is explained and supported with case studies and international legislative examples. But first and most importantly, the legal category needs to be formulated for platform providers, which are a special new breed of business enterprises that emerged only after the web enabled P2P technology. Section 3.2 discusses the fundamental principles of freedom of expression from the angle of whether they allow any further legislative interference and restriction of expression with a view to tackling disinformation and propaganda.

Chapter 4 describes and assesses the latest legislative attempts, whether they are in line with these standards, and how successfully they can counteract disinformation and propaganda. The potential of the European and national legislative instruments relating to elections and election campaigns are assessed, and their limits established. The possibilities of international legal reactions to foreign-generated informational interferences have been considered and analysed, as presented in section 5.2.

Chapter 5 also provides a look at future disruptive technologies. As these become significantly more accessible in the near future, more actors will utilise them and people will be even more exposed to their effects. For example, the Internet of Things will determine the everyday lives of people – and expose people to being tracked even more. Virtual reality has already been used in political campaigns. This immersive experience is less subjected to rational thinking and amplifies the effects of any potential manipulation. Innovative services are driven by the monetisation of personal data, which are provided by platform providers. This supply-and-demand chain needs to be reconsidered by decision makers and industry actors, in order to find a new model that respects human rights.

Based on the conclusions in **chapter 6**, the study offers a coherent set of policy recommendations (**chapter 7**). These are centred around two large sections: (i) how to strengthen democratic resilience, and (ii) media policy. For the fairness of the coming EP elections, imminent actions are recommended.

Detailed recommendations address the obligations of platform providers, which should relate to the architecture and administration of their platforms, rather than policing content. Besides platform operators' obligation to respect and enforce existing legal rules in their platforms – such as rules relating to advertising and data protection – regard ought to be taken of the platform providers' market share.

The aim of the policy recommendations is to ensure that platforms are safe and transparent, while the right balance between freedom of expression, protection of personal data and the right to receive information is maintained. In search of the balance, the interests of democratic stability should be kept in mind. In this context, citizens' human rights should prevail even if they conflict with corporate financial interests.

Methodology

Desk research

Several reports, research papers, articles and legal instruments emerged even during the writing of this paper, and Facebook has been continually adjusting its user policy to achieve positive changes to its reputation. The authors have made their best effort to keep track of these changes during the months of research, but they cannot guarantee that everything has been covered.

The research has been completed with authoritative academic books discussing media and communication theory, democracy and legal theory. The fact-finding chapters of the study are based on research reports, as well as other fact-based documents on instances of disinformation and the events during election campaigns published in the media and by monitoring organisations. Chapters assessing promising practices, existing laws, regulations and policies are based on legislative instruments, draft legislative instruments, policy papers and self-regulatory codes.

Disinformation actions

The research for describing and assessing the most important disinformation actions has relied mostly on four types of sources: news reports, official documents, communication from stakeholders and scholarly articles. Although there is much scholarly interest in the field, as the discussed events have happened recently, there are few peer-reviewed academic works published on them. For the selection and exploration of the cases we have had to rely heavily on journalistic investigation. New information related to the disinformation actions is published almost daily in the media; this has been included to ensure that the study is up to date. To reduce errors due to misreporting, all information included in the study have been cross-checked to ensure it was reported as fact by several media outlets, unless otherwise stated. When news reports referenced official communication, the original document was consulted whenever possible. In addition, the authors note the scarcity of quality empirical research into the dissemination of informational manipulation across digital platforms beyond Facebook and Twitter, as well as the lack of data on practices in different EU Member States.

Information manipulation campaigns abound; the most relevant cases have been selected to highlight different aspects. On foreign-influence campaigns, the two most recent elections in the US as well as the Brexit referendum have been picked because they have wide-ranging implications beyond their geographical locations. The 2017 French and German elections are included because they defied expectations by showing resilience to foreign disinformation. A recent campaign against several states originating in Iran is discussed to give an example of foreign meddling originating from a country other than Russia.

A further limitation comes from the fact that disinformation campaigns, by their nature, are covert operations; the originators usually deny involvement or at least offer alternative explanations. Even when investigations have progressed like in the US case, no court conviction declares that it can be attributable either to Russian individuals or companies, let alone the Russian state. Thus, the discussion in section 1.3 relies on assessments by intelligence services, state agencies, and scientific and journalistic investigation which has been published.

Election rules and best practices

When examining the election rules, 15 EU Member States were selected, with the selection criteria seeking to provide a diverse focus group and to ensure a balanced representation in all three aspects used for selecting the focus countries:

- equitable geographical distribution (taking into consideration all macro-regions of the EU);
- economic weight and voting power in EU decision-making processes; and
- potential risk of threatening the rule of law and system of democracy based on election results and societal phenomena.

Overview of the technological trends

The subsection giving an overview of the technological trends (5.1.1) is multidisciplinary and ambiguous due to its nature. It was compiled drawing on the publicly available sources, including scientific publications, online articles, reports and seminal studies. The authors have not attempted to review the patents nor have they had access to proprietary new technology that has not yet been made public. To remedy these limitations, the authors sought consultation from experts in the fields of data analytics, distributed ledger technologies and software engineering. Specifically, Alex Comunian, Gabrielle Pellegrinetti and Giulia del Gamba[13] contributed to the study by providing information about the development of the technologies in question and reviewing the relevant parts of the section. Their critical feedback has been taken into consideration, and necessary changes to the final text of the report have been made. The experts provided information by means of electronic communication.[14]

Interviews

Consultations and interviews have been undertaken to generate and check background knowledge for the study with relevant stakeholders and policy makers, in addition to experts on information technology. The European institutions contacted were DG Just, DG Connect, the EDPS and the EEAS. Within the IT sector Microsoft was contacted. (Facebook and Google were inaccessible despite repeated attempts.)

[13] The experts contributed in their private capacity.

[14] These three sections have been directly informed by 20 academic papers and books, 34 reports produced by the academic institutions, think tanks and regulatory authorities, 79 online articles from reputable sources, 6 policy documents by the international organizations and human rights mechanisms, and 14 piece of relates material, including statistical data and websites of digital platforms.

1. STATE OF PLAY

KEY FINDINGS

- There is an emerging consensus among public policy actors against using the term 'fake news' and in favour of using the term 'disinformation' to describe what is generally understood as false or misleading information produced and disseminated to intentionally cause public harm or for profit.
- Disinformation actions have happened all over the globe and they are diverse in scope and effect, but important details of recent disinformation actions have still not completely been revealed.
- Among the disinformation actions, some practices have been conceived as dangerous because of their divisive impact, rather than their misleading content. Any long-term policy initiative should address both disinformation and aggressive informational practices meant to purposefully cause social harm.
- Awareness of the threat appears to have a protective effect: in societies that have exercised great caution, no substantive harm could be observed (e.g. in Germany and France).
- The interests of the technology service providers (social media platforms and digital advertisers) and the actors behind the disinformation campaigns are to some extent aligned. Both are interested in capturing the scarce resource of the information economy – users' attention.
- Individuals are not only targets but also mediums for disinformation distribution, which has a participative nature.
- Instant messaging services offering group chat services are even less transparent and less controllable than social media. Their growing popularity carries the risk that disinformation and propaganda are submerged and less apparent to researchers and policy makers.
- Paid advertising reaches more of the target audience and broadens the overall reach of the messages, while organic content posted on the social media network can only reach those users who see the posts in their own newsfeed, or in feeds of friends, groups and communities engaging with this information.
- The use of micro-targeting and artificial dissemination techniques like social bots have been key in disseminating disinformation and changing the rules of the network.
- Some of the government-owned or government-sponsored media outlets, including media outlets popular among national minorities in foreign countries, are also known to be important vehicles of state-led disinformation campaigns, which raises the issue of mutual confidence between EU Member States.
- Advertising tools rely on the collection and analysis of personal data about the target audience. The data are collected directly from the individuals or observed from user activity online using tracking technologies, or obtained from third parties.

1.1 Summarising definitional challenges

Below we provide an assessment on the terms commonly used in the literature and parlance of institutions relating to the phenomena. In this study, we use the terms 'disinformation' and 'propaganda', by which we understand all the expressions shown in Table 3 (subsection 1.1.2).

1.1.1 Fake news

The term 'fake news' only emerged around the end of the 19th century due to the relative novelty of the word 'fake'. According to the website of the *Merriam–Webster Dictionary*, 'fake' was little used as an adjective prior to the late 18th century, and before that point, the most common collocation in use was 'false news'.[15]

[15] "The Real Story of 'Fake News'", https://www.merriam-webster.com/words-at-play/the-real-story-of-fake-news.

The data from Google Trends and the Web of Science show that the search for fake news and the number of peer-reviewed papers including the term 'fake news' have grown exponentially since November 2016.[16] The sudden popularity of this term is attributed to the 2016 US presidential elections and the current US President himself.[17]

An authoritative paper by economists Allcott and Gentzkow defines 'fake news' as "news articles that are intentionally and verifiably false, and could mislead readers".[18] The lawyers Klein and Wueller leave out the impact of the information on the reader and employ the following working definition: "the online publication of intentionally or knowingly false statements of fact".[19] Media scholars Bakir and McStay propose to define fake news in a disjunctive way: "as either wholly false or containing deliberately misleading elements incorporated within its content or context".[20] The philosopher Rini has offered one of the most extensive definitions so far:

> A fake news story is one that purports to describe events in the real world, typically by mimicking the conventions of traditional media reportage, yet is known by its creators to be significantly false, and is transmitted with the two goals of being widely re-transmitted and of deceiving at least some of its audience.[21]

The philosopher Gelfert made an attempt to build on the previous definition and suggests defining fake news as "the deliberate presentation of (typically) false or misleading claims as news, where the claims are misleading by design".[22] **Overall, as an extensive study by Tandoc et al. found, academic articles between 2003 and 2017 used the term 'fake news' to refer to a range of different phenomena including news satire, news parody, fabrication, manipulation, advertising and propaganda.[23]** Some scholars exploring the phenomena of fake news avoid defining it altogether or use it interchangeably with 'disinformation' or 'propaganda'.

There is a notable lack of consistency among human rights organisations using the term 'fake news'. Two resolutions by the Parliamentary Assembly of the Council of Europe (PACE) referring to fake news do not attempt to define the phenomena. In Resolution 2212 (2018), PACE considers "fake news", "propaganda" and "disinformation" as different forms of manipulation,[24] whereas in Resolution 2217 (2018), "fake news" is identified as a form of "mass disinformation campaigns", which constitute a technique of a "hybrid war".[25] The Joint Declaration by the special rapporteurs on freedom of expression acknowledges fake news in the title of the document, but talks exclusively about "disinformation" and "propaganda" throughout the main body of the declaration.[26]

Within the legislative domain, the concept of fake news is even more ambiguous, as evidenced by the recent debates around the efforts to introduce national 'anti-fake news' laws. In **France**, the law against the "manipulation of information" attempts to define fake news as "any allegation of a fact that is inaccurate or

[16] Ibid.

[17] Ibid.

[18] Allcott H. and Gentzkow M., Social media and fake news in the 2016 election. Stanford University, Journal of Economic Perspectives 31(2): 211-236, 2017.

[19] Klein, D. and Wueller J, Fake news: a legal perspective. Journal of Internet Law 20(10): 5-13, 2017.

[20] Vian B. and McStay A. Fake news and the economy of emotions: problems, causes, solutions. Digital Journalism 6(2): 154-175, 2018.

[21] Rini, R., Fake news and partisan epistemology. Kennedy Institute of Ethics Journal 27(2): 43-64, 2017.

[22] Gelfert A., Fake news: a definition. Informal Logic 38 (1):84-117, 2018, pp. 85-86.

[23] Tandoc E. et al., Defining "fake news" a typology of scholarly definitions. Digital Journalism 6(2): 137-153, 2018, p. 137.

[24] Resolution 2212 (2018) of the Parliamentary Assembly of the Council of Europe on the Protection of editorial integrity, paras. 8.7 and 9.5.

[25] Resolution 2217 (2018) of the Parliamentary Assembly of the Council of Europe on the Legal challenges related to hybrid war and human rights obligations, para. 3

[26] Joint declaration on freedom of expression and "fake news", disinformation and propaganda, 2017, https://www.osce.org/fom/302796

misleading", which is likely to "distort the fairness of the election", if propagation on the internet was made "deliberately" and "in an artificial or automatized and massive way".[27] In **Italy**, the bill proposed but not adopted in 2017 defined fake news as "false, exaggerated, or biased" news reports online.[28] In 2018, the Italian Ministry of Interior further aimed to combat fake news by introducing a system of reporting "manifestly unfounded and biased news, or openly defamatory content".[29] The **German** NetzDG law of 2017 did not define "fake news"[30] but served as an inspiration for some anti-fake news legislation across the world.[31]

As pointed out by Martens et al., there is no consensus on the definition of 'fake news'.[32] The definitions discussed above tend to be constructed, to a varying degree, around four dimensions: (i) type of information; (ii) falsity of information; (iii) intention of the author; and (iv) consequences of dissemination of information, including personal (perception of the receiver) and societal effects (disruption of democratic processes). With respect to the first two dimensions – type and falsity of information – narrow definitions of 'fake news' tend to focus on verifiably false news reports, whereas broader definitions also include any misleading or distorted information.[33] The latter better reflects the reality of manipulative stories – many of them are not wholly false, but mix deliberate falsehoods with well-known truths by selectively presenting partial truths, employing a false context or manipulating images alongside verified news stories.[34]

1.1.2 Misinformation and disinformation

Scholars have argued that the term 'fake news' is woefully inadequate to describe the complex phenomena of mis- and disinformation.[35] Similar problems with the term 'fake news' were identified by the European Commission's High-Level Expert Group on Fake News and Online Disinformation (HLEG). In its final report, the HLEG found the term 'fake news' to be "inadequate to capture the complex problem of disinformation" which involves not necessarily "fake", but fabricated content and practices going beyond the conventional "news".[36] Also, it is found to be misleading due to it being appropriated by some politicians to dismiss any content they regard as disagreeable.[37] The HLEG preferred the word 'disinformation' instead and used the following definition

[27] Hochmann, T., Shedding light or shooting in the dark – how to define fake news, 5 September 2018, https://verfassungsblog.de/shedding-light-or-shooting-in-the-dark-how-to-define-fake-news

[28] Letter of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression No OL ITA 1/2018, 20 March 2018, p. 1, https://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-ITA-1-2018.pdf

[29] European Center for Press and Media Freedom, Tackling fake news, the Italian way, 22 May 2018, https://www.rcmediafreedom.eu/Tools/Legal-Resources/Tackling-fake-news-the-Italian-way

[30] Human Rights Watch, German: flawed social media law, 14 February 2018, https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law

[31] See e.g., a proposed legislation in Philippines. Poynter, A guide to anti-misinformation actions around the world, https://www.poynter.org/news/guide-anti-misinformation-actions-around-world

[32] Martens, B. et al., The digital transformation of news media and the rise of disinformation and fake news. JRC Digital Economy Working Paper 2018-02, p. 5, https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf. For the perceptions of the public on the definition of 'fake news', consult: Nielsen, R. and Graves, L., "News you don't believe": Audience perspectives on fake news. Factsheet: October 2017, http://reutersinstitute.politics.ox.ac.uk/sites/default/files/2017-10/Nielsen%26Graves_factsheet_1710v3_FINAL_download.pdf

[33] Martens, B. et al., The digital transformation of news media and the rise of disinformation and fake news. JRC Digital Economy Working Paper 2018-02, pp. 10-11

[34] Gelfert A., Fake news: a definition. Informal Logic 38 (1):84-117, 2018, p. 100. See also a typology suggested by Wardle, C. and Derakhshan H. in Information disorder: Toward an interdisciplinary framework, Council of Europe, 2017, p. 17, https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

[35] Wardle C. and Derakhshan H., Information disorder: definitions in "Understanding and addressing the disinformation ecosystem", 2017, p. 6

[36] High level Group on fake news and online disinformation, A multi-dimensional approach to disinformation, 2018, p. 10, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50271; UK House of Commons Digital, Culture, Media and Sport Committee, Disinformation and 'fake news': Interim Report, 24 July 2018, p. 8, https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/363.pdf

[37] High level Group on fake news and online disinformation, A multi-dimensional approach to disinformation, 2018, p. 10

throughout the report: "all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit".[38]

The definition of the HLEG resembles a broad definition of fake news already offered by some scholars as discussed above. **However, being a less contentious and less politically-charged term,[39] 'disinformation' is increasingly favoured by the national and supranational institutions and bodies, including the European Commission,[40] the European Council,[41] the EEAS,[42] the UK House of Commons Digital, Culture, Media and Sport Committee,[43] and the Danish government.[44]**

'Disinformation' is often used alongside 'misinformation', but their usage also suffers from a lack of consistency. Some authors, including Losee and Fox, use it interchangeably, whereas Zhou and Zhang consider one a variation of the other. [45] While the *Oxford* online dictionary of the Oxford University Press and the *Collins English Dictionary* list 'misinformation' as a synonym for 'disinformation', the *Merriam–Webster* and *Oxford Living Dictionaries* make subtle distinctions between the two definitions. The EU's interinstitutional terminology database IATE (Inter-Active Terminology for Europe) specifically notes that disinformation should not be confused with misinformation, defined in IATE as "information which is wrong or misleading but not deliberately so".[46] Yet, the European Parliament resolutions use 'misinformation' and 'disinformation' interchangeably.[47]

In this regard, Wardle and Derakhshan draw a helpful distinction between disinformation, misinformation and mal-information based on the level of facticity and intent to harm.

[38] Ibid., p. 3

[39] Nielsen, R. and Graves, L., "News you don't believe": Audience perspectives on fake news. Factsheet: October 2017, p. 5.

[40] Communication from the Commission on Tackling Online Disinformation: A European Approach, COM (2018) 236 final (26 April 2018)

[41] European Parliament, Understanding propaganda and disinformation, November 2015, http://www.europarl.europa.eu/RegData/etudes/ATAG/2015/571332/EPRS_ATA(2015)571332_EN.pdf

[42] Ibid.

[43] UK House of Commons Digital, Culture, Media and Sport Committee, recommended the UK government to reject the term 'fake news' and to put forward an agreed definition of 'disinformation' and 'misinformation' instead.

[44] Ministry of the Foreign Affairs of Denmark, Strengthened safeguards against foreign influence on Danish elections and democracy, 7 September 2018, http://um.dk/en/news/NewsDisplayPage/?newsID=1DF5ADBB-D1DF-402B-B9AC-57FD4485FFA4. The plan released by the Danish government also uses a term 'influence campaigns' to capture the attempts of the foreign countries to influence national elections and referendums.

[45] Karlova, N. and Lee, J., Notes from the underground city of disinformation: A conceptual investigation. Proceedings of the American Society for Information Science and Technology 48(1), 2012, p. 1.

[46] European Parliament, Understanding propaganda and disinformation, November 2015.

[47] Ibid.

Table 3: Types of information disorders

	Definition	Example
Misinformation	When **false information** is shared, but **no harm** is meant	During the 2016 US presidential elections, a tweet about a 'rigged' voting machine in Philadelphia was shared more than 11 000 times. It was later established that the original tweet was a mistake made by a voter who had failed to follow the instructions exhibited on the voting machine.[48]
Disinformation	When **false information** is knowingly shared **to cause harm**	During the 2017 French presidential elections, a duplicate version of the Belgian newspaper *Le Soir* was created, with a false article claiming that Emmanuel Macron was being funded by Saudi Arabia.[49]
Mal-information	When **genuine information** is shared to **cause harm**	Examples include intentional leakage of a politician's private emails, as happened during the presidential elections in France.[50]

Source: Wardle, C. and Derakhshan H. in Information disorder: Toward an interdisciplinary framework (2017).

In the suggested typology, the determination about the nature of information is not objective, but relative – the information shared with malicious intent is recognised as disinformation, whereas the same information shared by a poorly informed party is viewed as misinformation. This subtle but important distinction may contribute to a better understanding of whether to assign responsibility to those involved in dissemination of disinformation.

1.1.3 Propaganda

Propaganda has been studied from the perspective of history, journalism, political science, sociology and psychology, as well as from the interdisciplinary perspective. As a result, unsurprisingly, different definitions of 'propaganda' have emerged. **According to Martin, of 26 definitions examined, "all agree that propaganda is the art of influencing, manipulating, controlling, promoting, changing, inducing, or securing the acceptance of opinions, attitudes, action, or behaviour". The opinions of the scholars differ as to whether 'propaganda' must be systematic and ordered.** For example, Albig, Bird and Doob consider that it must be organised, and Doob defines it in the following way: "Propaganda is a systematic attempt by an interested individual (or individuals) to control the attitudes of groups of individuals through the use of suggestion and, consequently, to control their actions."[51]

This definition does not carry any political connotation and can be applied to a variety of settings. For example, Carrey considered commercial advertising and public relations to be forms of propaganda.[52] Political advertising campaigns, especially active during the electoral period, have also been considered a form of propaganda, as well as the attempts of ideological movements to influence and recruit followers,[53] or deliberate actions by a third-country government to influence the democratic processes in neighbouring states. A broad definition of 'propaganda' is used in the *Oxford English Dictionary*: "systematic propagation of information or ideas by an interested party, esp. in a tendentious way in order to encourage or instil a particular attitude or response. Also,

[48] Electionland, Viral 'Rigged' Voting Machine Video Actually User Error, 8 November 2016, https://projects.propublica.org/electionland/pennsylvania/viral-rigged-voting-machine-actually-user-error/; https://www.nytimes.com/2016/11/09/us/politics/debunk-fake-news-election-day.html

[49] Wardle, C. and Derakhshan H. in Information disorder: Toward an interdisciplinary framework, Council of Europe, 2017, p. 21.

[50] Ibid.

[51] Martin, J., Definition of propaganda in "International Propaganda: Its Legal and Diplomatic Control". University of Minnesota Press, 1958, p. 10.

[52] Jowett, G. and O'Donnell, V., Propaganda & Persuasion. SAGE Publications, 12 Apr 20111, p. 6.

[53] Ibid., p. 8.

the ideas, doctrines, etc., disseminated thus; the vehicle of such propagation."[54] On the other hand, Tandoc et al. do not consider that propaganda must be systematic, but they limit the definition of the phenomenon to a political context and 'news'.[55]

The problem with the broad definitions of propaganda used by the scholars is that they equalise persuasion with manipulation. The definitions discussed above do not leave much room for delineation between legitimate political campaigning (protected political speech) and highly targeted, emotive political advertising. The Joint Declaration by the special rapporteurs on freedom of expression[56] is clearer in this respect, as from the outset it focuses on propaganda "designed and implemented so as to mislead a population, as well as to interfere with the public's right to know and the right of individuals to seek and receive, as well as to impart, information and ideas of all kinds".[57] The special rapporteurs focus on the intent of the propaganda, rather than its content, thus excluding from the scope of their intervention legitimate forms of persuasion.

NATO and the European Parliament frequently use 'propaganda' in their public communication, sometimes interchangeably with 'disinformation', and sometimes together with it. In this context, 'propaganda' is predominantly used to describe strategic information campaigns orchestrated by the Kremlin with the aim of influencing democratic processes in Ukraine and beyond.[58]

1.1.4 A missing link

The terms 'fake news', 'disinformation', 'misinformation' and 'propaganda' have been defined and redefined in different contexts. The summary of the existing research on the subject matter of the study demonstrates largely inconsistent and sometimes conflicting usage of nowadays popular terminology. The existing cacophony of definitions has become an obstacle to scoping the phenomena, which inadvertently has negatively impacted the ability to design an effective response to address it.

At the same time, there is an emerging consensus among public policy actors against using the term 'fake news' and in favour of using the term 'disinformation' to describe what is generally understood as false or misleading information produced and disseminated to intentionally cause public harm or for profit. Arguably, it can be considered an adequate term to describe the phenomena in question, with one caveat. The definition focuses on the dichotomy between the truth and falsehood, whereas analysis of the real-life cases of manipulation shows that the actors behind them do not necessarily position themselves relative to the truth but may simply be trying to produce the dividing effect.[59] In other words, dissemination of misleading information is one of the elements used by the state or domestic actors in their overall strategic effort.

[54] Oxford English Dictionary, 2nd ed., 1989.

[55] Tandoc E. et al., Defining "fake news" a typology of scholarly definitions. Digital Journalism 6(2): 137-153, 2018, p. 146.

[56] Joint declaration on freedom of expression and "fake news", disinformation and propaganda, 2017, https://www.osce.org/fom/302796

[57] Ibid, p. 1.

[58] European Parliament, Understanding propaganda and disinformation, November 2015.

[59] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 20, https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf. As explained in the study, the objective of the propaganda carried out by Kremlin is no longer to convert people to an ideology, but to weaken and divide the societies. Therefore, the actions have supported both far right and far left movement and endorsed contradictory narratives (p. 53). Also see analysis of Iranian information manipulation actions against the UK and the US on Facebook: Atlantic Council's Digital Forensic Research Lab, #TrollTracker: Facebook Uncovers Iranian Influence Operation, 26 October 2018, https://medium.com/dfrlab/trolltracker-facebook-uncovers-iranian-influence-operation-d21c73cd71be

Table 4: Examples of manipulative practices

Hypothetical scenario	Nature of claims	Objective	Manipulative practice
Domestic political actors present voters with false information about same-sex relationships in order to convince them to vote against same-sex partnerships in a national referendum.	False facts	Influence opinion and voting behaviour	Disinformation
A foreign state actor micro-targets local voters with two series of posts on social media. National minorities are exposed to the messages calling on them to 'assert their power', while the rest see posts emphasising the importance of 'a dominant nation'.	Not susceptible to proof	Sow distrust and polarise society	Manipulating through aggressive informational practice

Source: Authors.

To fully understand the scope of the problem, there is a need to acknowledge emerging practices that are dangerous because of their potential for divisiveness, rather than the misleading content. These aggressive practices may be employed to promote false or manipulated content, as well as genuine information or value judgment-like statements that cannot be objectively verified. Examples of such aggressive practices include divisive political advertising, highly targeted political advertising aimed at exploiting personal vulnerabilities, fears and beliefs, hacking and leaking private information, and verbal abuse perpetrated by the hired 'trolls'. Irrespective of the information in question, these practices are harmful due to their privacy-invasive nature, a potential to 'nudge' individuals into certain thinking or behaviour,[60] and to escalate societal polarisation. Instances of divisive advertising were found among Russian-bought ads during the 2016 US presidential elections:[61] "The day after the election, an event called for an anti-Trump rally in Union Square even as another ad called for Trump supporters to rally outside Trump Tower. In another instance, the ads promoted both a pro-Beyoncé and anti-Beyoncé event[62] in New York City."

The developing data-driven technologies discussed further in the study may potentially result in a proliferation of these practices and exacerbation of their impact. **Therefore, it is important that any long-term policy initiatives take into consideration the reality of informational manipulation,[63] addressing both false or misleading information (disinformation) *and* aggressive informational practices, meant to purposefully cause harm.**

1.2 Production, distribution and amplification of informational manipulation

1.2.1 Origin and nature of manipulative campaigns

An informational manipulation campaign can be launched by a state or a non-state actor or their proxies. There are no precise estimates as to what proportion of disinformation in the EU originates from foreign actors (e.g.

[60] Borgesius, F. et al., Online Political Microtargeting: Promises and Threats for Democracy. Utrecht Law Review, 14(1): 82–96, p. 87.

[61] What we can learn from the 3,500 Russian Facebook ads meant to stir up U.S. politics, 10 May 2018, https://techcrunch.com/2018/05/10/russian-facebook-ads-house-intelligence-full-list/

[62] As explained by the Guardian, 'In one particularly brazen example, ads were run promoting both a "Pro-Beyonce Protest Rally" and an "Anti-Beyonce Protest Rally" scheduled for the same time and place following the controversy over the artist's performance at the 2016 Super Bowl. The pro-Beyoncé ad was targeted at users designated as having African American behaviors. The anti-Beyoncé ad was targeted narrowly at people who had studied to become a police officer or whose job title matched a list of law enforcement or military titles […]'. See: #BlueLivesMatter and Beyoncé: Russian Facebook ads hit hot-button US issues, 10 May 2018, https://www.theguardian.com/us-news/2018/may/10/russia-facebook-ads-us-elections-congress. Other examples include psychometric targeting used in the 2016 US Presidential Elections, see: 'I made Steve Bannon's psychological warfare tool': meet the data war whistleblower**,** 18 March 2018, https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump

[63] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 20, https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf. p. 21.

Kremlin, Beijing, Iran[64]) and what is authored domestically (e.g. by government representatives, political parties, populist politicians, interest groups, profit-seeking individuals, independent trolls[65] or conspiracy theorists[66]). For instance, a piece by the *New York Times* has asserted that based on Facebook's data, the majority of the accounts behind false and misleading content in the US were domestic.[67] No comparable conclusions can be drawn about the EU in the absence of reliable data, but there is evidence of domestic actors, for instance referendum campaigners, using digital platforms to advertise factually misleading messages as part of the campaign.[68]

The tactics of Kremlin informational manipulation are so far the most researched ones, also because their manifestation during recent democratic processes in the EU and overseas has been particularly prominent.[69] Its distinct features are organisation and a wide network of allies aiding in the distribution of the messages. RAND Corporation has identified four types of such allies: government bodies (e.g. ministries, embassies); fake NGOs (financed or working closely with the state), other seemingly unrelated organisations that in reality are close to the governing authorities (e.g. motorcycle clubs) and religious, political and economic relays (political parties of other sovereign states, religious groups).[70] These allies can act as both initiators of a campaign and distribution mediums. For example, an embassy can 'produce' a report making false claims or post a report already produced by a false NGO on its website.

The narratives and formats of disinformation messages are usually dictated by the context and the target audience of the campaign.

Table 5: Key elements of informational manipulation campaigns

Context
Disinformation is not time- or place-restricted, but is **likely to intensify before/during significant democratic decision-making processes** such as referenda (e.g. the Dutch referendum on the Association Agreement between Ukraine and the EU, the 2016 UK referendum) and elections (e.g. the French presidential elections).[71] Referendums provide particularly fertile ground for disinformation as by nature they address issues society is divided over. The virality of disinformation is also more easily achieved at the

[64] Ibid. On Iran, see Atlantic Council's Digital Forensic Research Lab, #TrollTracker: Facebook Uncovers Iranian Influence Operation, 26 October 2018, https://medium.com/dfrlab/trolltracker-facebook-uncovers-iranian-influence-operation-d21c73cd71be

[65] 'Independent trolls' (not paid trolls) are not analysed in this report in detail. For more information see: Tucker, J. et al., Social media, political polarization and political disinformation: a review of the scientific literature. Hewlett Foundation, March 2018, p. 22, https://hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf

[66] Conspiracy theorists are not analysed in this report in detail, but as noted in the report 'Information Manipulation: A Challenge for Our Democracies' (cited above), 'they pose a particular difficulty because they are highly resistant to debunking, especially if attempted by the State. As conspiracy theories hold that certain people have disproportionate power with which to conceal their actions, these attempts can become absorbed into the narrative of the plot', pp. 34-35.

[67] Facebook Tackles Rising Threat: Americans Aping Russian Schemes to Deceive, 11 October 2018, https://www.nytimes.com/2018/10/11/technology/fake-news-online-disinformation.html; also see Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 47

[68] For example: Vote Leave's targeted Brexit ads released by Facebook, 26 July 2018, https://www.bbc.com/news/uk-politics-44966969; Italy's vote: Fake claims attempt to influence election, 3 March 2018, https://www.bbc.com/news/world-europe-43214136

[69] According to the National Endowment for Democracy, 'Even a partial list of elections where Russian-produced or -supported disinformation has featured includes the French, German, and American elections in 2016 and 2017; the 2018 Czech presidential election; and the 2017 vote on Catalonian secession from Spain.' (Issue in brief: how disinformation impacts politics and publics, 29 May 2018, https://www.ned.org/issue-brief-how-disinformation-impacts-politics-and-publics/).

[70] Linda Robinson et al, Modern Political Warfare. Santa Monica, CA: RAND Corporation, 2018, p. 56.

[71] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 39.

beginning of a **high-profile event or crisis** when people are paying attention and while trusted authorities have not yet provided an authoritative narrative to explain the situation.[72]

Audience

Geographical location, nationality, political beliefs, socioeconomic profile, age and other personal characteristics are taken into account when calibrating the message to the audience. For example, some of the Kremlin's disinformation efforts target Russian speakers and underprivileged communities abroad in order to feed on the frustration of these groups.[73] In the course of electoral campaigns, domestic actors may choose to target opposition supporters with false information about the voting requirements or reports about violence or long queues in the polling stations (voter suppression).

Narrative

The overall narratives do not materialise out of thin air, but rather exploit the pre-existing tensions in society (e.g. migration, crime, LGBT and reproductive rights), which appear to them to be more credible.[74] Some narratives are framed **positively** (e.g. supporting a political candidate) and others **negatively** (e.g. attacking the opposition's candidate) or even **inflammatory** (meant to inspire strong emotions like "fear, disgust and surprise").[75] And some messages used in informational manipulation campaigns may even seem benign, usually because they are designed to distract. For example, a tactic of Beijing's '50 cent party' is to post emotive comments online in order to provoke audience reaction against the individual commentator and divert attention from criticism of the government.[76]

Format

The informational manipulation campaigns include **messages in different formats** to reach the target audiences, including news pieces, blog posts, comments on articles or under the social media posts, 'memes', fake profiles of influencers, TV reports, documentaries, YouTube videos, Facebook event pages and hashtags.[77] Their content may be entirely fabricated or slightly manipulated, they may feature false connections (e.g. subtitles do not support the original footage) or be entirely genuine but provided in the wrong context (e.g. historical information featured as news). In some cases, as explained above, the format of the message may not include verifiable facts at all, but aggressively promote value judgment-like statements.

Source: Authors.

1.2.2 Digital amplification mechanisms

Informational manipulation campaigns are based on a set of coordinated yet dispersed activities that give an impression of a spontaneous action. They are launched on various mediums, both online and offline at different moments in time and rely on a combination of 'natural reach' (enabled by humans and traditional media) and automation (enabled by bots and advertising). This makes it difficult to trace the origins of the story and determine the culpable parties.

[72] Tucker, J. et al., Social media, political polarization and political disinformation: a review of the scientific literature. Hewlett Foundation, March 2018, p. 46, https://hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf

[73] Ibid., p. 76

[74] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 77

[75] Vosoughi, S. et al., The spread of true and false news online, Science 359(6380): 1146-1151, 9 March 2018, p. 1146.

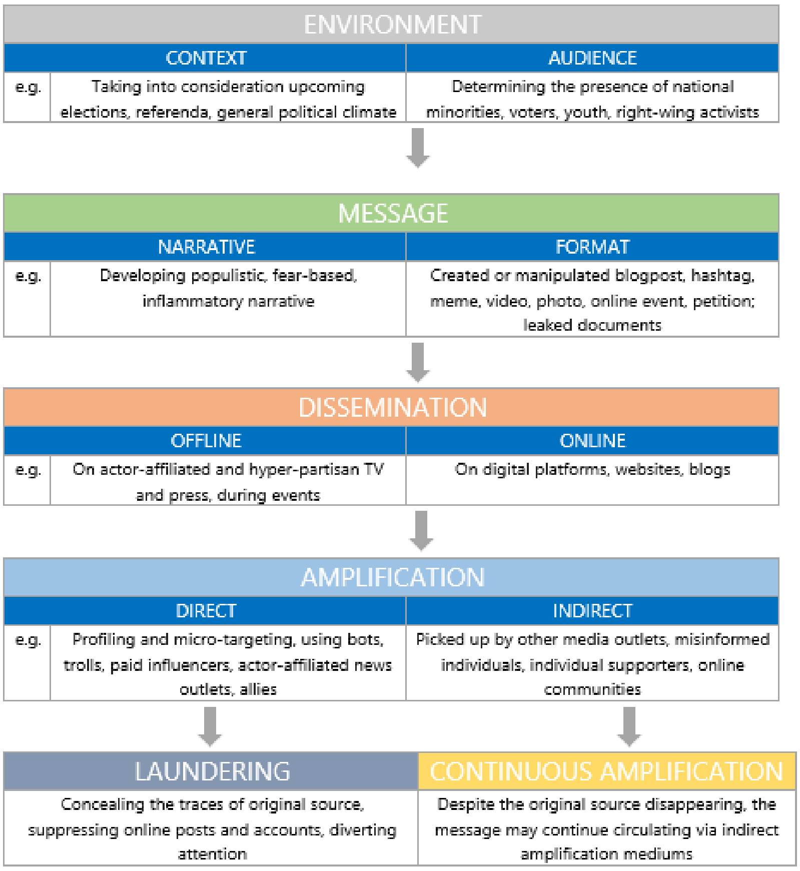
[76] Bradshaw, S., and Howard, P., Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation. University of Oxford: Working Paper, 2017(12), pp 8-9.

[77] For a comprehensive overview see: UNESCO, Journalist, Fake News and Disinformation, 2018, pp. 48-50, https://en.unesco.org/fightfakenews. Also see Brodnig, I., 7 types of misinformation in the German election, 2 November 2017, https://firstdraftnews.org/7-types-german-election/.

The modern-day informational manipulation campaigns rely on the digital tools widely used by the industry. **In fact, the tools and systems behind online informational manipulation (e.g. real-time bidding)[78] are no different from those employed to convince us to buy a fridge or a pair of new shoes.**

As such, the **interests of the technology providers (online platforms, social networks and digital advertisers) and the actors behind the disinformation campaign are to some extent aligned. Both are interested in capturing the scarce resource of the information economy – users' attention**[79] – and holding it for as long as possible. Actors achieve it by promoting sensational, controversial and engaging (manipulated) content, which in turn drives engagement with the platform and delivers ad revenue to the players in the digital ecosystem.[80] Annex 1 provides an overview of the digital platforms currently leveraged, to a higher or a lesser extent, to magnify the effects of the informational manipulation. Figure 2 depicts the interaction between digital and non-digital dissemination and amplification mechanisms.

Figure 2: Generic elements of an informational manipulation campaign


ENVIRONMENT
CONTEXT
AUDIENCE
e.g.
Taking into consideration upcoming elections, referenda, general political climate
Determining the presence of national minorities, voters, youth, right-wing activists
MESSAGE
NARRATIVE
FORMAT
e.g.
Developing populistic, fear-based, inflammatory narrative
Created or manipulated blogpost, hashtag, meme, video, photo, online event, petition; leaked documents
DISSEMINATION
OFFLINE
ONLINE
e.g.
On actor-affiliated and hyper-partisan TV and press, during events
On digital platforms, websites, blogs
AMPLIFICATION
DIRECT
INDIRECT
e.g.
Profiling and micro-targeting, using bots, trolls, paid influencers, actor-affiliated news outlets, allies
Picked up by other media outlets, misinformed individuals, individual supporters, online communities
LAUNDERING
CONTINUOUS AMPLIFICATION
Concealing the traces of original source, suppressing online posts and accounts, diverting attention
Despite the original source disappearing, the message may continue circulating via indirect amplification mediums

Source: Authors.

[78] Olejnik, L., Technological soft influence on elections, 10 August 2016, https://blog.lukaszolejnik.com/soft-influence-on-societies/

[79] The business model of social media is based on the attention economy. The more people use social media, the more attention social media can sell to advertisers - and the more data about the users' behaviour they can collect (The Economist, How the World Was Trolled (November 4-10, 2017), Vol. 425, No 9065, pp. 21-24).

[80] Ghosh D. and Scott B., #Digitaldeceit. The technologies behind precision propaganda and the Internet. Harvard Kennedy School, January 2018, pp. 3-4.

1.2.3 Digital platforms

While the use of online social media is rising sharply (e.g. the number of users on Facebook grew from 500 000 in 2010 to nearly 2 billion in mid-2017[81]), and the trust in mass media continues to decline, social networks have become an important source of political news and information in general.[82]

According to Allcott and Gentzkow, **social media** is also the most attractive vehicle for disinformation.[83] Their research showed that the largest share (41.8 %) of traffic to disinformation sites comes from social networks, while legitimate news sites are most reached by direct browsing (48.7 %), and social media there plays only a minor role (10.1 %).[84]

In contrast to the social media networks, the role of other digital platforms, such as **search engines** and **messaging applications**, is less thoroughly researched and thus less understood. Investigations by the *Guardian* have reported Google Search results on a particular political topic being dominated by the blogs promoting similar extreme viewpoints and pushing credible news sources out of the first page.[85] Google later acknowledged grappling with people who 'try to game the system' in order to bolster 'low quality' content and 'fake news'.[86] The characteristics of different digital platforms and their use in disinformation campaigns are discussed in more detail in Annex 1.

The actors behind the disinformation campaigns can introduce the misleading content onto the platforms and make use of a full suite of services available therein to amplify their messages.[87]

1.2.3.1 Advertising tools

According to Ghosh and Scott, "political disinformation succeeds because it follows the structural logic, benefits from the products, and perfects the strategies of the broader digital advertising market".[88] Organic content posted on a social media network can only reach those users who follow the feed of the disinformation provider or see the posts in feeds of friends, groups and communities that engage with this information. Paid advertising reaches more of the target audience and broadens the overall reach of the messages. As mentioned above and explained in more detail in Annex 1, many digital platforms come with integrated advertising tools that make use of the data already managed by the platforms themselves. In this case, the actors behind disinformation

[81] Constine, J., Facebook now has 2 billion monthly users… and responsibility, 27 June 2017, https://techcrunch.com/2017/06/27/facebook-2-billion-users/

[82] Reuters Institute for the Study of Journalism, Digital News Report 2018, pp. 38-39, http://media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf?x89475. Also see Allcott H. and Gentzkow M., Social media and fake news in the 2016 election. Stanford University, Journal of Economic Perspectives 31(2): 211-236, 2017, p. 212

[83] Allcott H. and Gentzkow M., Social media and fake news in the 2016 election. Stanford University, Journal of Economic Perspectives 31(2): 211-236, 2017, p. 221

[84] Ibid., p. 222. See also 'fake news' trajectory research by Albright J., The #Election2016 Micro-Propaganda Machine, 18 November 2016, https://medium.com/@d1gi/the-election2016-micro-propaganda-machine-383449cc1fba

[85] Catwalladar, C., Google, democracy and the truth about internet search, 4 December 2016, https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook; Solon, O. and Levin, S., How Google's search algorithm spreads false information with a rightwing bias, 16 December 2016, https://www.theguardian.com/technology/2016/dec/16/google-autocomplete-rightwing-bias-algorithm-political-propaganda; Also see: Ghosh D. and Scott B., #Digitaldeceit. The technologies behind precision propaganda and the Internet. Harvard Kennedy School, January 2018, p. 20; Meserole C. and Polyakova A., Disinformation Wars, 25 May 2018, https://foreignpolicy.com/2018/05/25/disinformation-wars/

[86] Gomes, B, Our latest quality improvements for Search, 25 April 2017, https://blog.google/products/search/our-latest-quality-improvements-search/

[87] See also in Conclusions.

[88] Ghosh D. and Scott B., #Digitaldeceit. The technologies behind precision propaganda and the Internet. Harvard Kennedy School, January 2018, p. 4.

campaigns may employ these tools to ensure a more targeted (by location, age group, interests) delivery of the message.

In addition, both commercial and non-commercial advertisers can enrich the datasets by collecting, analysing and integrating additional personal data about the target audience. The data can be collected directly from individuals (for example, canvass data, email, telephone), observed from user activity online using tracking technologies (for example, cookies, tracking pixels) [89] and obtained from third parties (data brokers, online campaigning platforms, data marketing services and social media platforms). The nature of personal data ranges from the name and phone number to location and lifestyle information (purchasing habits, favourite musical bands, etc.). These data are further analysed to better understand the personality and the likely beliefs of the audience (profiling). Profiling leads to the segmentation of the audience into different groups (e.g. likely/unlikely voters, women/men, lower/higher income, parents/non-parents), and different messages are developed for each group (targeting). Variations of messages are tested to identify the format and content that maximises the target audience engagement. The messages could reach the audience in the form of sponsored content across different digital platforms (e.g. ads on Google, Twitter, Facebook, YouTube) or banners on the websites. This process is called behavioural advertising or micro-targeting.

Micro-targeting is operationalised with the help of different players of the digital advertising ecosystem, including advertising platforms and networks, social media management platforms, social media networks and data analytics services. Personalised targeting and divisive advertising played a role in the 2016 US presidential elections, where it was employed by both domestic and foreign actors, and in the 2016 UK referendum. The latter is the subject of current inquiries by several UK authorities.[90] The ongoing investigation by the UK Information Commissioner's Office showed that many UK political parties use third-party digital campaigning platforms (such as NationBuilder), which enables parties to match voters' contact information with the data on Facebook and Twitter. Some of the data are bought from data brokers, and hence the legality of the data raises serious concerns. Facebook, Google, Twitter and Snapchat are most heavily used for political advertising purposes.[91] The use of political micro-targeting was also reported in Italy, Germany and the Netherlands, although to a lesser degree.[92] The situation in other EU Member States is less researched and thus less clear.

1.2.3.2 Bots

Another leg of disinformation automation is social robots, also known as bots, which grow in sophistication as AI advances. Bots are usually understood as automated or semi-automated accounts created to like, share, post or otherwise interact on the social networks.[93] They are intended to behave like humans and have properly filled-out biographies, profile photos taken from the internet, and are connected to other users.[94] Experts estimate that bot traffic now makes up over 60 % of all traffic online – up nearly 20 % from two years prior.[95] Oxford University

[89] For more information see Working Party 29, Opinion 2/2010 on behavioural advertising, 22 June 2010, p. 7.

[90] See e.g. UK House of Commons, Digital, Culture, Media and Sport Committee, Disinformation and 'fake news': Interim Report, 24 July 2018, p. 8, https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/363.pdf

[91] UK Information Commissioner's Office, Democracy disrupted. Personal information and political influence, 11 July 2018 p. 34, https://ico.org.uk/media/2259369/democracy-disrupted-110718.pdf

[92] Wong, J., 'It might work too well': the dark art of political advertising online, 19 March 2018; Bodó, B. et al., Political micro-targeting: a Manchurian candidate or just a dark horse? Internet Policy Review 6(4):2017, p. 8

[93] For more information see: Gorwa R. and Guibeault D., Unpacking the Social Media Bot: A Typology to Guide Research and Policy. Policy and Internet, 2018.

[94] Fuchs, M., Why Social Bots Threaten Our Democracy in "Das Netz – English Edition: Digitalization and Society", July 2017, pp. 89-91, http://irights-media.de/publikationen/das-netz-english-edition/

[95] Howard, P. and Kollanyi, B., Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum. COMPROP Research note 2016.1: Oxford University, p. 5, https://arxiv.org/ftp/arxiv/papers/1606/1606.06356.pdf

found that bots reached "positions of measurable influence during the 2016 US elections",[96] and although their prevalence during French and German 2017 elections was not substantial, in Germany they were particularly active in the context of the refugee debate.[97] The radical right-wing party AfD was reported to have also used bots (and trolls) before the German elections.[98]

Social bots operate in different ways. They can make online measures of support, such as the number of likes, look larger – which could create an illusion of popularity for a political candidate. They can also serve to game algorithms to push content on to curated social feeds[99] or to divert users to the fake news websites. For example, on Twitter, coordinated bots' campaigns were employed to spread rumours and false reports with the hashtag #RapeFugees. They are programmed to constantly scan Twitter for keywords and then publish automated comments wherever they are found.[100] In effect, social bots distort quality information flow and manipulate decision-making processes by 'manufacturing consensus' – giving the illusion of a candidate or argument's online popularity in order to build real political support.[101]

1.2.4 Other distribution mediums

Although the digital amplification mechanisms discussed above play an important role in enabling a wide reach of misleading messages, conventional mediums, such as **media and individual users**, are also directly and/or indirectly engaged in this process.

1.2.4.1 Media

The **government-owned or government-sponsored media outlets**, including media outlets popular among national minorities in foreign countries, are known to be important vehicles of state-led disinformation campaigns.[102] Freedom House observed this tactic used inter alia in Hungary and Russia.[103] Research also shows that **partisan media** promotes misperceptions aligned with their ideology, with the aim of planting a seed of doubt about the legitimacy of expert opinions or evidence[104] (e.g. climate change, vaccination). It is, therefore, instrumental in the disinformation actions initiated by the political parties or political groups. **Mainstream news**

[96] Ibid., p. 3

[97] Howard, P., Junk News and Bots during the French Presidential Election: What Are French Voters Sharing Over Twitter? COMPROP data memo: Oxford University, 22 April 2017; Neudert L-M. et al., Junk News and Bots during the German Parliamentary Election: What are German Voters Sharing over Twitter? COMPROP data memo: Oxford University, 19 September 2017. A recent study by the Swedish Defense Research Institute found that 2,618 automated Twitter accounts were sending 6% of the content bearing the hashtags #svpol and #val2018 gathered since March 2018, see: Sweden Democrats – anti-immigration, anti-EU party set to win more votes than ever, 6 September 2018, https://theconversation.com/sweden-democrats-anti-immigration-anti-eu-party-set-to-win-more-votes-than-ever-102675

[98] See also: Die betrügerischen Fake-Accounts, die dich wütend machen sollen. 21. Januar 2019. https://krautreporter.de/2762-die-betrugerischen-fake-accounts-die-dich-wutend-machen-sollen. See also: Wie Trolle im Wahlkampf manipulierten. 01. 03. 2018.

https://faktenfinder.tagesschau.de/inland/manipulation-wahlkampf-101.html

[99] Hern, A., Facebook and Twitter are being used to manipulate public opinion – report, 19 June 2017, https://www.theguardian.com/technology/2017/jun/19/social-media-proganda-manipulating-public-opinion-bots-accounts-facebook-twitter

[100] Fuchs, M., Why Social Bots Threaten Our Democracy in "Das Netz – English Edition: Digitalization and Society", July 2017, pp. 89-91, http://irights-media.de/publikationen/das-netz-english-edition/

[101] Wooley, S. and Guilbeault D., Computational Propaganda in the United States of America: Manufacturing Consensus Online (2017), Working Paper No 2017.5: Oxford University, p. 3-4, http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-USA.pdf,

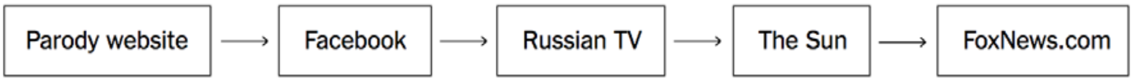
[102] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, pp. 71-72.

[103] Freedom House, Freedom on the Net 2017, p. 10, https://freedomhouse.org/sites/default/files/FOTN_2017_Final.pdf

[104] Weeks B., Why Partisan News—Not Just Fake News—Promotes Political Misperceptions, 12 April 2017, https://sites.lsa.umich.edu/learn-speak-act/2017/04/12/why-partisan-news-not-just-fake-news-promotes-political-misperceptions/

outlets may inadvertently become vehicles of disinformation by covering sensationalist claims or personalities. As explained by Marwick and Luis, "[a]nd even if the mainstream news was reporting on it in shock or disgust, it still led millions of viewers and readers to be exposed to these ideas".[105] The research by the Atlantic Council think tank demonstrates how the false story spreads using the media as one of the vehicles.

Figure 3: Lifecycle of a false story



Source: Atlantic Council and *New York Times*.[106]

1.2.4.2 Individuals

The success of a disinformation campaign can be measured by how widely **individuals and groups not directly related to the actor behind the campaign** endorse the campaign's narrative.[107] As such, individuals are not only targets but also mediums for disinformation distribution. According to MIT research on the spread of false content on Twitter, "robots accelerated the spread of true and false news at the same rate, implying that false news spreads more than the truth because humans, not robots, are more likely to spread it".[108] The motivation behind the human-enabled distribution of disinformation may differ from or coincide with that of the actor's.

Some individuals **take false manipulated information at face value** and further share it with their peers, for example, by 'liking', re-posting a message on social media or sharing it in the messaging application.[109] Some engage in the distribution of disinformation **for financial reasons** by taking advantage of the online advertising systems run largely by Google.[110] Others do it to "**signal shared ideological values" and to "militate against opposing perspectives"**.[111] This is increasingly leveraged in political campaigns. For example, elections in Italy saw the increase in individuals voluntarily surrendering their Twitter accounts to the application launched by a political party. By installing the application, all the Twitter messages posted by a political candidate were automatically retweeted by the supporters.[112] As a result, this led the said political candidate to gain considerable prominence on social media, as reported by major national news outlets.[113]

[105] Marwick, A. and Lewis R., Media, Manipulation and Disinformation Online. Data and Society, 2017, p. 22, https://centerformediajustice.org/wp-content/uploads/2017/07/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

[106] From: How Russian Propaganda Spread From a Parody Website to Fox News, 7 June 2017, https://www.nytimes.com/interactive/2017/06/07/world/europe/anatomy-of-fake-news-russian-propaganda.html

[107] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 74; Vian B. and McStay A. Fake news and the economy of emotions: problems, causes, solutions. Digital Journalism 6(2): 154-175, 2018.

[108] Vosoughi, S. et al., The spread of true and false news online, Science 359(6380): 1146-1151, 9 March 2018, p. 1146.

[109] The age of scepticism: from distrust to 'deepfake', https://www.ft.com/content/2fc9c1fa-d1a2-11e8-a9f2-7574db66bcd5

[110] Subramanian, S., Inside the Macedonian Fake-News Complex, 15 February 2017, https://www.wired.com/2017/02/veles-macedonia-fake-news/; Peinado, F., The business of digital manipulation in Spain, 24 May 2018, https://elpais.com/elpais/2018/05/24/inenglish/1527147309_000141.html

[111] Marwick, A., 'Beyond the Magic Bullet Theory of Fake News: Disinformation as Identity Expression'. Keynote speech at the iCS Symposium on Challenges to Studying Disinformation (University of Copenhagen), 27-28 October 2018.

[112] Chiuisi F. and Agosti C., The Influence Industry Personal Data and Political Influence in Italy, June 2018, pp. 12-13, https://ourdataourselves.tacticaltech.org/media/ttc-influence-industry-italy.pdf. An analogous technique was employed by the Russian Embassy in London through the 'Russian Diplomatic Online Hub'.

[113] Chiuisi F. and Agosti C., The Influence Industry Personal Data and Political Influence in Italy, June 2018, pp. 12-13, https://ourdataourselves.tacticaltech.org/media/ttc-influence-industry-italy.pdf. An analogous technique was employed by the Russian Embassy in London through the 'Russian Diplomatic Online Hub'.

The scenarios above highlight the **participatory nature of disinformation**.[114] According to Haigh et al., this includes the notion of 'peer-to-peer propaganda' as a situation in which "ordinary people experience the propaganda posts as something shared by their own trusted friends, perhaps with comments or angry reactions, shaping their own opinions and assumptions".[115] This trajectory further disguises the original actor of the disinformation campaign and on a certain level disperses the responsibility for its effects.

Finally, there is a separate group of individuals who significantly contribute to the dissemination of manipulated content and are **directly related to the actor behind the campaign, usually because they are paid by that actor**. These include recruited, influential, social media personalities ('influencers'[116]) and 'trolls' or 'sock puppets'.[117] The main functions of the trolls partially overlap with those performed by bots – they comment on posts, share links to disinformation websites and promote hashtags to give a certain issue an appearance of popularity and majority support ('astroturfing').[118] In addition, they are known to be more than relays and perform aggressive functions – actively engaging users in the discussion in the comments section, editing the content of Wikipedia pages, intimidating and harassing opposition and journalists.[119]

1.3 Analysis of disinformation/propaganda campaigns

1.3.1 Description and assessment of the recent, main disinformation actions with a formative effect on political opinion or which induced action

This subsection aims to describe and assess the major disinformation events of the last four years. As the study focuses on disinformation actions that have a strategic political objective, we have excluded from our research disinformation for financial gains, even if it had political consequences, such as the infamous case of "the Macedonian teenagers".[120]

The research is based on four types of sources: news reports, official documents, communication from stakeholders and scholarly articles. Thanks to the high international interest in the subject, several valuable papers are available that are based on field research and which classify the events under various aspects.[121]

The most violent case to date is the propaganda campaign in Myanmar, inciting hatred against the Rohingya minority. The recent Brazilian presidential election has been selected to illustrate the use of a digital platform – WhatsApp – that has not been much used for disinformation in Europe. The case of Hungary has been included as an instance of – in contemporary European terms – an overwhelming governmental disinformation campaign

[114] Asmolov G., The Disconnective Power of Disinformation Campaigns. Journal of International Affairs 71(1.5): Columbia, 18 September 2018, https://jia.sipa.columbia.edu/disconnective-power-disinformation-campaigns

[115] Ibid.

[116] Tucker, J. et al., Social media, political polarization and political disinformation: a review of the scientific literature. Hewlett Foundation, March 2018, p. 31, https://hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf

[117] 'Independent trolls' (not paid trolls) are not analysed in this report in detail. Gorwa R. and Guibeault D., Unpacking the Social Media Bot: A Typology to Guide Research and Policy. Policy and Internet, 2018, p. 9. An example of highly organized and institutionalized activity of trolls is a so-called 'Russian troll-factory', see Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 20, https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf. pp. 84-85.

[118] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 20, https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf. p. 87

[119] Ibid.

[120] For a more details on the discussed cases, see Annex 2

[121] See for example publications of the Oxford Internet Institute, among others: Bradshaw, Samantha, and Philip N Howard. "Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation." Working Paper. Oxford: Oxford Internet Institute, 2017. https://comprop.oii.ox.ac.uk/research/troops-trolls-and-trouble-makers-a-global-inventory-of-organized.social-media-manipulation/ as well as Bradshaw, Samantha and Philip N. Howard: Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation. Computational Propaganda Project, University of Oxford. 2018.

targeting the domestic population. The Italian elections and the American alt-right campaigns show the alleged use of disinformation by individuals and groups.

Information manipulation campaigns are multifaceted, and many factors are at play simultaneously. Describing them and their interplay in detail lies beyond the scope of this study. As current affairs with wide-ranging implications, these campaigns are also political and highly politicised. This, coupled with the fact that **most of the sources are news reports that have not gone through a rigorous peer-review process and are from newspapers with their own potential political prejudice makes this chapter vulnerable to charges of political bias.** The authors acknowledge this limitation and refer to section 1.1 (on definitional challenges) of this study, which delineates the exact technical characteristics of disinformation campaigns without regard to their content.

There is a long history of states and non-state actors taking disinformation action against domestic groups. The Oxford Internet Institute found that authoritarian regimes tend to target their own populations with social media campaigns, while in democracies, information campaigns are used by non-state (such as party) groups to target domestic populations.[122] A well-known example is **China's '50 cent army',** in existence since at least 2010, which has recently been revealed to post 448 million social media comments a year.[123] Another case worth mentioning is **Turkey**, where after the 2013 Gezi Park protests, the ruling AKP party recruited 6 000 people to conduct a disinformation campaign against its own population. The **'AK Trolls'** are particularly active on Twitter, where they spread pro-government messages, drown out critical voices (often with the help of bots) and abuse dissidents.[124] **Russia's Internet Research Agency** (IRA) also targets domestic populations, pushing the government's line. In these authoritarian regimes, the mainstream media are also strictly controlled by the government.

As discussed in section 1.1, politically motivated disinformation actions may be conducted by states and their proxies, by non-state groups or by individuals. It must be noted that **attribution in online disinformation campaigns is complicated, therefore it is not entirely possible to define the source, the funding of the disinformation campaign or whether it had a domestic or international effect**. Disinformation affects nearly all EU Member States and many nations worldwide. The cases selected for discussion here offer insights into different aspects of disinformation campaigns. Some of them are also particularly relevant for the future of the EU. For the constraints of the study, this subsection contains only information found to be the most crucial for assessment. For more detailed descriptions, as well as of other important cases that could not be included here, see Annex 2.

[122] See previous fn. first source.

[123] King, Gary, Jennifer Pan, and Margaret E. Roberts. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." American Political Science Review 111, no. 03 (August 2017): 484–501. https://doi.org/10.1017/S0003055417000144. This study found that contrary to earlier reports, the majority of posts do not aggressively push pro-government messages. Rather, they hijack the conversation with positive posts about China.

[124] "Freedom on the Net 2017: Manipulating Social Media to Undermine Democracy. Turkey Country Report." Freedom on the Net 2017. Washington, D.C.: Freedom House, November 14, 2017. https://freedomhouse.org/report/freedom-net/2017/turkey.

Table 6: Digital platforms used for disinformation operations

Digital platform	Main advantages for disinformation	Examples
Facebook	• The most popular social media platform worldwide • Wealth of user data • Very precise micro-targeting • Algorithm-driven feed curation prioritises content on the basis of popularity, not truth • Easy to game the system for popularity • Source of (dis)information seen as trustworthy • Groups create a feeling of community while keeping manipulated content out of sight, making it hard to trace or debunk	• The Internet Research Agency's campaigns in the US • Iran's disinformation campaigns in the US • Inciting hatred against the Rohingya • Attempts to spy on French President Emmanuel Macron
Twitter	• Easy to set up bots and semi-automated accounts that enable the fast spread of (dis)information • Anonymity, multiple accounts allowed • Ranking algorithm (trending hashtags) can be cheated	• 2016 US presidential elections • Brexit campaign • Turkey's AK trolls harassing government critics
Instagram	• Younger audience,[125] maybe more easily radicalised • Users do not expect to be confronted with partisan, political content; they are more vulnerable • Visually oriented app ideal for memes • 'Explore' feature promotes related content • Uses hashtags, which can be gamed	• IRA using Instagram to spread disinformation before the US presidential elections • Conspiracy theories/memes about billionaire George Soros[126]
Messaging apps	• (Dis)information comes from what is perceived as a trusted source • Messages spread in closed groups, not visible to fact-checkers on the outside • Encrypted conversations – hard to track • Group messages allow (dis)information to spread quickly	• WhatsApp: Violence in Mexico,[127] India,[128] and the disinformation campaign in Brazilian presidential elections
		• WeChat: Information manipulation targeting Chinese Americans[129]
YouTube	• Younger users, more easily radicalised	• RT's YouTube channel[131] • Conspiracy theories

[125] Larsson, Anders Olof. "The News User on Social Media: A Comparative Study of Interacting with Media Organizations on Facebook and Instagram." Journalism Studies 19, no. 15 (November 18, 2018): 2225–42. https://doi.org/10.1080/1461670X.2017.1332957.

[126] Ingram, David. "Attacks on Jewish People Rising on Instagram and Twitter, Researchers Say." NBC News. Accessed November 11, 2018. https://www.nbcnews.com/tech/tech-news/attacks-jewish-people-rising-instagram-twitter-researchers-say-n925086.

[127] Popken, Ben. "How WhatsApp Became Linked to Mob Violence and Fake News — and Why It's Hard to Stop." NBC News, November 2, 2018. https://www.nbcnews.com/tech/tech-news/how-whatsapp-became-linked-mob-violence-fake-news-why-it-n929981.

[128] "How WhatsApp Helped Turn a Village into a Mob." BBC, July 19, 2018, sec. India. https://www.bbc.com/news/world-asia-india-44856910.

[129] Chi, Zhang. "Study: Chinese-American Immigrants Fall Prey to WeChat's Misinformation Problem." Columbia Journalism Review, April 9, 2018. https://www.cjr.org/tow_center/wechat-misinformation.php.

[131] For an overview, see Bajoran, Donara. "YouTube's Kremlin Disinformation Problem." DFRLab (blog), May 3, 2018. https://medium.com/dfrlab/youtubes-kremlin-disinformation-problem-d78472c1b72b.

	• 'Autoplay' feature automatically plays content it deems related • Recommendation algorithm pushes viewers towards the extreme[130] • AI-curated 'trending news' ranks videos on the basis of popularity, not the quality of information	

Source: Authors.

1.3.2 Strategic political propaganda disseminated through social media sites

1.3.2.1 2016 US presidential elections

The disinformation campaign targeting US voters before the 2016 presidential elections was the first case when the global public had to face the enormous scale of online disinformation and propaganda. It is also the most well-documented recent disinformation campaign, and one of the few cases where investigators claim to have found clear evidence about the originator, pointing at Russia. The US Office of the Director of National Intelligence published a detailed report on Russia's meddling with the elections,[132] and to date, Special Counsel Robert Mueller has indicted 32 individuals, 26 of whom are Russian, and 3 Russian companies over election interference.[133] The disinformation actions in this wide-ranging and years-long campaign include Russian trolls commenting on news stories and maintaining several bogus Facebook and Twitter accounts,[134] with the "strategic goal to sow discord in the U.S. political system".[135] **The Russian trolls infiltrated both right- and left-wing communities online to stir controversy[136] by making emotionally charged posts about controversial issues. One important effect on society is the trenchant polarisation that can be observed in the US during and after the 2016 presidential campaign. At the same time, with such complex phenomena a cause–effect relationship is extremely difficult to prove empirically.**

It is complicated to quantify the impact the Russian campaign had on the US election results. Some analyses suggest that without the Russian influence, Donald Trump would not have won the elections,[137] and the knowledge that foreign interference may have played a role in Trump's victory can make the legitimacy of his presidency questionable in the eyes of some.[138]

In this campaign, the sources of disinformation actions were diverse: besides Russia, non-state groups and private individuals such as far-right (or 'alt-right') groups equally manipulated public opinion by spreading made-up

[130] Tufekci, Zeynep. "Opinion | YouTube, the Great Radicalizer." The New York Times, June 8, 2018, sec. Opinion. https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html.

[132] 'Assessing Russian Activities and Intentions in Recent US Elections' 2017.

[133] Prokop, Andrew. 2018. 'All of Robert Mueller's Indictments and Plea Deals in the Russia Investigation so Far'. Vox. 20 February 2018. https://www.vox.com/policy-and-politics/2018/2/20/17031772/mueller-indictments-grand-jury.

[134] Benedictus, Leo. 2016. 'Invasion of the Troll Armies: "Social Media Where the War Goes On"'. The Guardian, 6 November 2016, sec. Media. https://www.theguardian.com/media/2016/nov/06/troll-armies-social-media-trump-russian.

[135] "Internet Research Agency Indictment in the United States District Court for the District of Columbia." 2018. United States Department of Justice. p4. https://www.justice.gov/file/1035477/download, p. 4.

[136] Timberg, Craig, and Shane Harris. 2018. "Russian Operatives Blasted 18,000 Tweets Ahead of a Huge News Day during the 2016 Presidential Campaign. Did They Know What Was Coming?" Washington Post, July 20, 2018. https://www.washingtonpost.com/technology/2018/07/20/russian-operatives-blasted-tweets-ahead-huge-news-day-during-presidential-campaign-did-they-know-what-was-coming/.

[137] Jamieson, Kathleen Hall. 2018. Cyberwar: How Russian Hackers and Trolls Helped Elect a President What We Don't, Can't, and Do Know. Oxford: Oxford University Press. https://www.amazon.co.uk/dp/B07GJM18PL/ref=as_at?slotNum=1&linkCode=g12&imprToken=nU1ZSuTAQQjaa5h0wqES9Q&creativeASIN=B07GJM18PL&tag=tnyuk-21.

[138] Frum, David. 2018. "Trump's Crisis of Legitimacy." The Atlantic, July 17, 2018. https://www.theatlantic.com/ideas/archive/2018/07/is-trumps-presidency-legitimate/565451/.

news stories, conspiracy theories and memes virally on digital platforms.[139] Marwick and Lewis have found that the motivation behind media manipulation was a mixture of ideological conviction and financial interests as well as status and attention.[140] Fabricated news articles were found to heavily **favour Donald Trump**.[141]

Disinformation and manipulation were spread not exclusively through social media, but also via blogs, online journals, messaging applications, bulletin boards like 4chan and 8chan, and in the mainstream media. Among Trump voters, 40 % named *Fox News* as their main source for election news in the presidential election.[142] The right-leaning bias of *Fox News* is well-documented;[143] it is arguably closer to hyper-partisan websites than to the Anglo-American tradition of 'objective journalism'. Some even argue that it is not a journalistic operation but propaganda,[144, 145] or even a political actor in its own right.[146]

1.3.2.2 2018 US midterm elections

Online political activity before the 2018 US midterm elections received much international attention. Facebook announced on the day before the elections (on 5 November 2018) that it had removed 30 Facebook accounts and 85 Instagram accounts that may have been linked to Russia[147] (the number of the overall fake accounts was significantly higher, see below). On 19 October 2018, a Russian individual was charged with a criminal act by the US Justice Department for meddling with the midterm elections.[148]

It may have appeared that both the audience and the service providers were more aware of the risk of information manipulation. However, research has found that the amount of misinformation circulated online in the weeks leading up to the 2018 midterms actually increased compared with the 2016 presidential election; that users shared (again) more misinformation than true news; and that inauthentic and hyper-partisan news is getting into the mainstream in the US.[149]

[139] Marwick, A. and Lewis R., Media, Manipulation and Disinformation Online. Data and Society, 2017, https://centerformediajustice.org/wp-content/uploads/2017/07/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

[140] Marwick, A. and Lewis R., Media, Manipulation and Disinformation Online. Data and Society, 2017, https://centerformediajustice.org/wp-content/uploads/2017/07/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf. 27-32.

[141] Guess, Andrew, Brendan Nyhan, and Jason Reifler. "Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 U.S. Presidential Campaign." 2018. http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf.

[142] Gottfried, Jeffrey, Michael Barthel, and Amy Mitchell. "Trump, Clinton Voters Divided in Their Main Source for Election News | Pew Research Center," January 18, 2017. http://www.journalism.org/2017/01/18/trump-clinton-voters-divided-in-their-main-source-for-election-news/.

[143] Aday, Sean, Steven Livingston, and Maeve Hebert. "Embedding the Truth: A Cross-Cultural Analysis of Objectivity and Television Coverage of the Iraq War." Harvard International Journal of Press/Politics 10, no. 1 (January 2005): 3–21. https://doi.org/10.1177/1081180X05275727.

[144] Conway, Mike, Maria Elizabeth Grabe, and Kevin Grieves. "Villains, Victims and the Virtuous in Bill O'Reilly's 'No Spin Zone.'" Journalism Studies 8, no. 2 (April 1, 2007): 197–223. https://doi.org/10.1080/14616700601148820.

[145] Bard, Mitchell T. "Propaganda, Persuasion, or Journalism?: Fox News' Prime-Time Coverage of Health-Care Reform in 2009 and 2014." Electronic News 11, no. 2 (June 1, 2017): 100–118. https://doi.org/10.1177/1931243117710278.

[146] Yglesias, Matthew. "The Case for Fox News Studies." Political Communication 0, no. 0 (October 23, 2018): 1–3. https://doi.org/10.1080/10584609.2018.1477532.

[147] Gleicher, Nathaniel. "Election Update." Facebook Newsroom (blog), November 5, 2018. https://newsroom.fb.com/news/2018/11/election-update/.

[148] Gerstein, Josh. "U.S. Brings First Charge for Meddling in 2018 Midterm Elections." Politico, October 19, 2019. https://politi.co/2Ajbubq.

[149] Marchal, Nahema, Lisa-Maria Neudert, Bence Kollányi, and Philip N Howard. "Polarization, Partisanship and Junk News Consumption on Social Media During the 2018 US Midterm Elections." COMPROP DATA MEMO 2018. 5. Oxford: Oxford Internet Institute, November 1, 2018: at page 6. https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/11/marchal_et_al.pdf.

In addition to possible Russian interference, domestic trolls were also reportedly hard at work to spread disinformation.[150] In the days preceding the vote, Twitter removed 10 000 inauthentic accounts that tried to discourage Democrats from voting.[151] It would be interesting to see whether the domestic trolls overtook the role of the Russian trolls – this would be a sign of long-term success in the destruction of social norms and cohesion.

In the run-up to the elections, Facebook also removed 559 Pages and 251 accounts that it said used "sensational political content … to drive traffic to their websites, earning money for every visitor".[152] These were domestic sites that Facebook claims used clickbait political articles for monetary gains. This case illustrates some of the dilemmas Facebook faces when trying to police content. Owners of some of the accounts that were purged deny that they were running 'ad farms'. Instead, they claim to be legitimate political activists. The removal of their Pages resulted in accusations of Facebook censorship and of curtailing free speech.[153] Although Facebook claims to require the verification of identity of all political advertisers, to date there are no guidelines to distinguish citizens' justified political commentary from political advertisements.[154]

1.3.2.3 Brexit referendum

The UK's 2016 referendum on leaving the EU is probably the most painful for the EU for both practical and emotional reasons. Although less well-researched than the American case, **it appears that Russian disinformation campaigns targeted the UK's population, hoping to secure referendum victory for the Leave side**. The scope and reach of the campaigns appear to be smaller than in the US. The UK's Parliamentary Select Committee for Culture, Media and Sport, which is investigating the issue, stated in its interim report that it has "heard evidence of Russian state-sponsored attempts to influence elections in the US and the UK".[155]

Social media sites such as Facebook and Twitter deny any significant Russian meddling on their platforms, and researchers have also found comparatively little social media action.[156] Facebook was reluctant to internally investigate whether its facilities have been used by Russia to influence others.[157]

Twitter also claimed **not to have found significant Russian activity** regarding Brexit. Researchers identified **over 150 000 Twitter accounts** that listed Russian as their language and which tweeted about Brexit. **In the final days of the campaign, on the referendum day and the day after, they posted altogether over 45 000 tweets about Brexit, mostly to promote the Leave vote.**[158] **A prominent Brexit supporter on Twitter, with over 100**

[150] Collins, Ben. "In Secret Chats, Trolls Struggle to Get Twitter Disinformation Campaigns off the Ground." NBC News, November 6, 2018. https://www.nbcnews.com/tech/tech-news/secret-chats-trolls-struggle-get-twitter-disinformation-campaigns-ground-n931756.

[151] Bing, Christopher. "Exclusive: Twitter Deletes over 10,000 Accounts That Sought To..." Reuters, November 2, 2018. https://www.reuters.com/article/us-usa-election-twitter-exclusive-idUSKCN1N72FA.

[152] Gleicher, Nathaniel, and Oscar Rodriguez. "Removing Additional Inauthentic Activity from Facebook." Facebook Newsroom (blog), October 11, 2018. https://newsroom.fb.com/news/2018/10/removing-inauthentic-activity/.

[153] Dwoskin, Elizabeth, and Tony Romm. "Facebook Purged over 800 U.S. Accounts and Pages for Pushing Political Spam." Washington Post, October 11, 2018. https://www.washingtonpost.com/technology/2018/10/11/facebook-purged-over-accounts-pages-pushing-political-messages-profit/.

[154] Facebook's new authorization process for political ads goes live in the US. April 2018. https://techcrunch.com/2018/04/23/facebooks-new-authorization-process-for-political-ads-goes-live-in-the-u-s/

[155] "'Disinformation and "Fake News": Interim Report.'" House of Commons Culture, Media and Sport Select Committee, July 2018, para 2. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/36308.htm#_idTextAnchor033.

[156] For example, IRA-linked accounts paid Facebook altogether only $463 to show ads to British people "'Disinformation and "Fake News": Interim Report.'" House of Commons Culture, Media and Sport Select Committee, July 2018.

[157] Disinformation and 'fake news': Interim Report. 5. Russian influence in political campaigns. July 29. 2018. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/36308.htm#_idTextAnchor033

[158] Gorodnichenko, Yuriy, Tho Pham, and Olaksander Talavera. "Social Media, Sentiment and Public Opinions: Evidence from #Brexit and #USElection. Working Paper 24631." Cambridge, MA: National Bureau of Economic Research, May 2018. https://www.nber.org/papers/w24631.pdf. See also: "Russian Twitter Trolls Meddled in the Brexit Vote. Did They Swing It?"

000 followers, was exposed as a likely Russian troll.[159] Of the IRA-linked Twitter accounts identified for the US investigation, 419 also posted about Brexit, sometimes **anti-immigrant and anti-Muslim propaganda.** Researchers have also found evidence of bot activity[160] but it is unknown whether these bot networks were connected to Russia.

The Leave campaign was additionally implicated in the **Facebook–Cambridge Analytica data breach**. Both Leave.EU and Vote Leave face allegations of benefitting from the data breach. Moreover, Christopher Wylie, the Cambridge Analytica whistle-blower said Cambridge Analytica shared the data with Russian companies that have ties to the Russian intelligence services.[161]

The result of "the first major vote in the post-truth era"[162] came as a shock to many. Brexit is a hard blow to the cohesive force of the EU, and deeply affects the fundamental rights of the British people. British society remains bitterly divided about Brexit: the vote has torn the social cohesion of the UK and also threatens its territorial integrity. The suspicion of foreign interference raises deep concern for the functioning of democratic processes, including the legitimacy of the referendum.

1.3.2.4 2017 French presidential elections

Although much less well-documented and with no conclusive evidence, it appears that the French elections, and specifically the campaign of current French President Emmanuel Macron, were subject to disinformation campaigns launched by Russia. Macron's email was hacked, Russian agents tried to spy on him using Facebook by allegedly posing as friends of friends of his associates to get access to personal information and unsubstantiated rumours were spread about him – yet the disinformation efforts largely failed in France. Some explain this by saying the French are less susceptible to fake news because they prefer mainstream media.[163] Another claim is that the French institutions tasked with ensuring the integrity of the elections (such as the elections watchdog or the national security agency) worked better than they did in the US.[164] Others allege that the hackers made foolish mistakes, making it easier for Macron's team to denounce the hacks.[165] Still, others found that the **campaign attracted foreign Twitter users, rather than French voters** – after all, even the hashtags such as #MacronLeaks were in English.[166] One key aspect appears to be that **French voters and politicians were aware of the problem**. After the high-profile disinformation actions during the Brexit campaign, and particularly during the US presidential elections, the problem was widely discussed in France. This may be the crucial element that made French society better equipped to handle disinformation. At the same

The Economist, November 23, 2017. https://www.economist.com/britain/2017/11/23/russian-twitter-trolls-meddled-in-the-brexit-vote-did-they-swing-it.

[159] Dearden, Lizzie. "Pro-Brexit Twitter Account with 100,000 Followers Could Be Part of Russian 'Disinformation Campaign.'" The Independent, August 30, 2017. http://www.independent.co.uk/news/uk/home-news/david-jones-pro-brexit-ukip-twitter-account-russia-fake-bot-troll-trump-disinformation-followers-a7920181.html.

[160] Bastos, Marco T., and Dan Mercea. "The Brexit Botnet and User-Generated Hyperpartisan News." Social Science Computer Review, October 10, 2017: https://doi.org/10.1177/0894439317734157.

[161] "Whistleblower: Cambridge Analytica Shared Data with Russia." Euractiv.Com, May 17, 2018. https://www.euractiv.com/section/global-europe/news/whistleblower-cambridge-analytica-shared-data-with-russia/.

[162] Viner, Katharine. "How Technology Disrupted the Truth." The Guardian, July 12, 2016, sec. Media. https://www.theguardian.com/media/2016/jul/12/how-technology-disrupted-the-truth, para. 9.

[163] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 20. https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf.

[164] Bulckaert, Ninon. 2018. "How France Successfully Countered Russian Interference during the Presidential Election." Euractiv.Com (blog). July 17, 2018. https://www.euractiv.com/section/elections/news/how-france-successfully-countered-russian-interference-during-the-presidential-election/.

[165] Ibid.

[166] Ferrara, Emilio. "Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, June 30, 2017. https://papers.ssrn.com/abstract=2995809.

time, it must be noted that far-right candidate Marine Le Pen, supported by Russia[167] and Russian media,[168] came in second in the presidential elections.

1.3.2.5 2017 German federal elections

The 2017 German elections are noteworthy for the **lack of disinformation actions**. It is thought that greater awareness by all stakeholders prevented or discouraged attempts at manipulation. Especially after the earlier, high-profile disinformation actions (see the 'Lisa case' below), the risks of being manipulated were widely discussed in Germany.[169] Successful earlier disinformation actions may have also helped to prepare citizens to be more impervious to hoaxes. One possible explanation could be that after the much-discussed US case and the failure in France, Russia did not want to risk another wide-ranging operation. **German parties prepared for an eventual hack and disinformation campaign: they agreed not to use bots[170] or leaked data**.[171] Germany also adopted the Network Enforcement Act (see subsection 4.1.1), which introduced fines of up to EUR 50 million on social media companies if they fail to remove fake, hate-inciting or criminal content.[172] Additionally, Facebook took action against thousands of fake accounts before the elections, partnered with German authorities, and shared security tips with parties and candidates.[173]

The radical right-wing party AfD, however, has reportedly organised a coordinated social media campaign with the help of bots and a troll army counting 5 000 to 6 000 persons, each creating several social media accounts. The project, named 'Reconquista Germanica', "could not have been realised without Russian help", as maintained by an organiser.[174]

The highest profile disinformation action in Germany to date occurred a year before the elections in the infamous '**Lisa case**', in which a 13-year old Russian-German girl went missing for a day in January 2016. Russian Channel One television reported that she had been raped by Arab migrants. This news, despite being quickly debunked by the German police, got broad coverage in Russian media, and was widely distributed on social media. Demonstrations were organised on Facebook by Russian minorities in Germany. Taking the case to the diplomatic level, Russian Foreign Minister Sergei Lavrov told journalists that the German authorities had failed to conduct a proper investigation.[175]

[167] Gatehouse, Gabriel. "Who's Funding France's Far Right?," BBC.com, April 3, 2017, sec. Europe. https://www.bbc.com/news/world-europe-39478066.

[168] Klingová, Katarína, Daniel Milo, Veronika Víchová, Lóránt Győri, and Patrik Szicherle. "Information War Monitor for Central Europe: Pro-Kremlin Disinformation Outlets Have a Favourite French Presidential Candidate." GLOBSEC. Accessed December 21, 2018. https://www.globsec.org/publications/information-war-monitor-central-europe-pro-kremlin-disinformation-outlets-favourite-french-presidential-candidate/.

[169] Palma, Bethania. "Germany Election So Far Unaffected by 'Fake News.'" Snopes.com, September 23, 2017. https://www.snopes.com/news/2017/09/23/german-election-so-far-unaffected-by-fake-news/.

[170] Neudert, Lisa-Maria N. "Computational Propaganda in Germany: A Cautionary Tale." Project on Computation Propaganda. Oxford, UK: Oxford Internet Institute, 2017. http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-Germany.pdf.

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

[172] Győri, Lóránt, Péter Krekó, Jakub Janda, and Bernhard Weidinger. "Does Russia Interfere in Czech, Austrian and Hungarian Elections? A Study by Political Capital, European Values Think-Tank in Cooperation with DöW," 2017. https://www.kremlinwatch.eu/userfiles/western_experiences_eastern_vulnerabilities_20171012_15273208786863.pdf.

[173] Zuckerberg, Mark. "I Just Went Live a Minute Ago. Here's What I...," September 21, 2017. https://www.facebook.com/zuck/posts/10104052907253171.

[174] The organiser using the nickname Nikolai Alexander told to ARD: "Ohne russische Unterstützung wäre das Projekt in dieser Form wohl nicht möglich gewesen." AfD-Funktionär an Troll-Attacken beteiligt. 01. 03. 2018.

https://faktenfinder.tagesschau.de/inland/manipulation-wahlkampf-103.html.

[175] Meister, Stefan. "The 'Lisa Case': Germany as a Target of Russian Disinformation." NATO Review Magazine, July 15, 2016. http://www.nato.int/docu/review/2016/Also-in-2016/lisa-case-germany-target-russian-disinformation/EN/index.htm.

1.3.2.6 2018 Efforts by Iran to increase polarisation in the US and the UK

Although Russia appears by far to be the most active state using disinformation strategically on foreign populations, it is not the only one. In August 2018, and then again in October 2018, Facebook and Twitter removed 652 bogus Iranian accounts posting a lot of content on divisive issues, targeting states in the Middle East, Latin America, the UK and the US.[176] The effects of these actions are hard to quantify, but they may have contributed to the polarisation of societies. The operations had been going on since 2011, but their reach was relatively limited with fewer than 1 million followers. The method was to infiltrate both left and right groups and publish memes, pictures and other content on their behalf to further divide these communities. Analysis of a sample of their content shows a clear evolution: instead of pushing explicitly pro-Iranian messages, the accounts, **masquerading as American liberals**, shared content on **divisive issues** in the US, such as gun control and race relations.

1.3.3 Disinformation actions in the service of nationalism, populism, hate speech and extremism

1.3.3.1 Hungarian government campaigns against migrants and against George Soros

In 2013-14, Hungary appeared to be a target of Russian disinformation campaigns. Gradually, Hungarian pro-government media outlets allegedly took over spreading pro-Kremlin propaganda themselves.[177] "Pro-government disinformation matches Kremlin narratives without any direct influence from Russia", the Oxford Internet Institute found.[178]

As for local disinformation actions, the Hungarian government was reported to "[finance] an entire fake news industry",[179] running several information campaigns since 2015. Two of the bigger campaigns have been the anti-migrant campaign, which started in 2015, and the campaign against Hungarian-born US billionaire George Soros in 2017. The latter one culminated in a legislative package entitled "Stop Soros" in June 2018, imposing severe restrictions on civil society organisations.[180] These campaigns arguably consisted of spreading far-right stereotypes on every distribution channel available to the government, including billboards, television, radio, print media and social media. However, for online propaganda, the Oxford research rated the Hungarian government as "low capacity".[181]

After the anti-migrant campaign, the Hungarian population was found to be more xenophobic than at any time in the past 25 years.[182] With the campaign, the "ethno-nationalist boundaries of Hungarian-ness" have been reinforced.[183] The **government's permanent campaign has arguably had the effect of mainstreaming**

[176] Gilbert, David. "Iran Is Running an Online Disinformation Campaign on the Scale of Russia's Troll Farm." Vice News, August 22, 2018. https://news.vice.com/en_ca/article/594ekk/iran-russia-facebook-twitter-disinformation.

[177] Bayer, Lili. "Fidesz-Friendly Media Peddling Russian Propaganda." The Budapest Beacon, November 17, 2016. https://budapestbeacon.com/fidesz-friendly-media-peddling-russian-propaganda/.

[178] Bradshaw, Samantha, and Philip N Howard. "Online Supplement to Working Paper 2018.1 Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation." Oxford: Oxford Internet Institute, 2018. http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct_appendix.pdf, p. 29.

[179] Ibid. p. 29.

[180] Dunai, Márton. "Hungary Approves 'STOP Soros' Law, Defying EU, Rights Groups." Reuters, June 20, 2018. https://www.reuters.com/article/us-hungary-soros/hungary-approves-stop-soros-law-defying-eu-rights-groups-idUSKBN1JG1VN.

[181] Bradshaw, Samantha, and Philip N Howard. "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation." Oxford: Oxford Internet Institute, University of Oxford, 2018, p. 17. http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct2018.pdf

[182] Keszthelyi, Christian. "Xenophobia Skyrocketing in Hungary, Surveys Reveal." Budapest Business Journal, November 17, 2016. https://bbj.hu/budapest/xenophobia-skyrocketing-in-hungary-surveys-reveal_124920. It found the tendency that "xenophobia saw a hike in Hungary when the refugees had disappeared and the campaign against "migrants" by the Hungarian government accelerated."

[183] Thorleifsson, Cathrine. "Disposable Strangers: Far-Right Securitisation of Forced Migration in Hungary." Social Anthropology 25, no. 3 (August 1, 2017): 318–34, p. 318. https://doi.org/10.1111/1469-8676.12420.

and normalising extreme-right beliefs and language. Increasingly radical, the campaigns keep moving the range of ideas and the rhetoric that are acceptable in public discourse further to the right. Arguably, the campaign has continually been on the verge of legality for its content and distribution methods. The overwhelming nature of this single-issue campaigning has conceivably **erased policy discussions from the agenda.**

1.3.3.2 2018 Italian general elections

The situation in Italy might be seen by some as comparable to that in Hungary: both have allegedly pro-Russian politicians in power, as well as Russian-friendly media outlets.[184, 185] Italian journalists claim to have uncovered some signs of Russian-influenced campaigns[186]. As for locally produced disinformation, while Italian communications regulator AGCOM named 2017 as "the year of the emergence of 'fake news'",[187] concerns about disinformation go back earlier. Fact-checking site Pagella Politica found that half of the most widely shared stories about the referendum in 2016 were fabricated.[188] Italian so-called anti-establishment parties have long been accused of spreading disinformation.[189] Journalistic investigations have linked Movimento Cinque Stelle (M5S)[190] and Lega Nord to seemingly independent websites and social media accounts that push fabricated content. **False stories were found to be shared by leading politicians on social media, presenting Facebook with the problem of how to interfere without being accused of meddling with the elections**.[191] In the run-up to the elections, most of the disinformation shared concerned migrants.[192] Still, for all the talk about fake news, a study looking at data from 2017 found that the **disinformation sites were insignificant in terms of both reach and engagement times when compared with mainstream media.**[193] Beyond disinformation, journalistic research reportedly found that Lega Nord leader Matteo Salvini and M5S leader Luigi Di Maio used "inflammatory and visually arresting" content as well as viral videos and live broadcasting to dominate the election campaign on Facebook.[194]

[184] Plucinska, Joanna, and Mark Scott. "How Italy Does Putin's Work." Politico, March 3, 2018. https://www.politico.eu/article/italy-election-fake-news-sunday-bufale-misinformation-vladimir-putin-russia/.

[185] Horowitz, Jason. "Will Russia Meddle in Italy's Election? It May Not Have To." The New York Times, October 10, 2018, sec. World. https://www.nytimes.com/2018/03/01/world/europe/italy-election-russia.html.

[186] "Russian 'troll Factory' Tweets Tried to Influence Italian Voters." The Local Italy, August 2, 2018. https://www.thelocal.it/20180802/russian-troll-factory-tweets-attempted-influence-italian-elections

[187] "2017 Was 'the Year of Fake News in Italy', Regulator Warns." The Local, February 20, 2018. https://www.thelocal.it/20180220/fake-news-spread-italy-agcom-election.

[188] Kottasova, Ivana. "Did Fake News Influence Italy's Referendum?" CNNMoney, December 5, 2016. https://money.cnn.com/2016/12/05/media/fake-news-italy-referendum/index.html.

[189] Bradshaw, Samantha, and Philip N Howard. "Online Supplement to Working Paper 2018.1 Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation." Oxford: Oxford Internet Institute, 2018. http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct_appendix.pdf.

[190] Nardelli, Alberto, and Craig Silverman. "One of the Biggest Alternative Media Networks In Italy Is Spreading Anti-Immigrant News and Misinformation on Facebook." BuzzFeed, November 21, 2017. https://www.buzzfeed.com/albertonardelli/one-of-the-biggest-alternative-media-networks-in-italy-is.

[191] Plucinska, Joanna, and Mark Scott. "How Italy Does Putin's Work." Politico, March 3, 2018. https://www.politico.eu/article/italy-election-fake-news-sunday-bufale-misinformation-vladimir-putin-russia/

[192] Alaphilippe, A, C Ceccarelli, L Charlet, and M Mycielski. "Disinformation Detection System: 2018 Italian Elections." Brussels: EU Disinfo Lab, June 1, 2018. https://disinfo.eu/wp-content/uploads/2018/06/2018-Italian-elections-Case-report.pdf.

[193] Fletcher, Richard, Alessio Cornia, Lucas Graves, and Rasmus Kleis Nielsen. "Measuring the Reach of 'Fake News' and Online Disinformation in Europe." Reuters Institute for the Study of Journalism; University of Oxford, February 2018. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/Measuring%20the%20reach%20of%20fake%20news%20and%20online%20distribution%20in%20Europe%20CORRECT%20FLAG.pdf.

[194] Kalia, Ammar, and Caelainn Barr Angela Giuffrida in Rome. "Revealed: How Italy's Populists Used Facebook to Win Power." The Guardian, December 17, 2018, sec. World news. https://www.theguardian.com/world/2018/dec/17/revealed-how-italy-populists-used-facebook-win-election-matteo-salvini-luigi-di-maio .

1.3.3.3 Disinformation campaign against the Rohingya minority in Myanmar since 2013

While violence has been triggered occasionally in the examples above, in the case of Myanmar's Rohingya minority, disinformation campaigns were part of what the UN described as a genocide.[195] Facebook, extremely popular in the country, has been used by Buddhist nationalists to incite hatred against the Rohingya since 2013. In addition to regular Facebook accounts spreading hate, often with unsubstantiated reports, **about 700 military personnel were reportedly also instructed to create fake pages on social media**. These pages, pretending to be celebrity fan sites, published a high number of anti-Rohingya stories.[196] The **disinformation and propaganda campaign that preceded ethnic cleansing contributed to the spread of hatred**.

Furthermore, less researched examples should be mentioned where Russian influence is suspected. Such is the case of the Spanish referendum on the independence of Catalonia (1 October 2017), where the suspected Russian interference provoked conflict, with the alleged intention to discredit the Spanish democratic system.[197] Before the Czech presidential elections (January 2018), 118 websites were identified as promoting dubious content (71 Czech, 41 Slovak and 6 foreign outlets).[198] In the Czech presidential elections, pro-Russian incumbent Miloš Zeman's pro-EU opponent Jiří Drahoš was reportedly the subject of wide-ranging disinformation campaigns.[199] About 30 pro-Russian websites were found to have smeared Drahoš with allegations that he had collaborated with the Communist secret police, supported unrestricted immigration or was a homosexual paedophile.[200] At the same time, a study of 6 popular Czech disinformation websites found that they paid little attention to the elections, and even when they did, they relied on emotive language rather than false information as such.[201]

The Irish referendum on repealing the country's abortion ban (25 May 2018) is also an interesting case of disinformation-related challenges. Concerned about foreign influence, a civic initiative called the Transparent Referendum Initiative was set up, which tracked the origins of Facebook ads.[202] An analysis by the NGO openDemocracy found foreign groups, mostly from the US and Canada, as well as alt-right activists behind some of the ads.[203] Some of the groups that had paid for ads were untraceable.[204] On 8 May, Facebook blocked ads

[195] Mozur, Paul. "A Genocide Incited on Facebook, With Posts From Myanmar's Military." The New York Times, October 18, 2018, sec. Technology. https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html.

[196] Ibid.

[197] Alandete, David. "Russian Network Used Venezuelan Accounts to Deepen Catalan Crisis." El País. November 11, 2017, sec. In English. https://elpais.com/elpais/2017/11/11/inenglish/1510395422_468026.html.

[198] Globsec: What Do We Know About Disinformation Websites in the Czech Republic and Slovakia?

24.09.2018. https://www.globsec.org/news/what-do-we-know-about-disinformation-websites-in-the-czech-republic-and-slovakia/#psDVScSk4axMyDoC.99

[199] Krejčí, Markéta, Veronika Víchová, and Jakub Janda. "The Role of the Kremlin's Influence and Disinformation in the Czech Presidential Elections." Th Kremlin Watch Report. European Values, January 29, 2018. https://www.kremlinwatch.eu/userfiles/the-role-of-the-kremlin-s-influence-and-disinformation-in-the-czech-presidential-elections_15263778517686.pdf.

[200] Ibid. pp 1-2.

[201] Syrovátka, Jonáš, and Jaroslav Hroch. "Czech Presidential Election 2018." Czech Elections in the Era of Disinformation. Prague: Prague Security Studies Institute, 2018. http://www.pssi.cz/download/docs/545_presidential-election-2018-analysis.pdf.

[202] Dwyer, Craig. "How Digital Threats to Democracy Were Tackled During Ireland's Abortion Referendum." Media Policy Project, London School of Economy (blog), July 10, 2018. http://blogs.lse.ac.uk/mediapolicyproject/2018/07/10/how-digital-threats-to-democracy-were-tackled-during-irelands-abortion-referendum/.

[203] Provost, Claire, and Lara Whyte. "Foreign and 'alt-Right' Activists Target Irish Voters on Facebook Ahead of Abortion Referendum." openDemocracy, April 25, 2018. https://www.opendemocracy.net/5050/claire-provost-lara-whyte/north-american-anti-abortion-facebook-ireland-referendum.

[204] Fitzgerald, Cormac. "Concerns over Mystery Facebook Ads Claiming to Offer 'unbiased Facts' on Eighth Referendum." The Journal, May 1, 2018. http://www.thejournal.ie/eighth-referendum-ads-3986039-May2018/.

related to the referendum that originated outside of Ireland.[205] This addressed the problems with foreign ads but not misleading local ads.[206] Facebook also introduced a tool allowing Irish Facebook users to "see all of the ads any advertiser is running on Facebook in Ireland at the same time".[207] Even more drastic, on 9 May, Google banned all advertisement relating to the referendum.[208] While the goal may be commendable, Google was accused of depriving campaigning groups of an important platform to spread their messages. Moreover, neither measures addressed the problem of organic, non-promotional content containing disinformation.

1.3.4 Summary of events

As discussed in the methodology section, **disinformation operations are inherently secretive**; only the tip of the iceberg can be proven. Foreign countries accused of meddling – such as Russia and Iran – deny the charges. The Italian M5S co-founder Beppe Grillo called the accusations "ridiculous".[209] **Funding links are complicated to establish except in such rare cases as when Facebook releases the relevant information about ad purchases**. **Causal links to concrete real-world events are notoriously hard to establish as well, rendering the assessment of effects more difficult**.[210] For example, it has been claimed that xenophobia increased in Hungary based on surveys conducted after the anti-immigration campaigns, but it cannot be empirically proven that this was caused by the campaign.[211] Likewise, it cannot be claimed that ethnic violence in Myanmar was caused by the hatred campaign led with the use of Facebook.

The various dodgy methods of manipulation are discussed extensively in a previous section (1.2), some of them directed at dividing a group with simple distraction or divisive comments – which makes exposure even more difficult. In politically charged cases, identifying the target or the objectives may also be contentious. These are important limitations for Table 7, which summarises the events.

[205] "Facebook Will Not Be Accepting Referendum Related Ads from Advertisers Based Outside of Ireland | Facebook." Facebook Ireland (blog), May 8, 2018. https://www.facebook.com/notes/facebook-dublin/facebook-will-not-be-accepting-referendum-related-ads-from-advertisers-based-out/10156398786998011/.

[206] McSorley, Christina. "Google Abortion Poll Ban 'Outrageous.'" BBC.Com, May 10, 2018, sec. Europe. https://www.bbc.com/news/world-europe-44067607.

[207] Op. Cit. "Facebook Will Not"

[208] Satariano, Adam. "Ireland's Abortion Referendum Becomes a Test for Facebook and Google." The New York Times, May 25, 2018, sec. Technology. https://www.nytimes.com/2018/05/25/technology/ireland-abortion-vote-facebook-google.html.

[209] Horowitz, Jason. "Spread of Fake News Provokes Anxiety in Italy." The New York Times, December 22, 2017, sec. World. https://www.nytimes.com/2016/12/02/world/europe/italy-fake-news.html.

[210] Two researchers claim to have found some causal link between anti-refugee sentiment on Facebook and real-world violence against refugees in Germany (Müller, Karsten, and Carlo Schwarz. "Fanning the Flames of Hate: Social Media and Hate Crime." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, May 21, 2018. https://papers.ssrn.com/abstract=3082972). Their findings, however, have been critiqued (Cottee, Simon. "Can Facebook Really Drive Violence?" The Atlantic, September 9, 2018. https://www.theatlantic.com/international/archive/2018/09/facebook-violence-germany/569608/). In a separate study, these researchers established correlation between US President Donald Trump's Islam-related tweets and anti-Muslim hate crime (Müller, Karsten, and Carlo Schwarz. "Making America Hate Again? Twitter and Hate Crime Under Trump." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, March 30, 2018. https://papers.ssrn.com/abstract=3149103).

[211] In a ballot in Florida in 2010, a low-budget campaign targeted voters via Facebook. In the areas where the Facebook ad ran, the group did 19 percentage points better than in the areas the ad did not target. A post-election poll even showed that many voters who voted in accordance with the campaign cited arguments used in the ads when explaining his/her vote. ("(2) Case Study: Reaching Voters with Facebook Ads (Vote No on 8)." Facebook, August 16, 2011. https://www.facebook.com/notes/government-and-politics-on-facebook/case-study-reaching-voters-with-facebook-ads-vote-no-on-8/10150257619200882/.) Yet, this is correlation not causation.

Table 7: Major disinformation actions, their sources, goals, tools, sponsors and results

Event	Originator(s)	Target/event to be influenced/ objectives	Methods	Funding	Influence on public opinion/ impact on core values
2016 US presidential elections	Foreign government (suspected)[212]	To sow discord among the US electorate To reduce Democratic candidate Hillary Clinton's chances of winning	Setting up fake accounts on Facebook and Twitter, posing as Americans Spreading propaganda on RT and Sputnik Amplifying messages with bot networks Email hacking, releasing confidential information	Putin-ally Yevgeniy Prigozhin (indicted)	Delegitimising the office of the US president for some Undermining trust in public institutions Increasing polarisation between Democrats and Republicans
	Non-state action (alt-right groups)	To confuse political and advocacy groups To raise social tensions and polarise society Discredit the Democratic Party candidates	Creating memes, conspiracy theories and fabricated news stories Manipulating social media sites' algorithms to spread the content	Unknown	Hyper-partisanship Polarisation of society Demobilisation and disengagement Undermining public trust in the media
Brexit referendum (2016)	Foreign government (suspected)[213]	To promote the Leave campaign To damage the integrity of the EU	Setting up fake accounts on Facebook and some on Twitter Spreading propaganda on RT and Sputnik Alleged illegal campaign financing	Under investigation	Divisions in the UK Delegitimising the referendum for some
2017 French elections	Foreign government (suspected)[214]	To reduce En Marche candidate Emmanuel Macron's chances of winning To promote National Front candidate Marine Le Pen	Email hacking and the release of confidential information Spying on Facebook Spreading propaganda on RT and Sputnik Amplifying messages with bot networks	Russia (suspected)	Failed The integrity of the elections was not questioned

[212] Russia is suspected.

[213] Russia is suspected.

[214] Russia is suspected.

2017 German federal elections	n/a	n/a	n/a	n/a	No significant action was attempted
2018 disinformation operations against several states	Iran (suspected)	To polarise society in the US, the UK, Latin America and the Middle East	Setting up fake accounts on Facebook and Twitter Posting highly charged content on controversial issues Amplifying messages with bot networks	Iran (suspected)	Potentially creating distrust and polarisation
Anti-migrant, anti-Soros campaigns	Hungarian government	To strengthen loyalty with the ruling party Fidesz and its foreign policy Fear-based communication, to justify government policy To define an out-group as a culprit To set the agenda for public discourse	Newspaper ads, billboards, radio and television spots, online banners 'National consultation' letters to voters Public and pro-government private media spreading the content	Tax-payers	Normalisation of far-right ideas and rhetoric Increasing xenophobia Erasing other issues from the agenda
2018 Italian general elections	Political party (reportedly)[215]	To win the elections, to discredit political opponents and to rile constituents	Reportedly pushing divisive, fabricated content on a network of seemingly unaffiliated sites and social media accounts Amplifying the content via politicians' and party accounts	n/a	Increasing pre-existing anti-immigrant and anti-establishment sentiments
2018 Brazilian presidential elections	All candidates	To win the elections, to discredit political opponents and to rile constituents	Bombarding voters with disinformation on WhatsApp	n/a	Hyper-partisanship and polarisation

[215] See footnote 193.

Campaign against the Rohingya in Myanmar	Myanmar government, military (suspected)	To fuel anti-Muslim and nationalist sentiments	Spreading hate speech and false news on Facebook Setting up fake accounts on Facebook to spread the message	Myanmar government (suspected)	Propaganda to support a strategy to force 700 000 Rohingya to leave the country At least 6 700 Rohingya killed in Aug.–Sept. 2017 At least 288 Rohingya villages destroyed

Sources: Authors.

2. IMPACT ON DEMOCRACY, FUNDAMENTAL RIGHTS, THE RULE OF LAW AND THE EU

KEY FINDINGS

- Social media (and other web2.0 services) brought about a new form of public sphere (*Öffentlichkeit*) that is more inclusive than any other public sphere before. In that context, people who are more susceptible to manipulation tend to amplify misinformation and disinformation through their posts, comments, likes and shares, and they form the electoral basis for populistic politicians.
- The new public discourse is horizontally organised, based on the rules of the platform and is not conducive to central supervision. The volume of the content makes reliance on supervision insufficient, if not futile.
- Platform providers design the structure of communication, but they are not to blame for what content people share through that structure. They aggregate and classify information, and mediate services between citizens/consumers, corporate actors, NGOs, institutions and other players.
- The shared information basis and the common narratives of society that are the preconditions of democratic public discourse are being splintered by filter bubbles, and further ruined by micro-targeting.
- Participatory and deliberative democracy requires the provision of information and effective processes of consultation. Meaningful participation in a debate is only possible if individuals know what public policy is, how they are affected and what alternative solution there is.
- People still trust the traditional media sources, but less so online media sources.[216]
- User-generated and spontaneous information is on equal legal terms with strategically designed and artificially disseminated information. Susceptible users become weaponised as instruments for disseminating disinformation and propaganda.
- Regulation should demand an architecture from social media that provides an equal and fair setting for all opinions to be heard and for human rights to be protected.
- Some cases of disinformation and propaganda serve the strategic goal of overthrowing democratic systems. Popular support for this also signals a functional challenge to the operation of democracy, and requires the actions of militant democracy.[217]
- In line with the concept of militant democracy**,** mature constitutionalism implies the existence of robust precautionary measures in democratic systems to protect them against a future potential political force that, when entering government, replaces a constitutional government with an autocratic one.
- The postmodern risk society is vulnerable to fear creation and populism. Deliberative democracies need to respond to public fear, while also staying committed to deliberation and rationalism.

[216] Flash Eurobarometer (EBS) 464, Fake news and disinformation online. April 2018. These results show a contradiction with the results in 2017 which showed 34% trusted the "media", and 61% not trusted, but no differentiation between online and offline media was included.

[217] For a full description see K. Loewenstein, 'Militant Democracy and Fundamental Rights,' 31 *American Political Science Review* 417–433 and 638–658 (1937). For a most recent authoritative account of such a function of international legal mechanisms, see R. Dworkin, 'A New Philosophy of International Law' (2013) 41 *Philosophy and Public Affairs* 1, 2–30 (2013).

2.1 Changes in the structure of the public sphere accelerating the spread of disinformation and propaganda

Public discourse is an inherent basis for democracy, and in the past centuries it has been facilitated by the free press and mass media in democratic countries. Before explaining how disinformation and propaganda impact democracy, we provide an overview of how the online communication environment has changed after the millennium. This new public sphere has become fertile ground for the disinformation and propaganda campaigns, and in this respect, it can be regarded as their precondition. At the same time, it is also being transformed organically by the very discourse that includes disinformation as well as its criticism and reflections. Thereby it behaves like a moving target for scientific observation.

The public sphere (*Öffentlichkeit*)[218] consists of professional media on the one hand, and citizens' communication on the other. The latter used to be direct and unmediated, exercised in public spaces like coffee houses and town squares. The interactive social media platforms brought back these **virtual public spaces**, potentially engaging everyone with internet access in an interactive discussion. But these new 'coffee house' conversations are no longer unmediated, because the platform operators interfere with their algorithms. If they did not, the discussion would fall into a global cacophony.

The **emergence of social media signals a new age in the public sphere**. This new public sphere gives way to undercurrent voices in society that challenge the existing status quo of democratic governance. (How disinformation impacts the system of democracy is set out more specifically in sections 2.2 and 2.3.) Masses of voters, whose opinions were underrepresented earlier, learned to express themselves with the help of social media.[219] However, the impact on public discourse is still dominated by powerful actors, such as political parties, advertisers and big media outlets. The likes and the shares of users who are susceptible to manipulation have been **weaponised** to serve particular interests, rather than the interests of the voters themselves.

While **users feel empowered** by the easy access to information and knowledge, and increasingly dissatisfied with the ruling political class, they also may **feel excluded** from meaningful participation in traditional decision-making processes.[220] The transparency of democratic processes often reveals the hesitation of decision makers and the **malfunctions** of the state administration. Newly emerging fears and insecurity in society are not addressed appropriately by the ruling governments (see subsection 2.2.3. on the postmodern risk society). This **frustration** leads to **distrust** in the establishment, including the media. It also forms a breeding ground for **intolerance** and **susceptibility to extreme 'alternative' solutions.** In a hyper-pluralistic information environment **rational information is often dwarfed by the legion of other voices**. **Research has supported the hypothesis that populist communication is aligned with social media, partly due to the special characteristics of social media (direct access to the audience without journalistic interference, personal connection, potential for personalisation and targeting).[221]** Populist politicians are more likely to use social media as their communication channel to the public than television talk shows.[222]

This problem is only superficially a problem of the media; it also pertains to **the new structure of public discourse, which shapes power relations in society as well. The vehicles of the new publicity are a new**

[218] Habermas, Jürgen: The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society. MIT Press, 1991 (Sixth edition).

[219] It should be noted, that according to Eurobarometer statistics, only 7% of the internet using respondents said they contributed to online political discussions, while 25% follow political topics in a passive way. A staggering 59% do not follow neither participate in political debates online. EBS 477. Democracy and elections. 2018. September.

[220] Digital Hydra: Security Implications of False Information Online. Riga, November 2017. https://www.stratcomcoe.org/digital-hydra-security-implications-false-information-online, at 7.

[221] Ernst, Nicole - Sven Engesser, Florin Büchel, Sina Blassnig and Frank Esser (2017) Extreme parties and populism: an analysis of Facebook and Twitter across six countries. Information Communication and Society. 20:9, 1347-1364.

[222] Ibid.

genre of actors: 'platform providers', which design the structure of communication but which are not responsible for how people use this structure. Finally, in the big picture it signals a **general crisis** of the democratic processes, of which the **public sphere** is a crucial element.

It should be noted that the impact level of disinformation and propaganda depends – among many elements – on two main variables:

1) the pluralism of media and of ideas within the particular media landscape. A strong and lively public discourse – which presupposes **media freedom and pluralism** – makes the audience more resistant to disinformation and propaganda actions. When such content can be confronted with rational argumentation and various opposing views, from several sources, then the audience has a greater chance of discerning the truth. Therefore, those countries where media freedom and pluralism are endangered, typically where public service broadcasting or even a large part of the media is captured, are more vulnerable to experiencing the negative effects; and

2) the **level of organisation of the disinformation and propaganda campaign.** Sporadic disinformation events are not comparable with concerted campaigns using a variety of sources and communication channels at the same time. Such concerted efforts must be well-financed, which offers another anchor for investigation and control. Governmental actors are known to have larger financial resources at their disposal. And when the system of checks and balances is weak or deconstructed, there are slim chances of resisting a decay of democracy.

It is suspected that **when governmental actors apply disinformation and propaganda methods to political communication within a society, where media freedom and pluralism has already been limited**, then the targeted society has almost **no chance of resisting the influence.** Moreover, if governmental actors apply disinformation and propaganda within their domestic territory, they can evade legal responsibility domestically. Disinformation combined with media capture and state impunity can be regarded as **early warnings of rule of law backsliding**.

2.1.1 New features of the online media environment

Below we list features of the new public sphere that make it prone to exploitation by populistic political communication.

2.1.1.1 The control vacuum

Traditional media had a built-in 'social filtering' mechanism in its own hierarchical structure: only those views could be published that passed the scrutiny of editors and other participants in the process, including media owners. The scarcity of resources (and the consequent regulation) worked as a **natural filter** for the mass electronic media. Content was mostly created by media **professionals**, or was at least coordinated by studio anchors.

The **central and vertical** organisation made it possible for the state to regulate the gatekeepers (media outlets) in order to enforce national rules on publicly disseminated content. With respect to national media markets, the number of gatekeepers was smaller,[223] and they had to obey the stricter rules applying to electronic media. **Traditional mass media was vertical, centralised in its dissemination, solely one-way and elitist in its content**. Even though the commercial mass media and tabloid press were frequently accused of transgressing ethical rules, and they targeted non-elite social groups, their owners were embedded in the political and economic elite of society and the entertaining content was still filtered through the main value system of the

[223] According to some views, the number of gatekeepers is smaller today – provided that as gatekeepers are meant not the media outlets but the social media platforms and possibly search engines. Rasmussen, Terje: Internet and the Political Public Sphere. In: Sociology Compass. 8/12 (2014): 1315–1329, DOI 10.1111/soc4.12228. page 1320.

social elite. As a result, **society's information basis was hierarchically** organised and mediated by the mass media or other authorities in society, such as the church, academia and doctors.

The social media discourse realises a **horizontal interchange of ideas** that is not directly influenced by the social elites; state control is more limited than before, while authoritarian powers and informal political groups are exploiting this **'control vacuum'** to influence and manipulate the public discourse.

In addition, this new communication is cross-border, and the democratic public discourse is not isolated from foreign interference. What is more, AI instruments – which can be used by anyone with the financial means to operate them – influence the public discourse.

2.1.1.2 The hyper-democratic nature of social media

Underrepresented communities often rejected the ideals represented by the mainstream media, but their voices were excluded from the mainstream.[224] In contrast, in our networked society all citizens have the ability to engage and express their views without any entrance barriers such as literacy: due to its simple design, dominance of pictures and videos, along with short messages (with even autocorrect and style recommendations), it enables anyone with a smart device to let their voice be heard,[225] as noted sarcastically by Umberto Eco.[226] Social groups that used to be underrepresented in the traditional media are now empowered by social media: with their posts, but especially with their comments, likes and shares they can amplify content that used to be suppressed in the traditional mass media. Those users who are **least resilient** to manipulation, disinformation and propaganda are also among them. Organised disinformation campaigns exacerbate these gaps in media literacy as well as cultural and economic capital,[227] and **weaponise the likes and shares** of people susceptible to manipulation in order to amplify their content. In this way the "attention politicians" harvest the users' attention without even investing in advertising space or dissemination techniques.[228]

Box 1: Empowered citizens or exploited victims?

The same group of people can be regarded as **powerful, i.e. those who induce changes, or as vulnerable, i.e. susceptible to manipulation and to being weaponised** by populist politicians for their own advantage. The two scientific perceptions nevertheless describe the same phenomenon: **political campaigns can build on interactive social media users who possess the individual right to represent ideologies that would not be tolerated by the mainstream media, and who, whether consciously or not, advance the case of populism and autocracy in European democracies.**

On the other hand, the latest results of the Eurobarometer survey show that only 7 % of internet users actively follow and contribute to online discussions of political topics on online social networks during election periods,

[224] Caplan, Robyn and Danah Boyd: Who Controls the Public Sphere in an Era of Algorithms? Mediation, Automation, Power. 05.13.2016. https://datasociety.net/pubs/ap/MediationAutomationPower_2016.pdf at page 8.

[225] Bertin Martens, Luis Aguiar, Estrella Gomez-Herrera Frank Mueller-Langer: The digital transformation of news media and the rise of disinformation and fake news. An economic perspective. JRC Digital Economy Working Paper 2018-02. https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf

[226] "Social media give the right of speech to legions of imbeciles who previously spoke only at the bar after a glass of wine, without damaging the collectivity" – Umberto Eco told journalists in a conversation. https://www.pcworld.com/article/2938832/twitter-a-trap-for-italys-communications-gurus.html

[227] Lessenski, M. (2018). *Common sense wanted. Resilience to 'post-truth' and its predictors in the New Media Literacy Index* 2018. Retrieved from the Open Society Institute website: http://osi.bg/downloads/File/2018/MediaLiteracyIndex2018_publishENG.pdf

[228] See also: Hendricks, V.F. - Vestergaard, Mads: Reality Lost. Markets ofo Attention, Misinformation and Manipulation. Springer, 2018. (Open Access) at p. 32.

and only 25 % follow these topics and discussions passively.[229] This sends a strong message that a large majority of society is passive politically on social networks. Further research is needed on the constitution of these groups, and also what exactly qualifies as political discussions. For example, would the burkini or the issue of daylight saving time be considered political issues by the respondents?

2.1.1.3 New messaging technologies and platform polarisation

Social media at least provides some insight into the content circulated. But private messaging apps, such as WhatsApp, and newly emerging social apps, such as Everdays, enable **private communication between large groups of people**. WhatsApp messages have a history of circulating incitement to violence and effectively leading to atrocities. They played a significant role in the run-up to the presidential elections in Brazil (see subsection 1.3.4). Group chats on WhatsApp enable the sharing of **encrypted** messages with up to 256 people at a time – and as most users are members of several groups, messages spread virally. According to Nahema Marchal, a researcher at the Oxford Internet Institute, this service is analogous to **small versions of town squares,** where fake news can easily be shared.[230]

Social apps like the benign Everdays encourage the creation of groups among communities, potentially recommending new members based on personal data that is provided by the users.[231] While Everdays is an app to help the grieving, the same technology could organise groups along any social interest, including political preferences, as happened in the case of WhatsApp.

There is a risk that once social media platforms become safe, disinformation and propaganda will seek other, unregulated channels. This is a threat to democratic stability from two perspectives: (i) the opaque communication activity that still reaches masses of people; and (ii) an even deeper fragmentation of society into various social media platforms which collect people with the same worldview and which completely alienate others with different opinions.

The recommended tool to prevent these harms is **platform-neutral regulation,** especially in the fields of **data protection, political and public issue advertising**, interoperability, and requirements for transparency by operators of mass communication services. In addition is the need to increase media literacy – also with regard to the ethical principles of publishing and sharing content – for all age groups.

2.1.1.4 A new category of gatekeepers

The role of social media platforms is not comparable with earlier actors in the media industry. But we can find analogies in the networked economy: Über, eBay, Airbnb, Tinder, Google Search and so on, all **aggregate and classify information**, making it possible for customers to get to the service directly without a human agent.[232] Regulation could not come to terms with (literally could not *even find a legal term/definition of)* these aggregators, which hinders conceptual thinking of their roles and responsibilities.[233]

While various platform providers exist, at the time of writing Facebook provides the widest degree of flexibility to its users, because it embraces the most diverse themes and aspects of human life: social, political, commercial,

[229] Special Eurobarometer 477. Democracy and elections. Sept-Nov. 2018. http://ec.europa.eu/commfrontoffice/publicopinionmobile/index.cfm/Survey/getSurveyDetail/surveyKy/2198

[230] https://www.nbcnews.com/tech/tech-news/how-whatsapp-became-linked-mob-violence-fake-news-why-it-n929981

[231] https://everdays.com/partners

[232] Helberger, Natali - Jo Pierson, and Thomas Poell: Governing online platforms: From contested to cooperative responsibility. In: THE INFORMATION SOCIETY 2018, VOL. 34, NO. 1, 1–14 https://doi.org/10.1080/01972243.2017.1391913

[233] Various terms have been used to describe various groups of these actors, see more in chapter 3.1.

artistic, professional and family, and it cultivates almost all content formats, such as text, pictures, videos and links. This versatility is also reflected in its popularity: despite all the scandals it still leads all statistics in 2018.[234] But, no matter how powerful social media platforms may appear, they are not content producers; their role is limited to conveying and facilitating communication, with the added value of amplification. Making them responsible for deciding on transmitted content **would give them greater power than necessary,** because it would give them discretionary power to decide over citizens' speech. Without disputing the need to remove manifestly illegal content, we argue that the fundamental right **to free speech would best be served if platform providers were obliged to remain neutral intermediaries, make their algorithmic principles transparent and be subject to supervision.**

2.1.1.5 Community function of social media

The new communication style favours short messages, which tend to be **simplistic, emotional** and **surprising**. Research has supported the hypothesis that **fake news**, because of the mentioned characteristics, enjoys more popularity and **gets more likes and shares** than rational, truthful information.[235]

According to Michael Schudson, the public sphere was never based on purely rational independent debate.[236] **Communication's 'tribal' function** has always been at least as much of a determinant in society as the transmission of ideas. This group-creating function is perfectly realised by Facebook, which urges its users to post information about how they feel, what they read, etc. Its main purpose is to get people to **experience a 'sense of community'**. This function places communication back into the setting of the village (a 'global village').[237] People consume news and information primarily to reaffirm their connection to **a narrative about the world and their place in it**.[238]

Common narratives contribute to the construction of identity and social values, which promote cohesion and solidarity.[239] While the centrally forged and pushed narratives are associated with authoritarianism, democracies also need common narratives, including a **shared information basis and shared basic values within society**. In liberal democracies respect for human rights and minorities, and respect for the rule of law are among the dominant narratives. Some **governments are actively working on finding an alternative to liberal democracy,[240] therefore on replacing the liberal narrative of the mentioned values with their narratives relating to national pride, the demand for security, building up enemies**, etc. **The traditional media system**

[234] In 2018, Facebook leads all statistics with a great edge. WhatsApp - also owned by Facebook - is among the first few as well. See https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/, http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/, https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/

[235] "falsehood diffused significantly farther, faster, deeper, and more broadly than the truth in all categories of information". Vosoughi, S. - Deb Roy, Sinan Aral: The spread of true and false news online. Science 09 Mar 2018: Vol. 359, Issue 6380, pp. 1146-1151. DOI: 10.1126/science.aap9559. See also: Measuring the reach of "fake news" and online disinformation in Europe, Reuters Institute https://reutersinstitute.politics.ox.ac.uk/our-research/measuring-reach-fake-news-and-online-disinformation-europe/.

[236] Schudson, Michael: Was There Ever a Public Sphere? If So, When? Reflections on the American Case. In: Calhoun, C. (ed.) Habermas and the Public Sphere. MIT Press. 1992. 143-163.

[237] McLuhan, Marshall: Understanding Media. The extensions of Man. MIT Press, 1994 [1964]. See also: McLuhan, M.: The Gutenberg Galaxy. Toronto, 2011 [1962].

[238] Wardle, Claire - Hossein Derakhshan: Information Disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe report DGI(2017)09. At page 78.

[239] See also: Inglis, Fred: Media Theory. Blackwell, 1990. Pp. 179-180.

[240] "Orban said the European parliamentary vote must prove that there was an alternative to liberal democracy", " 'We are facing a big moment: we are saying goodbye not simply to liberal democracy ... but to the 1968 elite,' he said. Hungarian PM sees shift to illiberal Christian democracy in 2019 European vote. July 28. 2018. https://www.reuters.com/article/us-hungary-orban/hungarian-pm-sees-shift-to-illiberal-christian-democracy-in-2019-european-vote-idUSKBN1KI0BK. See also: Breaking down democracy. *Chapter 5: Illiberal Democracy*. The Rise of 'Illiberal Democracy'. https://freedomhouse.org/report/modern-authoritarianism-illiberal-democracies

regularly recreated the dominant narrative[241] and public service media still does so, but social media provides space for those opinions that would not be tolerated by mainstream media.

Box 2: The meaning of narratives in communication theory

The narrative paradigm was developed by Walter R. Fisher in 1978, which he further clarified, elaborated and developed later until 2009. According to this paradigm, human culture and values are largely conveyed by storytelling, including news reporting, rather than rational argumentation and discussion of theories. Especially when it comes to combatting issues in the public sphere, narratives can have greater influence than rational argumentation.

Although the narrative paradigm has also been criticised on a theoretic basis,[242] its concept became widely used in communication theory. In the context of this study, the authors rely on its assumption that **not all human communication is entirely rational**, and not always based on arguments. Especially popular communication is to a great extent moral,[243] with the **central elements of emotion and community.[244] Societies express and shape their social values through narratives. The sources of historical narratives are difficult to uncover. The recent narratives have partly been formed by popular culture, partly by the media itself and partly by political actors (but the media plays a central role in all cases).** [245] **Besides reflecting existing social values, they also shape and amend** these to a slight extent. Slight it must be, because a narrative with values detached from existing social values would be perceived as fake.[246]

Emotions have been playing an increasing role in political communication as also described in "The Permanent Campaign" by Sidney Blumenthal.[247] According to Carey, the purpose of communication is partly **ritual**: to construct and maintain a community, represent commonly shared ideas and create a symbolic order that is meant to reinforce, rather than inform.[248]

Spin doctors, speakers and strategic advisers have been dominating political discourse. As the complexity of public issues grows, decisions (like voting) are based more on emotions and social identity, as opposed to reasoned argumentation. This process has been described by Francis Fukuyama as well. [249]

2.1.1.6 Information bubbles

The user perception of the online media environment is that it offers an unlimited variety of information and ideas. This **diversity is just illusory:** different news items may have common sources, users are confined to their

[241] Critical media theory research provides evidence that traditional media, even if sometimes critical with the government, at large supported the system of governance, the dominant narrative, and the "ideology" of liberal democracy. (See: Chomsky-Herman: Manifacturing Consent, 2002)

[242] For example: Barbara Warnick (1987) The narrative paradigm: Another story, Quarterly Journal of Speech, 73:2,172-182, DOI: 10.1080/00335638709383801. See also: Fisher, W.R. (1995) Narration, knowledge and the possibility of wisdom. In. Rethinking Knowledge: Reflections across the Disciplines. by Goodman, Fisher W.R. (eds.) at p. 170.

[243] Walter R. Fisher (1984) Narration as a human communication paradigm: The case of public moral argument, Communication Monographs, 51:1, 1-22, DOI: 10.1080/03637758409390180

[244] McLuhan, Marshall: Understanding Media. The extensions of Man. MIT Press, 1994.

[245] Partly emerging in popular culture, narratives represent the values, identity and morals through entertainment. Among others, thousands of American cinematographic and television productions received funding from Pentagon (Curran, 2011, 140-158.) - which is of course picked up and amplified by RT.com "The Pentagon & Hollywood's successful and deadly propaganda alliance", 2018 Mach 12.

[246] See the notion 'narrative fidelity' at Fisher, Walter R. Fisher (1984) Narration as a human communication paradigm: The case of public moral argument, Communication Monographs, 51:1, 1-22, DOI: 10.1080/03637758409390180 at 272.

[247] Blumenthal, Sidney: Permanent Campaign. 1980. Beacon Press.

[248] Carey, James W. (1989): *Communication as Culture. Essays on Media and Society*. New York & London: Routledge.

[249] Fukuyama, Francis: Identity: The Demand for Dignity and the Politics of Resentment. Farrar, Straus and Giroux, 2018.

filter bubbles and micro-targeting addresses one person but not that individual's neighbour (discussed below in more detail). The bubble effect may be partly due to human cognitive factors that drive users to aggregate in echo chambers supporting their favourite narrative (confirmation bias).[250] Content selection algorithms assist users in reaching this desired goal, and through micro-targeting further splinter the bubbles. The effect is a **fragmented public sphere**.

In the age of legacy media, the audience – even if not consuming the same content – could have an overview of the available content on offer. Citizens, in order to exercise their right to make an informed decision when voting, need a **common basis of information on public matters, including the entire political programmes of political parties.** "In a well-functioning democracy, people do not live in [an] echo chamber or information cocoons."[251] Micro-targeting deliberately limits the audience of a certain content, in order to raise the likelihood that it gets the attention of a certain part of the audience. In commercial matters, the practice is relatively accepted, as it may reduce the number of irrelevant advertisements that we encounter day by day. However, in the democratic public discourse, micro-targeting intentionally deprives those citizens who are not addressed from information directed at the target group – for example, enabling political parties to share only those fragments of their political programmes with the targeted voters these would be likely to support.[252] Beyond being an unfair practice, this splinters the shared information basis of society, **reduces understanding** between people with different beliefs and **exacerbates polarisation**. The increasing polarisation affects friendships and family relationships in countries torn by such manipulative propaganda, most recently the US.[253] Micro-targeting disinformation or propaganda causes **double harm: partly to those who receive the disinformation or propaganda, and partly to those who do not and thus are unaware of what content their fellow citizens are exposed to.** Beyond this, relevant micro-targeting in political issues is usually **based on special categories of (sensitive) personal data relating to the political or worldview preferences** of the targeted persons.

Recent studies have shown that some algorithms are designed to increase diversity with the goal of increasing user satisfaction. Recommendation systems usually combine several aspects, including the surprise element: serendipity.[254] This supports the assumption that **algorithms could be tailored to counteract the unwanted processes** which generate partly naturally from networked communication, like polarisation and echo chambers, and also that users could be offered the option to pick the aspects they would like to combine in their personal algorithm package.

Research also shows that most social media users are embedded in ideologically diverse networks, and exposure to political diversity has proved to have a positive effect on political moderation. The data suggest that social media usage involving participation in several networks reduces mass political polarisation and echo chambers.[255]

[250] See also: Bessi a, Zollo F, Del Vicario M, Puliga M, Scala A, Caldarelli G, et al. (2016) Users Polarization on Facebook and YouTube PLoS ONE 11(8):a0159641. doi:10.1371/journal.pone.0159641.

[251] Sunstein, Cass: #Republic: Divided Democracy in the Age of Social Media. 2007. Princeton University Press.

[252] See more in: See: Zuiderveen Borgesius, F.J., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., Bodo, B. and de Vreese, C., 2018. Online Political Microtargeting: Promises and Threats for Democracy. *Utrecht Law Review*, 14(1), pp.82–96. DOI: http://doi.org/10.18352/ulr.420

[253] "113 million people [...] think that the 2016 presidential election impacted their relationships with loved ones who support the opposing party." https://www.refinery29.com/en-us/2017/11/182043/talking-about-politics-with-family-friends-tips-holidays, See also https://www.psychologytoday.com/us/blog/lifetime-connections/201711/do-you-dare-talk-politics-family-holiday-gatherings and dozens of other sources.

[254] Möller, Judith - Damian Trilling, Natali Helberger & Bram van Es (2018): Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity, Information, Communication & Society, DOI: 10.1080/1369118X.2018.1444076.

[255] Bertin Martens, Luis Aguiar, Estrella Gomez-Herrera and Frank Mueller-Langer, The digital transformation of news media and the rise of disinformation and fake news - An economic perspective; Digital Economy Working Paper 2018-02; JRC Technical Reports. https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf, 27. oldal

2.1.2 Dissemination strategies

At the technological level social media provides a level playing field for all social actors to let their voice be heard; but in reality, a variety of tools can be utilised to amplify messages.

Marketers in the advertising industry have developed **sophisticated psychological and technological methods** of learning users' preferences in detail, categorising them into narrowly defined groups and targeting them with precisely designed communication practices. **Social media results provide immediate feedback and allow the constant adaptation of the algorithms – offering an excellent arena in which to study human behaviour and to design tailored manipulative campaigns**. The personal information may be used to find the most vulnerable persons and touch upon their vulnerabilities.[256] This practice provokes human rights concerns when used for commercial purposes, but **gravely violates human rights and democratic values when used in conjunction with political communication.**

Political bots can send out messages on a significantly larger scale than ordinary human users would, while human trolls spend hours working hard at sending strategically formulated messages as their main activity.[257] When these actors operate from concealed identities – as they usually do – the **individual users are misled regarding the nature of the communication in which they are engaging**. Such **a fraudulent activity is mixed with active user participation** through sharing and liking – this blend creates a sense of uncontrollable anarchy for potential supervisors and legislators.

2.1.3 The realignment of the media landscape

When we mention the drastic change of the public sphere, we should note that both elements, the media environment and citizens' communication practices, have dramatically changed. While the two are clearly intermingled, below we highlight some important aspects of the media environment, acknowledging that the two cannot be treated entirely separately.

As soon as online journals appeared, print media was shattered and lost its traditional income sources. Many of the previously strong and successful newspapers have found new business models, but even more had to close their doors – layoffs dominated the industry in past decades, causing a **crisis of journalism**. Distribution methods have changed considerably and are diverse; notably, social media distribution does not generate revenues for media outlets. Only 32 % of users access their online newspaper sites directly and at least 23 % access them through social media.[258]

The content-overladen and hyper-pluralistic public sphere has generated a culture of relativism: every fact and opinion can be found online and so can their counterparts. Statistics show that traditional media is still relatively trusted by users, in contrast with online media (63-70 % vs 26-27 %).[259] But many users do not distinguish consciously between well-established, high-quality news sites and ephemeral, sensational misinformation or disinformation sources. This is also illustrated by previous results on media trust. When no distinction was made in the questionnaire between online and offline media, the results were significantly lower: only 34 % trusted "the media" and 61 % did not trust it, although there were significant differences between the

[256] O'Neil, Cathy: Weapons of Math Desctruction (2016) Crown Publishing, New York.

[257] Bradshaw, S. and P. Howard (August 2017) Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/07/Troops-Trolls-and-Troublemakers.pdf

[258] Bertin Martens, Luis Aguiar, Estrella Gomez-Herrera and Frank Mueller-Langer, The digital transformation of news media and the rise of disinformation and fake news - An economic perspective; Digital Economy Working Paper 2018-02; JRC Technical Reports. https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf

[259] They trust in news and information they receive through radio (70%), television (66%) and printed media (63%). However, less than half (47%) trust online newspapers and magazines, and lower proportions trust video hosting websites and podcasts (27%) and online social networks and messaging apps (26%). Flash Eurobarometer (EBS) 464, Fake news and disinformation online. April 2018. These results show a contradiction with the results in 2017 which showed 34% trusted the "media", and 61% not trusted.

Member States, from 15 % to 61 %.[260] Even though high-quality and trustworthy news channels and papers still exist, these typically will not be accessed by those groups of users who are the most vulnerable to manipulation.

The established media outlets in democratic countries usually maintain proper journalistic standards. Dubious sites that regularly engage in disinformation would not register with journalistic associations, although this assumption is subject to further research. As a third genre, governmentally-sponsored propaganda media from the illiberal side of the political spectrum would once again not submit to journalistic self-regulation. Therefore, the actions of the established media and of media self-regulation are limited, but not entirely useless: they can frame the issue of disinformation and serve as trustworthy news sources.

2.2 The distortion of democratic processes

Political propaganda and disinformation have been with us since democratic states came into being. In 480 BC the Athenian naval commander Themistocles spread a series of – what we would call today – fake news, suggesting that many of the Greek troops that had joined the Persians in the wake of their successes were unreliable and on the verge of revolt. When the news reached Persian ruler Xerxes, he chose not to deploy these troops.[261] Julius Caesar "was a master propagandist … He understood the need to use such symbols of power and sophistication as a means of converting subject populations to the Roman way of life. This was far less expensive than maintaining elaborate garrisons of legionnaires and induced obedience."[262] In the US, the Federalist Papers written by the founding fathers – under the pseudonym Publius – were also a means of successful political propaganda around the adoption of the federal Constitution to replace the Articles of Federation, in order to persuade the citizens of states to accept and ratify the document.[263] (See more on the definitions of disinformation and propaganda in section 1.1.)

Political persuasion by propaganda poses problems with regard to both procedure and outcome. Even if democratic processes and the dictates by the rule of law and fundamental rights are followed, **if democracy is not militant enough, it may lead to the acquisition of power by non-democratic forces** (Table 8, scenario B). If, however, democracy is manipulated and the techniques of persuasion do not follow the rule of law and infringe fundamental rights, the procedure itself might be a threat to democracy, albeit it will not automatically result in democratic decline. Procedures may be violated by democratic forces, which will, beyond distorting election procedures in all other aspects, respect the concept of democracy (scenario C). Yet more often manipulations of procedure aim at overthrowing democracy itself (scenario D).

[260] Eurobarometer 461. Designing Europe's Future, April 2017. EBS 464 also showed differences, the largest in the television sector: 28% vs. 74%

[261] Jowett, Propaganda Through the Ages, 49-92, 51-52.

[262] Jowett, Propaganda Through the Ages, 49-92, 55.

[263] Martin J. Manning, Herbert Romerstein, Historical Dictionary of American Propaganda, 101-102.

Table 8: Distortions in democratic processes and their consequences

		Outcome	
		Democracy upheld	**Democracy challenged**
Process	Democratic processes, the rule of law and fundamental rights are **respected**	A. Ideal case *Example*: Clean elections result in a democratic change of government.	B. Tools of militant democracy fail to prevent an undemocratic power from gaining support and overthrowing the liberal democratic order once it wins an election. *Example*: A political party that should have been banned wins the elections.
	Democratic processes, the rule of law and fundamental rights are **manipulated**	C. Manipulation of democratic processes benefit a democratic power or does not achieve its aims. *Example*: Gerrymandering is conducted by an otherwise democratic government or it loses elections despite the use of chatbots.	D. Undemocratic tools or means in violation of the rule of law and fundamental rights are used to overthrow liberal democracy. *Example*: By violating privacy and data protection laws, voters are profiled and targeted with fake news, resulting in an authoritarian overtake of power.

Source: Authors.

2.2.1 The triangular relationship between democracy, the rule of law and fundamental rights

Democratic processes, the rule of law and human rights are not only values per se, but they also serve the purpose of keeping liberal democracy alive, to make sure that all, including forces in power, abide by the law and arbitrariness is prevented. This corresponds to the triangular nature of democracy, the rule of law and fundamental rights, where these three values are "inherently and indivisibly interconnected, and interdependent on each of the others, and they cannot be separated without inflicting profound damage to the whole and changing its essential shape and configuration".[264] **The rule of law without democracy is a contradiction, while democracy without the rule of law may turn into the dictatorship of the majority.**[265] Also, fundamental rights are closely interlinked with the other two values. The relation between freedom of expression, the right to receive information and informed participation in a democracy clearly illustrates this interdependency. One is capable of making informed choices during elections, or meaningfully participating in public debates only in the

[264] Carrera, S, Guild, E & Hernanz, N 2013, 'The Triangular Relationship between Fundamental Rights, Democracy and the Rule of Law in the EU, Towards an EU Copenhagen Mechanism', Study, CEPS, Brussels.

[265] According to the Venice Commission's Rule of Law Checklist, "The Rule of Law promotes democracy by establishing accountability of those wielding public power and by safeguarding human rights, which protect minorities against arbitrary majority rules." At least some of the elements on the checklist, including – legality, legal certainty, prevention of abuse or misuse of powers, equality, and access to justice – will be hampered in a non-democratic system. Legality for example may be infringed, when there are not sufficient checks on the executive's lawmaking powers, or the law-making procedures are distorted. European Commission for Democracy through Law (Venice Commission), Rule of Law Checklist, 18 March 2016.

And *vice versa*: the rule of law is mentioned in the Preamble to the Statute of the Council of Europe as one of the "principles which form the basis of all genuine democracy".

For more details see S. Carrera, E. Guild and N. Hernanz (CEPS), *The Triangular Relationship between Fundamental Rights, Democracy and the Rule of Law in the EU, Towards an EU Copenhagen Mechanism*, Study for the European Parliament, 2013, available at http://www.europarl.europa.eu/RegData/etudes/etudes/join/2013/493031/IPOL-LIBE_ET%282013%29493031_EN.pdf possession of knowledge about the facts. As proven by influential thinkers such as John Rawls,[266] the idea of deliberative democracy[267] presupposes exchange of information conducted along a fair procedure and a discussion based on reason, instead of self-interest or political power. According to this theory, the focus is more on the process or its legitimacy, and the source of its coercive power than the outcome. Citizens should be persuaded by the force of the more reasonable or better justified argument than private interests, biases, prejudices or views that cannot be convincingly explained and proven. As Habermas claimed, "the democratic procedure for the production of law evidently forms the only postmetaphysical source of legitimacy".[268] Only those laws can be regarded as legitimate that have been adopted based on the agreement of all citizens, in a fair and discursive process equally open to all.[269] Such procedures will not necessarily generate consensus, nor will they lead to the truth or even be just, but instead they will lead to results that are fair and reasonable, and that can also be subjected to revision if new information and further deliberation warrants so. The more the elements of a deliberative democracy are met, the more justifiable the outcome is likely to be, and the more legitimate the democratic process will be.[270]

Participatory and deliberative democracy require the provision of information and effective processes of consultation. Meaningful participation in a debate is only possible if individuals know what public policy is, how they are affected and what alternative solution there is.[271]

This is also translated into the rights language. According to its European understanding, freedom of expression is integral to discovering the truth, and individuals have to access potentially relevant information held by the state to enable their informed participation in democracy. In other words **the passive side of free speech incorporates the public's right to receive information**.[272] When developing its arguments in favour of the right to receive information, the European Court of Human Rights (ECtHR) noted that nearly all of the 31 member states of the Council of Europe surveyed have enacted legislation on freedom of information, and furthermore states agreed on the Convention on Access to Official Documents,[273] which are further proof of a common ground.[274]

> [I]n circumstances where access to information is instrumental for the exercise of the applicant's right to receive and impart information, its denial may constitute an interference with that right. The principle of securing Convention rights in a practical and effective manner requires an applicant in such a situation to be able to rely on the protection of Article 10 [of the European Convention on Human Rights on freedom of speech].[275]

[266] John Rawls, 'The idea of public reason revisited', The University of Chicago Law Review, vol.64, no.3, 1977.

[267] It is not the objective of the present paper to give a full account of deliberative democracy, but rather to highlight elements that are relevant for the EU constitutional analysis of disinformation. For authoritative texts on the matter see for example James Bohman, Deliberative Democracy (Cambridge, MA.: MIT Press, 1998).

[268] Jürgen Habermas, Between facts and norms: contributions to a discourse theory of law and democracy, Cambridge: MIT Press, 1996, 447-8.

[269] *Id.*, 107.

[270] Gutmann, A. and D. Thompson (1996). Democracy and disagreement. Cambridge: Harvard.

[271] ECtHR, Leander; Gaskin v United Kingdom, (1990) 12 E.H.R.R. 36, para 49; G. (M.) v United Kingdom, (2002) 36 E.H.R.R. 22, para 30; Guerra v Italy, (1998) 26 E.H.R.R. 357, para. 60; Oneryildiz v Turkey (2005) 41 E.H.R.R. 325, paras. 89-90, 108; McGinley and Egan v United Kingdom (1998) 27 E.H.R.R. 1, paras 101-2; LCB v United Kingdom (1998) 27 E.H.R.R. 212, paras 38-9 and Roche v United Kingdom (2006) 42 E.H.R.R. 599, paras. 165-6; Társaság A Szabadságjogokért v. Hungary, 14 April 2009; Guseva v. Bulgaria

[272] ECtHR, Bladet Tromsø and Stensaas v. Norway [GC], §§ 59 and 62, Sdruženi Jihočeské Matky v. the Czech Republic (dec.) (2006), Társaság a Szabadságjogokért v. Hungary (2009)), Youth Initiative for Human Rights v. Serbia (2013), and Österreichische Vereinigung zur Erhaltung, Stärkung und Schaffung v. Austria (no. 39534/07, 28 November 2013

[273] Council of Europe Convention on Access to Official Documents, Tromsø, 18 June 2009.

[274] ECtHR, Case of Magyar Helsinki Bizottság v. Hungary, Application no. 18030/11, 8 November 2016

[275] *Id.* at § 155.

As Ignatieff put it, "[o]ne of the things about a democracy that people forget is how important knowledge is ... If you don't have knowledge then all you get to go on is tweets and Facebook, and rumors and fantasy and paranoia ... You need knowledge in order to make choices."[276] Whereas case law so far has only addressed information retained by the government, the tsunami of irrelevant information, disinformation and manipulative propaganda poses very similar questions to the problems addressed by Strasbourg jurisprudence thus far.

In transmitting knowledge, the free press,[277] civil society[278] and academia[279] play a crucial role. Little wonder that states engaging in rule of law backsliding[280] and destroying democracy are capturing the media,[281] shrinking the space for NGOs[282] and curtailing academic freedom.[283] The ECtHR also acknowledged the "censorial power of an information monopoly" when public bodies fail to release information requested by certain entities.[284]

But distorting avenues of access to knowledge is not only a characteristic of states in rule of law backsliding. There are certain phenomena, such as online media and social networks (see more in section 2.1) and the postmodern risk society (below), whose interplay leads to distortions of freedom of expression and participatory democracy.

2.2.2 The dilemma of speech regulation vs freedom of expression in a flexible online environment

With the emerging models of mass communication citizens are no longer passive recipients of news, and initially with the expansion of internet access and social media usage there was hope that the production and distribution of information would be replaced with a decentralised, and therefore more democratic, system of discourse. Over time the debate took a U-turn, exploring whether the new technologies fundamentally endanger democracy as we know it. As Sajó noted,

> instead of creating a common space for democratic deliberation, the internet and social media enabled fragmentation and segmentation. Discourse is limited to occur within self-selecting groups and there are tendencies of isolation. Views are more extreme and less responsive to external arguments and facts, resulting in polarization around alternative facts.[285]

[276] See: https://www.youtube.com/watch?v=FoMd_Bxn5rg&feature=youtu.be&fbclid=IwAR2uHR_3m4w5DbB1ndIBRbvYHfHT_eIB9_V9kaaXvfi7_X16dDYclajnDEg Professor Michael Ignatieff is Rector of the Central European University, an entity being specifically singled out and attacked by a piece of Hungarian legislation. There is also a related infringement proceeding pending in front of the Court of Justice, see http://europa.eu/rapid/press-release_IP-17-1952_en.htm

[277] ECtHR, Bladet Tromsø and Stensaas v. Norway [GC], §§ 59 and 62.

[278] ECtHR, Animal Defenders International v. the United Kingdom [GC], § 10.

[279] "There is no Chinese wall between science and a democratic society. On the contrary, there can be no democratic society without free science and free scholars. This interrelationship is particularly strong in the context of social sciences and law, where scholarly discourse informs public discourse on public matters including those directly related to government and politics." See the concurring opinion of Judges Sajó, Vucinic and Kuris in ECtHR, Mustafa Erdoğan and Others v. Turkey, Applications nos. 346/04 and 39779/04, 27 May 2014

[280] Müller, Jan–Werner. 2013. 'Safeguarding Democracy inside the EU: Brussels and the Future of Liberal Order'. *Working Paper* No. 3 (Washington, DC: Transatlantic Academy).

[281] Petra Bárd - Judit Bayer, A comparative analysis of media freedom and pluralism in the EU Member States. Study for the LIBE Committee, PE 571.376 EN, 2016, http://www.europarl.europa.eu/supporting-analyses.

[282] Małgorzata Szuleka, First victims or last guardians? The consequences of rule of law backsliding for NGOs: Case studies of Hungary and Poland, CEPS Paper in Liberty and Security in Europe, 2018/6, https://www.ceps.eu/system/files/MSzuleka_RoLandNGOs.pdf.

[283] Michael Ignatieff, Stefan Roch (eds.), Academic Freedom: The Global Challenge, Budapest: CEU Press, 2018.

[284] ECtHR, Társaság a Szababságjogokért v. Hungary.

[285] See the Conference e-book by the European Centre for Press and Media Freedom (ECPMF), Promoting dialogue between the European Court of Human Rights and the media freedom community. Freedom of expression and the role and case law of the European Court of Human Rights: developments and challenges, 2017, https://ecpmf.eu/files/ecpmf-ecthr_conference_e-book.pdf, at 5.

However, this in itself is no reason to limit freedom of expression: "communication also means freedom to fake news, a freedom enjoyed and abused by individuals, political movements and governments".[286]

Limiting it poses a number of difficulties, starting with the practical problem that politicians are unlikely to be willing to regulate speech, since they are interested in upholding the "competition to find the most attractive alternative truth". At the same time the question emerges of whether the cure is not worse than the illness, i.e. **whether an interference limiting free speech in the name of democracy would not ultimately result in government monopoly on deciding which views are right and which are to be dismissed as fake.**

But **with the emergence of artificial intelligence and non-human mass distribution of ideas (through chatbots for example), certain views held by real people cannot even emerge in a cacophony of false information. Certain viewpoints may therefore simply be outnumbered and vanish.** That alone might trigger possible state intervention to stop the unfair competition between human-generated versus machine-generated distribution of content, and to enable human voices and opinions to reappear. The **idea is not to censor false information,[287] and this is particularly important in relation to elections,[288] but to create an equal setting for all opinions to be heard.**

2.2.3 Postmodern risk society

Our postmodern world is characterised by political and existential uncertainties – that is partially an inevitable consequence of the lack of an absolute (source of) knowledge, and is partially artificially created. **Knowledge is replaced by the culture of risks,[289] fear creation and punitive populism, i.e. harsher and harsher criminal legislation, allegedly responding to the demands of the people, without impact assessment, efficiency, necessity or proportionality analyses. Symbolic or expressive justice without criminological underpinnings and more generally the over-emotional tone of politics ultimately lead to the establishment of a control society.[290] Control societies employ a complex web of collective strategies through which fear, angst, anxiety, phobia or even hysteria are created and recycled.** Fulfilling the dreams of any businessperson, risks as a "bottomless barrel of demands"[291] are endlessly being identified in ever-newer types, forms and levels of insecurities, whether in the context of migration, terrorism, health or environmental hazards. Insecurities and reliance on beliefs and emotions are fed by lack of knowledge, and thus become self-producible, self-referential and tautological.[292]

[286] *Id.*

[287] Cf. ECtHR, Bladet Tromsø

[288] See e.g. Case of Bowman v. The United Kingdom

[289] Mary Douglas and Aaron Wildavsky, *Risk and culture: an essay on the selection of technical and environmental dangers*, Berkeley: University of California Press, 1982; Ulrich Beck, *Risikogesellschaft: auf dem Weg in eine andere Moderne,* Frankfurt am Main: Suhrkamp, 1986, Anthony Giddens, *The consequences of modernity*, Cambridge: Polity Press, 1995.

[290] David Garland, *The culture of control*, Oxford: Oxford University Press, 2001

[291] For the English terminology see the translation by Mark Ritter: Ulrich Beck, *Risk society: towards a new modernity*, London: Sage Publications, 1992

[292] "Durch die Produktion von Risiken werden die Bedürfnisse endgültig aus ihrer naturhaften Restverankerung herausgelöst und damit aus ihrer Endlichkeit, Erfüllbarkeit. Hunger kann man stillen, Bedürfnisse befriedigen; Risiken sind ein "Bedürfnis-Faß ohne Boden", unabschließbar, unendlich. Anders als Bedürfnisse können Risiken nicht nur […] manipuliert werden. Es können durch wechselnde Risikodefinitionen ganz neuartige Bedürfnisse – und damit Märkte – *geschaffen* werden. […] An die Stelle vorgegebener und manipulierbarer Bedürfnisse als Bezugspunkt der Warenproduktion tritt das *selbstherstellbare* Risiko." "Mit Risiken – könnte man mit Luhmann sagen – werden die Wirtschaft "selbstreferentiell", unabhängig von der Umwelt menschlichen Bedürfnisbefriedigung." Beck, *supra* note 1, 74. The market logics of security expand the forms of control also horizontally. Areas traditionally kept for the relation between the state and the individual are becoming privatized, decentralized and interconnected. For individuals living in the 21st century it becomes increasingly difficult to oversee who controls them, when, on what grounds, in whose interests. When relocating responsibilities for control important segments of state power are shifted to private parties.

Once it has been given up, a rational discourse is close to impossible to re-establish due to the multiple forms of vulnerability of ordinary citizens who are laypersons. In postmodernism **most threats are immaterial and invisible; knowledge about them is mediated through experts and as such are dependent on interpretation.**[293] Both perception and effective regulation in this regard are dependent on highly technical forms of scientific information, which are debated even among scholars in the field.[294] Knowledge has ceased to exist in the original sense,[295] including the notion of certainty of ideas. What we are left with is better or worse ways to interpret contemporary societies. Yet assuming for the sake of the hypothesis that experts manage to agree on scientific truth, the problem of incomprehensiveness for the average layperson voter arises. Even if citizens have access to the scientific knowledge of the day, they may not have the capacity to comprehend it. Instead of introducing scientific, rational, objective elements into the public debate, the dangers of a risk society are augmented. Ignorance turns into angst, "liquid fear"[296] from the yet unknown risks lurking around ready to swoop down any moment as soon as identified.

Deliberative democracies need to respond to public fear, where responsiveness needs to be "complemented by a commitment to deliberation, in the form of reflection and reason giving".[297] The tragedy of a risk society is that there is no room for an identification and solid evaluation of risks or the tools employed to ensure security. The debate becomes a tragedy of errors:[298] its internal logics would even allow an ineffective solution employed against a non-existing problem, with the price of deconstructing the rule of law. But more often there is a seed of truth, which makes disinformation even more likely to be persuasive.

Populistic politicians typically touch upon the voters' real fears and concerns, and appear to represent them sensitively, offering seemingly functional and fast responses. In contrast, democratic decision-making appears to be complicated, distant and inefficient, and requires the investment of time and energy to be meaningfully involved. In short, democratic processes seem to be beyond the reach of ordinary citizens.

The tendencies above threaten democratic processes and the institution of democracy. Instead of alleviating lack of knowledge by social inclusion, education and the empowerment of citizens, and in lieu of generating a meaningful debate, we see an opposite phenomenon on the side of some governments. On the one hand they abuse this ignorance, and on the other create even more insecurities through misinformation, disinformation, mal-information or fake news. Populism gives easy to understand but oversimplified and false answers to complex questions, and emotional politics builds on the worst characteristics of the people, demanding harsher intrusions into human rights and limitations of the rule of law in the name of greater security. As a result of the lack of a valid debate, people become inept in their own lives and the threat of insecurity is exacerbated by the loss of their cognitive sovereignty. Voters **call for strong and intrusive government measures**. **Demands for more certainty and security lead to a lack of limitation of state powers, a quasi-emergency situation.** This is all **opposing the state of a rule of law,** where those in power are supposed to be bound by predefined rules, as anyone else.

2.3 Impact on EU democracy and cohesion

Manipulations of people's access to information may come from within a democracy, from third countries or non-state actors. The multi-level system of the EU adds an extra layer to the complexity of the issue. Both the

[293] Barbara Adam, Ulrich Beck and Joost van Loon (eds.), *The risk society and beyond: critical issues for social theory*, London: Sage Publications, 2000, 3.

[294] Jean-Francois Lyotard, La condition postmoderne: rapport sur le savoir, Paris: Editions de Minuit, 1979.

[295] See Giddens, *supra* note 1, 38-39.

[296] Zygmunt Bauman, *Liquid Fear*, Cambridge: Polity, 2006.

[297] Sunstein, *supra* note 7, 1.

[298] David A. Green, "Public Opinion Versus Public Judgment About Crime: Correcting the 'Comedy of Errors'," 46 *British Journal of Criminology* 1, 131-154 (2006)

manipulation of public opinion over social media platforms and the emergence of risk societies (see subsection 2.2.2) need to be tackled in order to maintain participatory and deliberative democracy. The challenge is that both these phenomena are in the interest of those in power and are to a large extent created or enhanced by the government. The latter issue of risk creation (fear-based communication), including emotionalism, scapegoating and attaching collective responsibility to certain social groups, is well known from European history. But now we witness propaganda also conducted primarily by governments. In Hungary, for example, a country currently undergoing rule of law scrutiny by the EU in the form of an Article 7(1) procedure, exclusively government agencies are exploiting social media platforms to spread disinformation, whereas in other EU Member States it is a mixture of government agencies, politicians and parties, private contractors and civil society organisations.[299] Either way, some **national governments may have a vested interest in not tackling the issue via legal or policy measures**.

In the absence of national measures putting a halt to the above phenomena, in this section we explore the extent to which the EU will be affected by disinformation and propaganda.

2.3.1 Militant democracy: The history of an idea

Our starting point is the concept of militant democracy. Militant democracy is a term of art in constitutional law, coined in 1937 by Karl Loewenstein.[300] He distinguishes constitutional government and dictatorship, which he sometimes equates with the terms emotional government, disciplined or authoritarian democracy. The former is signified by the rule of law, rationality and calculability, the preservation of a well-delineated sphere for public law and respect for fundamental rights. The latter is characterised by legalised opportunism, where positive law is not bound by constitutional legality, but is reduced to unchallengeable command.[301] Loewenstein authored his influential articles in the wake of fascism, but the lessons he draws for constitutional law are instructive even today.

He envisages two possible explanations for authoritarianism as the new model of government: either democracy (which he uses interchangeably with liberal democracy) is doomed and authoritarianism is a “spiritual flame shooting across the borders”,[302] which implies that resistance is a waste of time; or, the authoritarian type of ruling is just a technique of some political elites of not only acquiring power, but also retaining it at all costs. If the latter holds true – which Loewenstein believes to be the case – then democracy is still a valid form of government, and must resist attempts to be overthrown by various techniques, or in other words “democracy must become militant”.[303]

Learning from the politics of the day, he is preoccupied with the in-built tension of a democracy, i.e. that **democratic tolerance can be used to destroy democracy itself.** As Joseph Goebbels infamously stated, “[i]t will always remain one of the best jokes of democracy that it provides its own deadly enemies with the means with which it can be destroyed”.[304] He illustrates this point at length with the lack of militancy (i.e. tools of self-preservation or resilience) of the Weimar Republic, and especially the Weimar constitution, against subversive movements.[305]

[299] Samantha Bradshaw, Philip N. Howard, Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation, 2018, http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct2018.pdf

[300] For a full description see K. Loewenstein, ‘Militant Democracy and Fundamental Rights,’ 31 *American Political Science Review* 417–433 and 638–658 (1937).

[301] K. Loewenstein, ‘Militant Democracy and Fundamental Rights,’ 31 *American Political Science Review* 417–433, 418, 432.

[302] *Id.* at 422.

[303] K. Loewenstein, ‘Militant Democracy and Fundamental Rights,’ 31 *American Political Science Review* 417–433, 418, 428.

[304] Quoted and the original wording translated by András Sajó, Militant Democracy and Transition towards Democracy, in: András Sajó (ed.), Utrecht: Eleven International Publishing, 2004, 209-230, 214.

[305] For a most recent authoritative account of such a function of international legal mechanisms, see R. Dworkin, ‘A New Philosophy of International Law’ (2013) 41 *Philosophy and Public Affairs* 1, 2–30 (2013).

Before listing possible tools of militant or self-preserving democracy, Loewenstein states what the tools of militant democracy do not incorporate – and this is emotionalism. Whereas liberal democracy was something people died for when absolutism was doomed, "democracy *à la recherche d'une nouvelle mystique* seems hopeless, if not ridiculous. Democratic romanticism is of itself a contradiction."[306]

Authoritarianism is held together not by violence, but emotionalism. Yet democracy can never trigger the same intensity of emotions; therefore, one cannot fight fire with fire in this context. In the second part of his paper, Loewenstein lists various forms of militant democracy that might work or which proved to be successful (until 1937) in countering attacks on liberal democracy,[307] concerning these aspects:

- legislative measures that deal with high treason (rebellion, insurrection, armed uprising, sedition, etc.);
- prohibition of subversive movements and parties;
- prohibition of paramilitary organisations;
- forestalling the creation of military bands;
- control over gun use, abuse of parliamentary institutions, incitement to violence or hatred;
- balancing freedom of assembly and public order;
- balancing freedom of speech and press and the slandering or ridiculing of political institutions;
- the practice of morally aiding and abetting political criminals;
- loyalty of the police and public officials;
- training of the police; and
- foreign influence, especially financing of foreign forces trying to destroy liberal democracy.

2.3.2 Militant democracy and constitutional resilience, and their precautionary nature

Loewenstein developed his theory in a special historical context, which is instructive even today in many aspects. In general, a constitutional democracy has to stand up against various forms of emotionalism. As András Sajó pointed out, "radical politics of emotions has a penchant for lying. (What differentiates it from 'ordinary' politics is that it is only capable of existing with lies.)"[308] It must go beyond party banning, not least because it will not be sufficiently effective. Instead it has to forestall misleading the people, insist on rational problem-solving and resist any kind of emotional manipulation.

Mature constitutionalism today implies the existence of robust precautionary measures in democratic systems to make them resilient against a future potential political force, which, when entering into government, would replace constitutional government by an autocratic one.[309]

The tools of militant democracy follow a **precautionary logic**. Once democracy is overthrown, there are slim chances for it to be restored by way of abiding by the rules of limiting government – a notion the undemocratic force will have attempted to abandon in the first place. Rules directed at limiting government will not work with a party that denies the notion of liberal democracy and demolishes checks and balances.[310] Take party banning for example – there has to be a prompt procedure that would have the power to ban the political party in question before it enters government, gains a parliamentary majority or captures the judiciary, all of which would

[306] *Id.* at 428.

[307] K. Loewenstein, 'Militant Democracy and Fundamental Rights,' 31 *American Political Science Review* 638–658 (1937).

[308] András Sajó, Militant Democracy and Transition towards Democracy, in: András Sajó (ed.), Utrecht: Eleven International Publishing, 2004, 209-230, 212.

[309] Otto Pfersmann, Shaping Militant Democracy: Legal Limits to Democratic Stability, in: András Sajó (ed.), Utrecht: Eleven International Publishing, 2004, 47-68.

[310] Cf. Dieter Grimm, How can a democratic constitution survive an autocratic majority?, VerfBlog, 13 December 2018, https://verfassungsblog.de/how-can-a-democratic-constitution-survive-an-autocratic-majority/.

result in impunity. To be efficient, laws should prosecute banned parties as early as possible. The German Federal Criminal Code, for example, sanctions even attempts to maintain a political party that has been declared unconstitutional by the Federal Constitutional Court or which is a surrogate organisation of a banned party.[311] Or recall national criminal laws prohibiting the attempt to overturn the constitutional order by force, or threatening to do so, but not the actual overturn of the constitutional order.[312] The logic is that once it happens, the Criminal Code provision would be of no use anyway.[313]

While recognising the precautionary nature of an efficient toolbox of militant democracy, an important difficulty arises. The tools must not be so cautious that any human behaviour which has the slightest chance of leading to a distortion to democracy would be banned. At first sight the precautionary logic might be at odds with the risk-taking attitude of liberty, but **in order to preserve itself some risk-averse attitude is needed, based on the seriousness of the danger, the likelihood by which it would occur and its immediacy**. A **specific historical experience** in a given jurisdiction might be a further influencing factor, which may trigger measures to halt any attempt of a U-turn towards certain ideologies that had tragic consequences in the past. The **interrelated nature of events** should also be considered, where individual events on their own may not but their cascading nature may lead to the collapse of democracy, the rule of law and fundamental rights.[314]

The question is how far one may go in entrenching these institutions, procedures and rights, or in other words how to be resilient against anti-democratic forces but still able to call a system democratic. Sajó warns us against the claim that militant democracy is a contradiction in terms. Such criticism, he claims, is illusory, if not hypocritical. There is a well-defined difference between disagreeing with certain democratic policies and denying democracy as such as the main process of decision-making.[315] Of course the techniques used must be in conformity with democracy, the rule of law and fundamental rights, but they must be readily available for the state. This is the point where the interrelated triangular nature of democracy, the rule of law and fundamental rights[316] that some of the authors previously argued for becomes apparent. As AG Maduro put it, we need

> to ensure that the political necessities of today do not become the legal realities of tomorrow. [The courts'] responsibility is to guarantee that **what may be politically expedient at a particular moment also complies with the rule of law without which, in the long run, no democratic society can truly prosper**.[317]

But even future-oriented precautionary **tools to enhance resilience should not be overestimated**: "there is no foolproof constitutional design that can immunize liberal democracy from the pressures of backsliding. At best, constitutional design features serve as speed bumps to slow the agglomeration and abuse of political power; they cannot save us from our worst selves completely."[318] Modern tools go beyond speech limitations and party

[311] See Article 84 of the German Federal Criminal Code on the continuation of a political party declared unconstitutional

[312] See e.g. Article 254 of Act C of 2012 on the Criminal Code of Hungary.

[313] Except retroactively making perpetrators criminally liable once the constitutional order has been restored. But as a tool of militant democracy a provision on a completed overturn of democracy is of no use.

[314] *Id.* at 215-217, 229-230.

[315] András Sajó, Militant Democracy and Transition towards Democracy, in: András Sajó (ed.), Utrecht: Eleven International Publishing, 2004, 209-230, 211.

[316] S. Carrera, E. Guild and N. Hernanz (CEPS), *The Triangular Relationship between Fundamental Rights, Democracy and the Rule of Law in the EU, Towards an EU Copenhagen Mechanism*, Study for the European Parliament, 2013, available at http://www.europarl.europa.eu/RegData/etudes/etudes/join/2013/493031/IPOL-LIBE_ET%282013%29493031_EN.pdf, 4–15, and Annex 1 of the study.

[317] Joined cases C-402/05 P and C-415/05 P, Opinion of Mr Advocate General Poiares Maduro delivered on 16 January 2008, Yassin Abdullah Kadi and Al Barakaat International Foundation v Council of the European Union and Commission of the European Communities.

[318] Tom Ginsburg, Aziz Z. Huq, Mila Versteeg, The Coming Demise of Liberal Constitutionalism?, The University of Chicago Law Review, Volume 85, Issue 2 (March 2018) 239–255, 253. See also James D. Best, Constitutional Speed Bumps, http://www.whatwouldthefoundersthink.com/constitutional-speed-bumps

banning, and vary from constitutionally embedded provisions that cannot even be amended by the ordinary process of constitutional amendment (*Ewigkeitsgarantien*) to the development of a constitutional culture.[319] Anything in between, such as fair elections, the rights of the parliamentary opposition,[320] constitutional adjudication, judicial independence, media pluralism, press freedom, civil control and academic freedom, and even the sheer existence of a constitution and a bill of rights may contribute to the survival of democracy. Tomasz Tadeusz Koncewicz identifies three complementary safeguards of democracies: the rule of law and the constitution, trust in the binding power of law, and mechanisms of supranational and international control.[321] It is this latter that will be addressed in subsection 2.3.3, when exploring the role that the EU might play in entrenching the concept of democracy.

2.3.3 The EU as a tool of militant democracy

Should national rules of militant democracy fail, and election laws be curbed, constitutional courts being captured, ordinary judges unduly influenced, media pluralism destroyed, participatory democracy dismantled or civil society harassed, international fora will still be able to remedy deficiencies to some extent.[322] But let us not forget that whereas from a rogue Member State's viewpoint, the tools the EU employs can be seen as external tools of militant democracy, for the EU these same tools are internal.

First, **a state's departure from the European consensus on rule of law standards will ultimately hamper the exercise of individuals' rights EU-wide.** All EU citizens in the given state will be detrimentally affected; furthermore, a **lack of limits to illiberal practices may encourage other Member State governments to follow**, so that rule of law violations become contagious.

What is more, the given state's participation in the EU's decision-making mechanism jeopardises the legitimacy of the EU and its legal instruments and policies.[323] Systemic violations of Article 2 of the Treaty on European Union (TEU) values will **undermine mutual trust-based instruments**, for example in the terrains of EU asylum law and in EU criminal justice.[324] Apart from these substantive problems, the principle of primacy would also be jeopardised. Member States would invoke the human rights argument in order to permit exemptions from the principle of primacy of EU law. The German Federal Constitutional Court (GFCC) is most illustrative for retaining the right to be the ultimate reviewer of EU law in the form of fundamental rights, ultra vires or constitutional identity review.[325] Whereas the GFCC takes a firm stance on protecting its own review powers on the constitutional permissibility of EU law, it only does so in order to grant EU values a higher level of protection; moreover, in the overall assessment it almost always comes to EU law-friendly conclusions. But its insistence on being the final arbiter of EU law may encourage other domestic apex courts to follow suit, and to opt out from the principle of primacy, whereas these latter fora may use the same claims for a less EU law-friendly

[319] Christoph Grabenwarter, *Constitutional Resilience, VerfBlog,* 6 December 2018, https://verfassungsblog.de/constitutional-resilience/.

[320] Matthias Kumm, *How populist authoritarian nationalism threatens constitutionalism or: Why constitutional resilience is a key issue of our time, VerfBlog,* 6 December 2018, https://verfassungsblog.de/how-populist-authoritarian-nationalism-threatens-constitutionalism-or-why-constitutional-resilience-is-a-key-issue-of-our-time/.

[321] Tomasz Tadeusz Koncewicz, The Democratic Backsliding and the European constitutional design in error. When will HOW meet WHY?, VerfBlog, 18 December 2018, https://verfassungsblog.de/the-democratic-backsliding-and-the-european-constitutional-design-in-error-when-how-meets-why/.

[322] On international mechanisms correcting the failure of domestic law to protect minorities see for example A. Verdross, *Die Einheit des rechtlichen Weltbildes auf Grundlage der Völkerrechtsverfassung*, Tübingen: Mohr, 1923.

[323] Petra Bárd, Sergio Carrera, Elspeth Guild, Dimitry Kochenov (2016) An EU mechanism on Democracy, the Rule of Law and Fundamental Rights, Brussels: Center for European Policy Studies (CEPS), 65-66.

[324] Petra Bárd, Wouter van Ballegooij, Judicial independence as a precondition for mutual trust? The CJEU in Minister for Justice and Equality v. LM, New Journal of European Criminal Law, 2018, Volume 9, Issue 3, forthcoming.

[325] For such attempts see e.g. BVerfGE 37, 271 – Solange I, 7 BVerfGE 73, 339 – Solange II, BVerfGE 102, 147 – Bananenmarktordnung, BVerfGE 89, 155 – Maastricht, BVerfGE 123, 267 – Lissabon, BVerfG, 21.06.2016 - 2 BvR 2728/13; 2 BvR 2728/13; 2 BvR 2729/13; 2 BvR 2730/13.

interpretation or even for lowering the level of human rights protection, especially in a state of constitutional capture to be discussed *infra*.

Second, the EU itself is founded on representative democracy,[326] with strong participatory rights for citizens of the EU,[327] the rule of law and fundamental rights.[328] Not only in the Member States, but also at the EU level, open democratic societies depend on whether there are public debates that allow well-informed citizens to express their will through free and fair political processes.[329] As the Commission noted,

> [d]isinformation erodes trust in institutions and in digital and traditional media, and harms our democracies by hampering the ability of citizens to take informed decisions. Disinformation also often supports radical and extremist ideas and activities. It impairs freedom of expression, a fundamental right enshrined in the Charter of Fundamental Rights of the European Union (Charter).[330]

This holds true at both the Member State and the EU level. **Should elections to the European Parliament be distorted by disinformation and propaganda, trust in European institutions will be shaken, and the legitimacy of the elected representatives will be directly affected.**

Militant democracy not only incorporates exceptional techniques, such as party banning, but also **we argue for viewing enforcing democracy, the rule of law and fundamental rights as tools to make the Member States and the EU itself resilient to political forces seeking to destroy democracy**, one of the core values behind European integration. Even seemingly technical rules to fight false information, and thereby revive deliberative democracy, contribute to upholding the EU's value system.

Table 9: Attempts to manipulate democratic processes in the EU

		Outcome	
		Democracy upheld	**Democracy challenged**
Process	Democratic processes, the rule of law and fundamental rights are **respected**	A. *No issue arises from the perspective of Article 2 TEU.*	B. Tools of militant democracy fail to prevent an undemocratic power from gaining support, and overthrowing the liberal democratic order once it wins an election. *The EU needs to exercise its supervisory powers.*
	Democratic processes, the rule of law and fundamental rights are **manipulated**	C. Manipulation of democratic processes benefit a democratic power or does not achieve its aims. *Assessment is needed of whether the manipulation itself is a threat to Article 2 TEU values.*	D. Undemocratic tools, or means in violation of the rule of law, and fundamental rights are used to overthrow liberal democracy. *A clear case for triggering the mechanism protecting Article 2 TEU values.*

Source: Authors.

[326] Article 10(1) TEU: "The functioning of the Union shall be founded on representative democracy."

[327] Article 10(3) TEU, first sentence: "Every citizen shall have the right to participate in the democratic life of the Union."

[328] Cf. Article 2 TEU.

[329] Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Region, Tackling online disinformation: A European Approach, Brussels, 26.4.2018, COM(2018) 236 final, 1.

[330] Article 11, Charter. Article 6(1) of the Treaty of the European Union confers binding force on the Charter and states that it "shall have the same legal value as the Treaties."

Ideally, EU Member States today operate along scenario A, meaning they respect democratic rules. Gaining power through democratic elections and then attempting to overthrow democracy in a functional democracy based on the rule of law is more of a hypothetical possibility (scenario B), due to the in-built mechanisms, institutions and procedures of public law limiting government. In a democracy based on the rule of law emerging as an institutional ideal, correction mechanisms compensate for the deficiencies of a majoritarian government.

Undemocratic tools to gain power are likely to mushroom around election times, so as to gain or retain power, irrespective of whether the government of the day or the government to be in other aspects respects democratic values (scenario C). National and public international law, just like monitoring entities, show special interest in overseeing elections. Once certain irregularities can be traced, the question is whether the problem can be tackled in the national setting, via for example election boards and ultimately the judiciary, and if not, what role can external entities like the EU play. The EU interest is even more visible when the undemocratic forces attempt to capture the state (scenario D).

For the sake of elections at the European Parliament, however, EU mechanisms to oversee the fairness of elections are both external tools as elections ultimately are conducted nationally, but also internal tools, since elections to an EU institution (and in fact the single democratically elected institution) are at stake. **The EU has a vital interest and an unequivocal obligation to protect and enforce values enshrined in Article 2 TEU, out of which democracy and the rule of law, including free and fair elections, stand out. It is therefore obliged to stand up against abuses, whether procedural (scenario C) or substantive (scenario D**).

2.3.4 Enforcement of values in the EU setting

Under current treaty law, the EU has two options to tackle democracy or rule of law violations in the Member States. It may initiate infringement proceedings in accordance with Article 258 of the Treaty on the Functioning of the European Union (TFEU) or take political actions relying on Article 7 TEU. Infringement proceedings are narrower and broader than Article 7 TEU procedures at the same time. While the former have to involve an element of EU law, the latter may go beyond the EU law domain. However, the infringement procedure may be employed to tackle any failure within EU law of whatever gravity, whereas the Article 7 TEU mechanism is there to address a 'serious' or a 'serious and persistent' breach of values enshrined in Article 2 TEU, including the rule of law. But even for an infringement action involving a dispute over the rule of law to reach the judicial phase, violations must be going on for quite some time. Also, entering the judicial phase of the infringement procedure the subject matter of which is democracy or the rule of law, proves that the Commission and the Member State concerned are not on the same page as regards the foundational values they are supposed to share, respect and promote. In other words, neither Article 7 TEU nor Article 258 TFEU procedures are preventive, or precautionary, using the above terminology.

The EU already possesses a number of instruments assessing Member States' compliance with the rule of law or its elements, including the legally binding EU Charter of Fundamental Rights.[331] For example, since 2012 these include the EU Justice Scoreboard,[332] which feeds into the EU yearly cycle of economic policy coordination, or 'European semester', to foster structural reforms at the national level.[333] Its aim is to identify shortcomings and good examples, and to foster structural reforms at national levels. The Scoreboard nonetheless has some major weaknesses; it is criticised for being

[331] Sergio Carrera, Elspeth Guild, Nicholas Hernanz, *The triangular relationship between fundamental rights, democracy and the rule of law in the EU, towards an EU Copenhagen mechanism* (2013) 4-15, 42-57. Available at http://www.ceps.eu/system/files/Fundamental%20Rights%20DemocracyandRoL.pdf, http://www.europarl.europa.eu/RegData/etudes/etudes/join/2013/493031/IPOL-LIBE_ET%282013%29493031_EN.pdf

[332] The EU Justice Scoreboard: Towards more effective justice systems in the EU, http://ec.europa.eu/justice/newsroom/effective-justice/news/150309_en.htm.

[333] See Communication from the Commission, Annual Growth Survey 2015, COM (2014) 902 final. For a study of the European semester method refer to 2013 CEPS Study.

> incapable of catching the most atrocious violations: it does not sufficiently detect internal linkages, thus it examines individual elements but fails to supply a qualitative assessment of the whole.[334] The Scoreboard does not foresee any coercive action or sanctions/penalties in a situation where an EU Member State may be seen as performing poorly on the above-mentioned indicators.[335]

Other mechanisms, such as the EU Anti-Corruption Reporting Mechanism for Periodic Assessment ('EU Anti-corruption Report') or the Cooperation and Verification Mechanism for Bulgaria and Romania,[336] involve important segments of the rule of law, but are limited in material and territorial scope. Additionally, these cannot be seen as supervisory mechanisms.[337]

As a way out of these difficulties, and building on several other past EP resolutions,[338] in its Resolution adopted in a Plenary session on 8 September 2015[339] the Parliament called on the Commission to draft an internal strategy on the rule of law "accompanied by a clear and detailed new mechanism". On 25 October 2016 the EP passed a Resolution inviting the Commission to initiate legislation on a comprehensive rule of law, democracy and fundamental rights scoreboard (DRF Resolution).[340] The EP's legislative initiative report called on the Commission to submit a proposal by September 2017 for the conclusion of a Union pact for democracy, the rule of law and fundamental rights (DRF Pact). The document was accompanied by a thorough European added value assessment.[341] More than a dozen Member States unifying under the slogan "Friends of the Rule of Law" welcomed the idea and took the lead in moving this initiative forward.[342]

The point of departure is that a scoreboard is a 'process' encompassing a multi-actor and multi-method regular cycle.[343] This was also the approach the EP took in the four subparts of the Pact that have been designed: (i) the annual European report on democracy, the rule of law and fundamental rights (European DRF report); (ii) the annual inter-parliamentary debate on the basis of the European DRF report; (iii) arrangements for remedying

[334] Kim Lane Scheppele, 'The Rule of Law and the *FrankenState*: Why Governance Checklists Do Not Work' (2013) 26 Governance 4, 559–562.

[335] Petra Bárd and others, *An EU Mechanism on Democracy, the Rule of Law and Fundamental Rights*, (CENTER FOR EUROPEAN POLICY STUDIES 2016, Brussels) 8-9.

[336] For the latest Cooperation and Verification Mechanism see http://ec.europa.eu/cvm/progress_reports_en.htm.

[337] European Commission Decision establishing an EU Anti-corruption reporting mechanism for periodic assessment ('EU Anti-corruption Report'), 6 June 2011, http://ec.europa.eu/dgs/home-affairs/what-we-do/policies/pdf/com_decision_2011_3673_final_en.pdf.

[338] See, for example, European Parliament resolution of 27 February 2014 on the situation of fundamental rights in the European Union (2012), European Parliament resolution of 3 July 2013 on the situation of fundamental rights: standards and practices in Hungary (pursuant to the European Parliament resolution of 16 February 2012).

[339] European Parliament resolution of 8 September 2015 on the situation of fundamental rights in the European Union (2013-2014) (2014/2254(INI)), 8_TA-PROV(2015)0286

[340] European Parliament resolution of 25 October 2016 with recommendations to the Commission on the establishment of an EU mechanism on democracy, the rule of law and fundamental rights (2015/2254(INL)), P8_TA-PROV(2016)0409.

For the time being the Commission followed up on this document in a rather hostile manner, which can be regarded as part of an inter-institutional dialogue on the matter. See Commission response to text adopted in plenary, SP(2017)16, 17 February 2017. For an assessment see Petra Bárd, Sergio Carrera, 'The Commission's Decision on 'Less EU', Safeguarding the rule of law: a play in four acts' (2017), CEPS Policy Insights https://www.ceps.eu/publications/commission's-decision-'less-eu'-safeguarding-rule-law-play-four-acts

[341] Wouter Van Ballegooij, Tatjana Evas, 'An EU mechanism on democracy, the rule of law and fundamental rights, interim European added value assessment accompanying the legislative initiative report (Rapporteur Sophie in 't Veld)' European Parliamentary Research Service, October 2016, PE.579.328; Annex I, Laurent Pech and others, 'Assessing the need and possibilities for the establishment of an EU scoreboard on democracy, the rule of law and fundamental rights'; Annex II, Petra Bárd and others with a thematic contribution by Wim Marneffe, (2016) 'Assessing the need and possibilities for the establishment of an EU scoreboard on democracy, the rule of law and fundamental rights'.

[342] http://www.liberalforum.eu/en/news/details/interview-with-mep-sophie-int-veld.html, https://euobserver.com/opinion/136030

[343] Petra Bárd and others, *An EU Mechanism on Democracy, the Rule of Law and Fundamental Rights*, (CENTER FOR EUROPEAN POLICY STUDIES 2016, Brussels) 73.

possible risks and breaches; and (iv) a DRF policy cycle within the Union institutions (DRF Resolution, points 5-7 and 15).

As some of the authors of the present study have suggested earlier, the entity **checking Member State compliance** with the values enshrined in Article 2 TEU at the heart of the democracy, rule of law and fundamental rights (DRF Pact) should **monitor whether Member States are 'on track' or 'off the track' in relation to these values.**[344] We have argued that if these values are respected or if deficiencies are remedied in the national setting by in-built correction mechanisms, and suggestions and obligations imposed by international fora are respected, a 'sunshine policy' may be followed, "which engages and involves rather than paralyses and excludes", and where value control "is owned equally by all actors".[345] Here an exchange of good practices may contribute to the fairness of elections and the fight against fake news. Yet, once a Member State systematically undermines democracy and deconstructs the rule of law, there is no reason to presume the good intentions of those in power to engage in a sunshine approach and share knowledge in order to have free and fair elections Europe-wide – something they may already have abolished internally or are progressing towards jeopardising national suffrage. With regard to countries falling under this latter category, we have suggested initiating systemic infringement procedures, Article 7 procedures and introducing effective sanctions. Should the DRF Committee come to the conclusion that a country is at high risk of not respecting Article 2 TEU values, one of the sanctions vaguely referenced by Article 7(3) TEU could be to conduct EP elections only if special external safeguards are introduced.

2.4 Direct impact on human rights

Disinformation and propaganda have direct impacts on human rights in two major respects:

a) violating **human dignity**, which includes decisional autonomy and **privacy** as well. The way artificial intelligence, bots and algorithms relate to humans, specifically to disseminate targeted messages based on user profiling and classification, is deeply injurious to these human rights. The underlying process of amassing vast amounts of **personal data** calls into question fundamental principles of data protection, such as purpose limitation and data minimisation; and

b) causing an injury to political rights. **Freedom of expression** and the **right to information** are equally necessary to a meaningful exercise of the voting rights.

2.4.1 Human dignity, autonomy, privacy and data protection

Human dignity is the root of all other human rights, expressed in Article 1 of the Charter of Fundamental Rights of the EU, Article 1 of the Universal Declaration on Human Rights and the Preamble of the International Covenant on Civil and Political Rights.

Interacting with robots is a new phenomenon that is not yet crystallised by legal jurisprudence. However, the deception of a user about the identity of a communication partner is clearly injurious to the person's dignity.

The ethical report by the European Group on Ethics in Science and New Technologies (Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems)[346] puts human dignity in first place among the ethical principles. It emphasises that a particular concept of human dignity is "that we are aware of whether and when

[344] Petra Bárd, Sergio Carrera, Elspeth Guild, Dimitry Kochenov (2016) An EU mechanism on Democracy, the Rule of Law and Fundamental Rights, Brussels: Center for European Policy Studies (CEPS),

[345] G.N. Toggenburg and J. Grimheden, 'The Rule of Law and the Role of Fundamental Rights: Seven Practical Pointers', in: C. Closa and D. Kochenov (eds.), *Reinforcing Rule of Law Oversight in the European Union*, Cambridge: Cambridge University Press, 2016.

[346] European Group on Ethics in Science and New Technologies: Artificial Intelligence, Robotics and 'Autonomous' Systems. March 2018. http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf

we are interacting with a machine or another human being". Also, the Code of Practice on Disinformation[347] formulates the intention to clearly identify robots and artificial intelligence.

When algorithms decide on what political views and information users should encounter, that violates users' autonomy. It is often regarded as a service, but services are supposed to be chosen knowingly.

The right to privacy is ensured by Article 8 of the European Convention on Human Rights, Article 7 of the EU's Charter of Fundamental Rights, Article 12 of the Universal Declaration on Human Rights and Article 17 of the International Covenant on Civil and Political Rights. The right to the protection of personal data emerged only later in the second half of the 20th century, but the EU's Charter of Fundamental Rights also guarantees the right to protection of personal data (Article 8).

Privacy takes on specific importance in relation to artificial intelligence, algorithms and social bots. The European Group on Ethics in Science also mentions the right "to not be profiled, measured, analysed, ... or nudged".[348] It mentions the limits to the determinations and classifications concerning persons, made on the basis of automated systems, as part of the protection of **human dignity.**

As explained already in the report, disinformation follows the structural logic of commercial industry and its digital dissemination is backed up by the advertising (ad tech) ecosystem, which is meant to aggregate user attention and to maximise profit by selling advertising.[349] As such, the users are transformed from active agents into passive 'consumers' of information. For advertisers and digital platforms, the truthfulness of information is of secondary importance, as long as the individual's attention is captured. This has a direct impact on the **dignity and privacy** of an individual, and directly **affects protection of the individual's personal data.**
Participants in the ad tech industry ecosystem are interested in collecting vast amounts of personal data about individuals to better serve them with targeted ads that cater to their unique interests. **Non-commercial advertisers, including internal actors (nation states, political parties, civic movements, etc.) and external actors (foreign states, civic and religious movements, etc.) leverage this industry to their advantage**.[350] It may encompass both legitimate purposes, including a purpose to further political debates, and purposes of disinformation and interference with the national decision-making processes.[351]

The real impact of political micro-targeting on individuals is not yet thoroughly researched. As such, micro-targeting is a notion that encompasses a range of data-driven processes aimed at creating personalised messages tailored to target audiences.[352] **On one end of the spectrum there are psychometric profiling techniques aimed at exploiting personal vulnerabilities and appealing to individual fears.**[353] **Such practices are deeply problematic from a dignity and privacy point of view.** On the other end, there are less dignity-invasive practices, for example, user targeting based on age group or country of residence. A common denominator for

[347] EU Code of Practice on Disinformation. https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation. Sept. 26, 2018.

[348] European Group on Ethics in Science and New Technologies: Artificial Intelligence, Robotics and 'Autonomous' Systems. March 2018. http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf at 19.

[349] Ghosh, D., and Scott, B., Digital Deceit. The Technologies Behind Precision Propaganda on the Internet, 23 January 2018, https://www.newamerica.org/public-interest-technology/policy-papers/digitaldeceit/

[350] European Data Protection Supervisor (EDPS), Opinion on online manipulation and personal data, 3/2018, p. 11, https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf

[351] See e.g., Thompson, I., How Irish anti-abortion activists are drawing on Brexit and Trump campaigns to influence referendum, 2 May 2018, https://www.opendemocracy.net/5050/isobel-thompson/irish-anti-abortion-campaigners-brexit-trump-data-companies

[352] European Data Protection Supervisor (EDPS), Opinion on online manipulation and personal data, 3/2018, supra note 39, https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf

[353] Borgesius, F. et al., Online Political Microtargeting: Promises and Threats for Democracy. Utrecht Law Review, 14(1): 82–96, p. 87.

these practices is, as pointed out above, an underlying process of amassing vast amounts of personal data. Such data are often stripped of its original purpose(s)[354] and may be used by the actors and for ends the individual is largely unaware of, **in contravention of the data protection principles.**

In the EU, the processing of personal data is regulated primarily by the GDPR, which became effective on 25 May 2018, whereas the confidentiality of electronic communication (including rules on cookies and direct marketing) is governed by the ePrivacy Directive.[355] Political parties, digital platforms, data brokers and ad tech companies processing data about the data subjects in the EU have to respect their obligations under the GDPR, including an obligation to process data in line with the six data protection principles and to be accountable for such processing.[356] **The notion of personal data is a comprehensive one and includes data that are not only provided or observed from the behaviour of the individual online, but also inferred about that person (e.g. sexual orientation, race or political convictions).[357]**

The processing of special categories of data, including data about political opinions, is generally prohibited unless specific justifications under Article 9 of the GDPR apply. In this respect, political parties and political campaigns[358] may process data about individuals if, for instance, these data have been made manifestly public by the individual, if that person has given explicit consent, if there is a substantial public interest on the basis of EU or national law,[359] or if special categories of data relate solely to members or former members of the party but only for internal disclosure purposes (e.g. a political party would not be able to lawfully disclose such data to a third party (e.g. a data analytics company) without the individual's consent).[360]

Although the GDPR does not mention 'political micro-targeting' as such, it does impose stricter rules on commercial and non-commercial targeted advertising (which may or may not be based on profiling) that produces sufficiently significant effects on individuals, i.e. significantly affects the circumstances, behaviour or choices of individuals.[361] Although the European Data Protection Board or the courts have not explicitly acknowledged that political micro-targeting does reach the threshold of 'significant effect', this seems to be the position the European Commission recently advanced in its guidance document: "Given the significance of the exercise of the democratic right to vote, personalised messages which have for instance the possible effect to stop individuals from voting or to make them vote in a specific way could have the potential of meeting the criterion of significant effect."[362]

[354] See e.g., from the investigation carried out by the UK Information Commissioner's Office: '[t]he company, which provides advice on pregnancy and childcare, sold the information to Experian Marketing Services, a branch of the credit reference agency, specifically for use by the Labour Party. Experian then created a database which the party used to profile new mothers in the run-up to the 2017 General Election.' Information Commissioner's Office, Investigation into the use of data analytics in political campaigns, a report to Parliament 6 November 2018, p. 60 https://ico.org.uk/media/2260277/investigation-into-the-use-of-data-analytics-in-political-campaigns-20181107.pdf. Also see France 24, Austria's Post Office under fire over data sharing: https://www.france24.com/en/20190108-austrias-post-office-under-fire-over-data-sharing

[355] We discuss ePrivacy Directive in more detail in chapter 3.1

[356] Article 5 of the GDPR.

[357] Article 29 Data Protection Working Party, Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (n 19), p. 8; Article 29 Data Protection Working Party, Guidelines on the Right to Data Portability (2017) 16/EN WP 242 rev.01 10, https://ec.europa.eu/newsroom/document.cfm?doc_id=44099. On the overview of the CJEU jurisprudence on this issue see Wachter, S. and Mittelstadt, B., A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI, 5 October 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829

[358] Please note that so-called 'household exception' under GDPR would not apply to political parties and political campaigns.

[359] Also see Recital 56 of the GDPR.

[360] Also see European Commission, Commission guidance on the application of Union data protection law in the electoral context, 12 September 2018, p. 5, https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-data-protection-law-electoral-guidance-638_en.pdf

[361] Article 22 of the GDPR.

[362] European Commission, Commission guidance on the application of Union data protection law in the electoral context, 12 September 2018, p. 8, https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-data-protection-law-electoral-guidance-638_en.pdf

Although commendable, this position seems to relate the ‘significant effect’ of political micro-targeting to voting behaviour, although personalised messages may not be aimed directly at influencing voting behaviour as such, but rather at influencing individuals’ understanding of certain events, processes and for example, increasing political polarisation in general. It is unclear whether micro-targeting in a non-commercial setting that is not aimed directly at influencing individuals’ behaviour would necessarily be recognised as meriting special protection under Article 22 of the GDPR. If it were, Article 22 prohibits solely automated decision-making based on special categories of personal data (e.g. data about political opinions) unless it is based on explicit consent or is necessary for reasons of substantial public interest, on the basis of EU or Member State law.[363]

The recent investigation by the UK Information Commissioner’s Office into the use of data analytics in political campaigns has identified particular concerns related to “purchasing of marketing lists and lifestyle information from data brokers without sufficient due diligence, a lack of fair processing, and the use of third party data analytics companies, with insufficient checks around consent”.[364] **This demonstrates that intermediaries (data brokers, data analytics companies and other actors within the ad tech ecosystem) have an important role to play in generating the impact on individuals’ dignity, privacy and data protection.** Although in theory the strict rules of the GDPR and of the ePrivacy Directive do apply to them (even if the entities themselves are not established in the EU),[365] the coherent and timely enforcement of the data protection rules, especially in cases where an alleged perpetrator is established outside the EU, is an ultimate concern. Currently, a majority of Member States have opted against enshrining collective redress mechanisms in their national laws,[366] which sufficiently limits the possibility of the data subjects to effectively defend their rights. **Equally, the resources available[367] to and independence of the data protection authorities[368] vary significantly across the Member States, which may effectively result in a lack of de facto protection against harms to individual privacy and data protection.**

2.4.2 Freedom of expression and the right to information

Privacy, freedom of expression and the right to information are all parts of the political rights of citizens. Their goal is to ensure citizens' participation in the democratic decision-making process. Freedom of expression has several theoretical justifications. According to the instrumental theory (also called the democracy argument), freedom of expression is necessary because it enables citizens to engage in public discussion and thereby to participate in the governance of their community.[369] According to the constitutive justification, freedom of speech is necessary because it enables individuals to 'develop their faculties', to realise their own potential and autonomy through expressing themselves – and to become responsible moral agents of a just political society.[370] A further argument for free speech is the process of constantly searching for truth: even falsehood and mistakes should be tolerated because only through the consideration of these can society reach the truth.[371]

[363] Article 22(4) of the GDPR

[364] Information Commissioner’s Office, Investigation into the use of data analytics in political campaigns, a report to Parliament 6 November 2018, p. 8, https://ico.org.uk/media/2260277/investigation-into-the-use-of-data-analytics-in-political-campaigns-20181107.pdf.

[365] Article 3 of the GDPR

[366] Article 80(2) of the GDPR. Internal Association of Privacy Professionals, EU Member State GDPR Derogation Implementation Tracker, 11 December 2018, https://iapp.org/resources/tools/eu-member-state-gdpr-derogation-implementation/

[367] Fazlioglu, M., Analyzing changes in DPA Income and Staff, from 2011 to 2016, 11 December 2017, https://iapp.org/news/a/analyzing-changes-in-dpa-income-and-staff-2011-2016/

[368] GDPR misuse in Romania: “independence of DPA” and “transparency” – keywords or buzzwords? 17 December 2018, https://www.gdprtoday.org/gdpr-misuse-in-romania-independence-of-dpa-and-transparency-keywords-or-buzzwords/

[369] Barendt, Eric: Freedom of Speech. Oxford University Press, 2005. p. 19-20.

[370] Dworkin, Ronald: Freedom's Law: The Moral Reading of the American Constitution. Oxford University Press, 1999. p 200.

[371] Mill, John Stuart: On Liberty. Boston. 1863. p. 50-58.

While this argument is of course criticised from many perspectives, one strong defence for it may be that the suppression of speech – even apparently false – creates a suspicion of authority. One of the relevant criticisms of Mill's theory on the search for truth is that Mill overvalued intellectual discussion and the need for all individuals to debate public affairs. The harm caused by the belief in falsehood is short term, and the benefits of a continual, uninhibited debate can be reaped only in the long term, if at all.[372] In fact, Mill was very well aware of the high price that is sometimes paid for truth: "History teems with instances of truth put down by persecution. If not suppressed forever, it may be thrown back for centuries."[373]

One further aspect is that **the competition between ideas is not necessarily fair: tools of an authoritarian media policy or financial tools to amplify content can decide the fight**, rather than the merit of the ideas. As Barendt puts it, "some constraints may be required to ensure that false propositions do not drive out truths".[374] Neither of the mentioned justifications are unlimited, and many forms of speech are indeed limited by national laws and even allowed so by international human rights conventions (see section 3.2).

Meiklejohn[375] held that the main purpose of free speech is for citizens to receive all information which may affect their choices in the process of collective decision-making and, in particular, in the voting process. "The voters must have it, all of them."[376] The primary purpose of the constitutional protection of free speech is to ensure that citizens, so far as possible, will understand public matters.[377] The international human rights conventions list the right of access to information next to freedom of expression (for example, Article 19 of the International Covenant on Civil and Political Rights states that "this right shall include freedom to seek, receive and impart information and ideas of all kinds"). Thus, the passive **side of freedom of expression is freedom of information; if this is missing, and people lack trustworthy information, it hinders the formation of their opinion and consequently their political decision-making (voting).**

The discovery of truth, the understanding of public matters and informed decision-making can take place **only if the public discourse as a whole is diverse enough so that citizens can have options to learn several different opinions and ideas**. If disinformation and propaganda reach a level such that they essentially distort public discourse (e.g. regularly false information occupies the mainstream agenda or hostility and xenophobia take root in society, or all media outlets are dominated by one political power), then the citizens' exercise of their political rights is stifled.

[372] Barendt, supra note. p. 9.

[373] Mill, supra note p. 56.

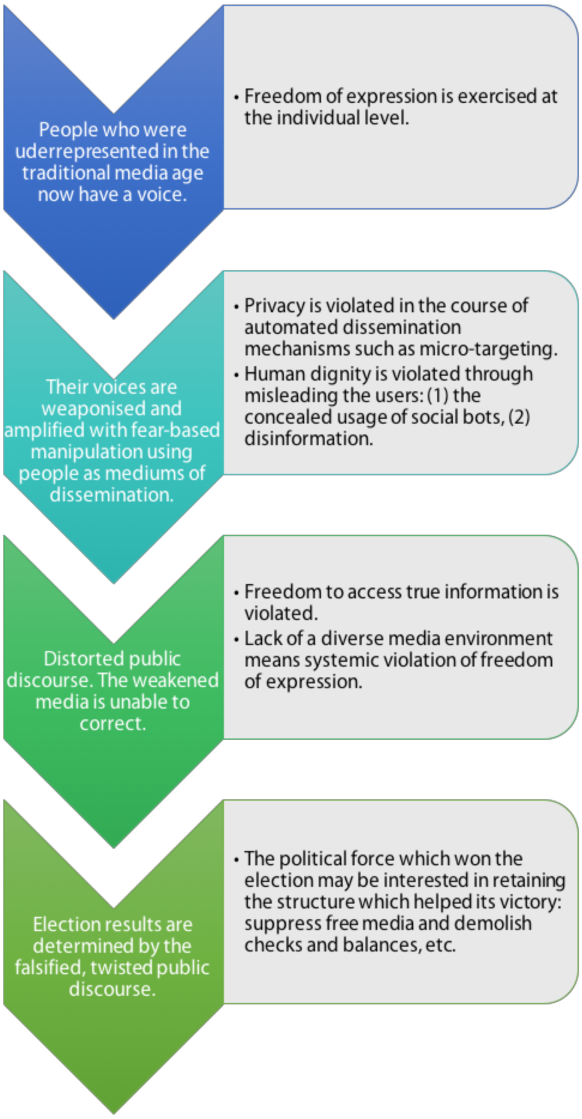
[374] ibid.

[375] Sadurski, Wojciech: Freedom of Speech and Its Limits. Kluwer Academic Publishers, 2014. p.20.

[376] Meiklejohn, Alexander: Free Speech And Its Relation to Self-Government. The Lawbook Exchange, Clark, New Jersey, 2004. p. 88.

[377] Ibid. p. 89.

Figure 4: How the process of spreading disinformation and propaganda in social media affects human rights



People who were uderrepresented in the traditional media age now have a voice.

- Freedom of expression is exercised at the individual level.

Their voices are weaponised and amplified with fear-based manipulation using people as mediums of dissemination.

- Privacy is violated in the course of automated dissemination mechanisms such as micro-targeting.
- Human dignity is violated through misleading the users: (1) the concealed usage of social bots, (2) disinformation.

Distorted public discourse. The weakened media is unable to correct.

- Freedom to access true information is violated.
- Lack of a diverse media environment means systemic violation of freedom of expression.

Election results are determined by the falsified, twisted public discourse.

- The political force which won the election may be interested in retaining the structure which helped its victory: suppress free media and demolish checks and balances, etc.

Source: Authors.

In the current media environment, the individual's freedom of expression can be realised to a maximum. But the reception side of information is controversial: on the one hand, anyone searching for it has access to all kinds of information. On the other hand, for many individuals, disinformation, hate speech and racism have occupied their information environment.

False information in itself (if it does not violate others' reputation, for example) enjoys the protection of freedom of expression, but when the whole environment of public discourse becomes occupied and dominated by falsehood, it frustrates the primary purpose of freedom of expression.

As long as disinformation originates from small media outlets and individuals, a strong professional media system can counteract its negative effect. Yet a crisis-stricken media that lost its reputation cannot effectively counteract the effects of disinformation and propaganda campaigns. And if the source of this is the government, a weak media cannot fulfil its watchdog role of disclosing the malpractices.

3. EXPLORING THE LIMITS OF EXISTING LEGAL REGULATIONS

KEY FINDINGS

Responsibility of platform providers

- There is as yet no standardised definition in the literature for platform providers or social media. This absence has hindered the development of a conceptual framework on their rights and responsibilities. The term 'platform providers' is proposed to designate those services that convey third-party content with value added services, of which 'social media' is a sub-category.
- Platform providers should not be responsible for third-party content; but they should be responsible for administering their platform rules: the transparency of their algorithms, ensuring that their algorithms have no viewpoint-based discrimination, for distinguishing sponsored content and ads from other content, identifying and disabling fake accounts, protecting the privacy of users, including those who are not members of their services.
- For the purpose of dealing with illegal content, the 'notice-and-notice' system has benefits compared to the 'notice-and-takedown' system: it causes less harm to freedom of expression, to user autonomy, and does not force platform providers to decide on the lawfulness of third-party content.
- Dominant social media platforms should accept a certain level of social responsibility because of their impact on the public sphere, just like traditional mass media companies do.[378]

International background legislation

- The standards of the ECHR require that any restriction on speech is regulated by law, has a legitimate reason such as national security, territorial integrity or prevention of disorder, shall be necessary to achieve this goal, and the restriction must remain proportionate to the goal.
- ECtHR jurisprudence has established the precedent of accepting speech restrictions for the goal of territorial integrity and preventing disorder; and the restriction of public issue advertising on mass media has also been found legitimate.
- Prohibition of incitement to violence, discrimination or hostility would be in accordance with international standards, as was specifically pointed out by the Joint Declaration of the UN, OSCE, OAS and ACHPR. Article 20 of the ICCPR and Article 4 of the ICERD require signatories to prohibit by law any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence (ICCPR) or dissemination of ideas based on racial superiority (ICERD).
- Article 17 of the European Convention on Human Rights is an example for militant democracy: the Convention does not protect actions that aim at the destruction of any rights inherent in the Convention.
- The principles of freedom of expression would not enable content restriction on the basis of falsity alone.

3.1 The role and responsibility of social media

Social media companies – particularly the most popular, Facebook – have had an increasingly influential effect on public communication, and consequently democratic public discourse in the past decade. These platforms have not only become an interface for users to engage with each other, but also to enable the messages of the advertising and political sectors to find their way to users. In the ever more sophisticated techniques of tailoring, targeting and disseminating messages, platform providers have provided the fuel: personal data.

Facebook founder Mark Zuckerberg appeared in Washington and in the European Parliament to represent his company and face questions. It appeared that he had not anticipated the ways that his technology was used.[379]

[378] See the United Nations, Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 6 April 2018, paragraph 72. (https://undocs.org/A/HRC/38/35).

[379] See https://www.youtube.com/watch?v=o0zdBUOrhG8, also EP Motion for a Resolution 2018/2855(RSP) 10.10.2018.

Considering the rapid development of automated technology and artificial intelligence in the future (see also section 5.1), such **surprises may multiply unless regulators consciously prepare for them**.

Platform providers are widely seen and sometimes blamed for the harms caused by disinformation and propaganda campaigns, and the distortion of public discourse in general.

Clearly, they have a function that did not exist earlier, and their rights and responsibilities are not delineated by policy measures or regulation. The authors of this study hold the view that the responsibilities of platform providers should extend to the actions that they perform: collecting and handling personal data, allowing user registration, designing and using algorithms. Their responsibility should not extend, however, to the actions of their users: publishing and disseminating content. Importantly, platform providers should not be allowed to control content, or to decide on the nature of content. Delineating their responsibilities requires technological knowledge from the side of the legislator.

3.1.1 State-of-the-art relating to responsibility of platforms

3.1.1.1 Definition

Internet service providers' responsibility was regulated by the e-Commerce Directive[380] in 2000, the most appropriate framework at the time: access providers, hosting providers and search engines could claim immunity of legal responsibility for content that they carried, provided that they did nothing more than the minimum duty of these providers (which was accurately set out in Articles 12-14. of the Directive). As soon as the directive was ready, it was outdated: it did not mention the new type of user interface that enabled anybody to publish content on the web, and which so profoundly changed the way the internet was used that influential experts used the term 'web 2.0' - and the previous web was retrospectively dubbed web 1.0.

Before 2000, publishing content on the internet involved reserving a domain name, building a webpage, and using coding language every time the owner wanted to make changes to the webpage. Ordinary users could passively search and browse through the web, without interacting (except for certain bulletin board services or chat-rooms). Sometime between 1997 and 2004, online services emerged that offered a platform for the everyday lay user to simply enter their content – using a user-friendly platform – that would immediately appear online: blogs, comment sections and the new social media services such as MySpace and Facebook. By now, online actors built on this technology are among the most important players of the web's content infrastructure: including eBay, Über, Tinder, Airbnb, and many aggregators and facilitators of human communication beyond Facebook. Although they have become crucial players in everyday life, they themselves do not publish, but mainly convey information. They could be compared to shopping malls which do not 'sell' anything: they just lease their premises to retail shops, and operate the infrastructure of the building.

The aim of the e-Commerce Directive was to liberalise innovation and to create a safe and secure online environment. Regarding the immunity of internet service providers, it followed American jurisprudence, which by then – after several rounds in court – was clearly defined, adding nuances to the application of Section 230 of the Communication Decency Act. This firmly declared that "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider".[381] The dilemma in American jurisprudence was whether to treat ISPs as 'publishers' responsible for all content they publish, or as 'distributors' who are only responsible if they had actual knowledge of the content. The Directive chose the 'distributor' liability clause, which – although this deserved some justified criticism by freedom of expression advocates – served the industry relatively well. However, the formulation of service

[380] 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').

[381] 47 U.S. Code § 230 - Protection for private blocking and screening of offensive material

(c)Protection for "Good Samaritan" blocking and screening of offensive material (1)Treatment of publisher or speaker: No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider

provider activities was **too narrow to be applied for the new services** that became prevalent in the web 2.0 age. These actors were not access providers, nor hosting providers, but also not content providers. They **still do not have a legal category** with different terms being used to describe various groups of these actors: they were called "internet intermediaries" by the Council of Europe Committee of Experts on Internet Intermediaries; YouTube was a "video-sharing platform provider" according to the 2018 Audiovisual Media Services (AVMS) Directive; the Code of Conduct on countering illegal hate speech online uses the wide category "IT companies" because Microsoft is included; the Commission Recommendation on measures to effectively tackle illegal content online mentions "those online platforms"; the Communication from the Commission on tackling online disinformation mentions "online platforms that distribute content".

The Joint Declaration calls them "intermediaries". The European Council decision (March 2018) called them "social networks and digital platforms". The Proposal for a Regulation on preventing the dissemination of terrorist content online uses the expression: "hosting services that allow the upload of third-party content", adding that "such providers of information society services include social media platforms, [etc.]". If they could be considered hosting services indeed that is a different position compared to legal decisions (see below).

The **Proposal for a Regulation on preventing the dissemination of terrorist content online adds to the confusion by calling all its subjects "hosting service providers", defining these as including "social media platforms, video streaming services, video, image and audio sharing services, file sharing and other cloud services to the extent they make the information available to third parties and websites where users can make comments or post reviews".**

We consider that the **wording 'platform providers' is the most accurate for those services that convey third-party content with value added services**, **of which 'social media' is a sub-category.**

3.1.1.2 Liability

Fortunately, the recently passed amendment to the **AVMS Directive clarifies that video-sharing-platform-providers can enjoy exemptions from liability as defined in the chapters mentioned, with reference to Articles 12-15 of the e-Commerce Directive.**[382] This means that **video-sharing platform providers are not liable for third-party content unless they had explicit knowledge**, and did not remove the content expeditiously despite that knowledge (1); it also means that they are not obliged to monitor content actively and search for illegal activity or content (2). This is a significant development, even though we would welcome the notice-and-notice system instead of the notice-and-takedown system,[383] which is regularly exploited by malicious notifiers to have their competitors' content removed or suspended (see below at 3.1.2).

Although promising, this does not fully clarify the situation with respect to all platform providers, including those social media platforms that are not subject to the AVMS Directive (unlike YouTube). Beyond the new amendment to AVMSD, by now several other instruments appear to express a consensus view that intermediaries should not be liable for third-party content – other than for expeditiously removing them, in case they are notified. The AVMSD could not be applied – without meaningful reform – to social media platforms in general, because: (1) its scope extends to those services which provide or transmit audiovisual content, (2) it applies to "television-like" services, which are more similar to mass media rather than an individual user's content. Although the door was opened for YouTube, which is required to self-regulate under AVMSD, but a horizontal regulation of all 'platform providers' without regard to the type of content that they transmit, promises better results – the transparency of advertisements, the protection of personal data, the transparency of algorithms is not related to the content type.

[382] In Recital (48), Article 28.a 5. and 28.b.1.

[383] See a detailed explanation in: Petra Bárd, Judit Bayer, A comparative analysis of media freedom and pluralism in the EU Member States. Study for the LIBE Committee, PE 571.376 EN, 2016, http://www.europarl.europa.eu/supporting-analyses.

The **Council of Europe Recommendation (2018)** on the roles and responsibilities of internet intermediaries held that "States should ensure, in law and in practice, that intermediaries are not held liable for third-party content which they merely give access to or which they transmit or store" but they should be "co-responsible, if they do not act expeditiously to restrict access to content or services as soon as they become aware of their illegal nature"[384] also adding that "State authorities should not directly or indirectly impose a general obligation on intermediaries to monitor content".[385] The Recommendation uses the term 'internet intermediaries', but from the context it is clear that it focuses mainly on 'platform providers'.

The Communication from the Commission on tackling online disinformation encouraged the creation of the **Code of Practice on Disinformation.**[386] This **self-regulatory document** was created in September 2018 by the major online platforms, leading social networks, advertisers and the advertising industry. The Commission also published the reflections of the Sounding Board on the Code of Practice.[387] Its critical opinion signals that the platform providers did not leave their comfort zone when drafting their rules. The Sounding Board found that the so-called Code "contains no common approach, no clear and meaningful commitments, no measurable objectives or KPIs [Key Performance Indicators], hence no possibility to monitor process, and no compliance or enforcement tool: it is by no means self-regulation, and therefore the Platforms, despite their efforts, have not delivered a Code of Practice." The Sounding Board encouraged the Commission to keep a close eye on events in the period leading to the European Parliamentary elections, and regularly evaluate developments.

The Code of Practice contains – in very cautious language – commitments to make "reasonable efforts" (not even "best efforts") towards disclosing "issue-based advertising", to improve the situation in fields such as the identification of automated bots or the "impermissible use of automated systems" (points 4, 5, 6). There are a few clear commitments regarding the differentiation of advertisements from editorial content, which is already a basic principle in most jurisdictions (point 2). A vague commitment is also offered on enabling the public disclosure of political advertising, possibly including the actual sponsor identity and the amounts spent (point 3). The hesitant approach could signal two things: a) that the platform providers do not yet see any technical solutions to these expectations, or b) perhaps they are still undecided how much effort would be needed to find a balance between their financial interests on the one hand and avoiding legal proceedings on the other. (See more on the Code of Practice below at 4.2.3.2.).

The Council of Europe Recommendation (2018) on the roles and responsibilities of internet intermediaries and the Council of Europe Study on the Human Rights Dimensions of Automated Data Processing Techniques (In Particular Algorithms)[388] both declared that states should not impose a general obligation on internet intermediaries to use automated techniques to monitor information that they transmit, store or give access to.[389] No such monitoring is required by the German Network Enforcement Act, either. The **Proposal for a Regulation on preventing the dissemination of terrorist content online** is a new important tool for tackling illegal online content. On the one hand, its Explanatory Memorandum explicitly acknowledges that the immunity granted by

[384] Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries. (Adopted by the Committee of Ministers on 7 March 2018 at the 1309th meeting of the Ministers' Deputies) at 1.3.7. and

[385] Ibid at 1.3.5.

[386] Communication from the Commission on tackling online disinformation: a European Approach. Brussels, 26.4.2018 COM(2018) 236 final. p.7.

[387] The Sounding Board's Unanimous Final Opinion on the So-Called Code of Practice *24 September 2018.* https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=54456. See also: Sounding Board on disinformation looks to Action Plan to address short-comings of so-called Code of Practice. 4. Dec. 2018**.** https://www.ebu.ch/news/2018/12/sounding-board-on-disinformation-looks-to-action-plan-to-address-short-comings-of-so-called-code-of-practice

[388] Study on the Human Rights Dimensions Of Automated Data Processing Techniques (In Particular Algorithms) And Possible Regulatory Implications - Prepared By The Committee Of Experts On Internet Intermediaries (MSI-NET) - https://rm.coe.int/algorithms-and-human-rights-study-on-the-human-rights-dimension-of-aut/1680796d10

[389] Ibid. Recommendation no. 6. page 46.

Articles 14 and 15 of the e-Commerce Directive shall not be affected. The obligation of service providers extends to following concrete decision of the authorities, as it is also required by the e-Commerce Directive. On the other hand, it outlines a **vague obligation**, **a "derogation from this principle"[390] obliging "hosting service providers", "where appropriate, [to] take proactive measures to protect their services against the dissemination of terrorist content**". Article 9 clearly puts a monitoring obligation on hosting service providers, to use automated tools in respect of content that they store, adding that they shall be complemented by human oversight and verifications before decisions to remove or disable content. This obligation **effectively removes immunity for third-party content, and puts a disproportionate burden on hosting service providers.**

The broadness of this requirement raises serious concerns: besides posing a risk to the rights of hosting service providers, as well as potentially content providers, it opens the door again for divergent regulation in Member States, the avoidance of which is the primary goal of the Regulation**.** This part of the proposed Regulation – among others – has also been **criticised by a Joint Letter issued on 7 December 2018 by three United Nations Special Rapporteurs.**[391]

By now, the situation has matured sufficiently enough to give such platform providers **a place in the legal system**. Their responsibilities should relate to **their actual activities**: the design and usage of algorithms, their handling of personal data, sponsored content and its conveyance to users, and the amplification of certain content to the detriment of other content. The first attempt at self-regulation shows the necessity of legal regulation. As Mark Zuckerberg said in his hearing before the European Parliament: **the question is not whether regulation is needed, but what the regulation would be.**

3.1.2 Discussion: the 'notice-and-takedown' regime and its critique versus the 'notice-and-notice' regime

Article 12-14 of the e-Commerce Directive defines the difference between "mere conduit" – usually called access provider 'caching', meant to give exemption to search engines when they create temporary copies of content, and "hosting", whereas Article 15 declares that providers are not obliged to monitor content for illegal activity.[392]

The only conditional exemption applies to hosting providers: exemption is granted only if the provider did not have actual knowledge of illegal activity, and upon obtaining such knowledge acted expeditiously to remove or to disable access to the information concerned. This obligation structure was named in the literature as the **notice-and-takedown obligation**, and has been extensively discussed and also criticised by freedom of expression advocates and academics. The reason is that it induces ISPs to remove all content that was brought to their notice. A careful investigation into whether the notice was justified or not, and if a (controversial) content was lawful or not, would be contrary to the interests of the ISP. This system can also be exploited to further the interests of the notifier: by simply asking for the removal of a competitor's content, or of pages critical of the notifier, illegitimate advantage can be gained.[393]

The Code of Conduct on countering online hate speech follows a similar logic, and the expectations of the regulator are clearly recognisable when the increasing ratio of the content removed in proportion to the content

[390] Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online. Brussels, 12.9.2018 COM(2018) 640 final. 2018/0331 (COD) https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-preventing-terrorist-content-online-regulation-640_en.pdf, page 3.

[391] Mandates of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression; the Special Rapporteur on the right to privacy and the Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism. 7. 12. 2018. OL OTH 71/2018.

[392] 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') Articles 12-15.

[393] See in: Liability of Internet Service Providers for Third Party Content. A comparative analysis with policy recommendations. VUW Law Report Special Edition. Wellington, New Zealand. 2007. 1-109. p. 91-93.

notified (59 % from 28 %) is acclaimed as progress – without there being independent information whether the notices were in fact well-founded or not.[394]

A system which pays more respect to the freedom of speech is applied in the UK's Defamation Act 2013 Section 5, and in the Copyright Modernization Act, or Bill C-86 of Canada.[395] Under the so-called **notice-and-notice regime,** the ISP should if possible forward the notice to the actual content provider, or otherwise remove the content. This system empowers content providers and users to settle their dispute responsibly. In Canada, even this regime was quickly exploited by copyright holders until legislative amendment excluded this possibility by prohibiting inclusion of settlement demands and similar information in the notice.[396]

The **German Network Act** and **the current codes of conduct**, probably as an effect of the AVMS Directive, apply the **notice-and-takedown** principle. In sum, this procedure can be regarded as moderately chilling, as it puts a burden on intermediaries and on content providers – the advantage goes to those behind the notice. The Oxford research has identified recent cases of **malevolent** notice-and-takedown practice, in order **to suppress legitimate voices** online. Both human-operated and automated accounts were used to **falsely mass-report** legitimate content or users in Armenia, China, Ecuador and Russia.[397] The notified content or accounts are temporarily suspended even if finally, after consideration, the platform provider would reinstate them. It should also be noted that users of platform services encounter material almost exclusively of their online friends, and they always have the option to **silence a feed that they do not like**.

3.1.3 Case law of the ECJ and of the ECtHR relating to the responsibility of intermediary service providers

L'Oréal v eBay

Articles 12-14 of the Directive have not yet served as the basis for much case law. In *L'Oréal* v *eBay*,[398] the ECJ found that eBay – called an "intermediary service provider" – was not entitled to rely on the exemption from liability provided by Article 14 of the Directive, because its activity was not confined to technical and automated processing of the data relating to the offers that it stored, but it played an active role, providing the customer with assistance consisting in particular of optimising the presentation of the offers or promoting them. It further declared that it is for the national courts to carry out this assessment.

The argumentation of the Court is not sufficient to draw conclusions whether intermediary service providers could in theory be subject to the e-Commerce Directive or not. It did not consider the question that providing assistance to the customer was an automated service offered for any customer, and that it did not entail that the "operator" (yet another term used) had actual knowledge about the items to be promoted or offered for sale. Instead, it referred the decision to the national courts, adding that the national measures should not require an operator of an online marketplace to monitor the goods offered for sale through its platform. However, it did not base this opinion on Article 15 of the Directive, which explicitly declares that no such monitoring should be required. Thus, although the Court was silent on this issue, through *argumentum a contrario* we may conclude that it did not hold that the case of eBay should be handled under the Directive.

Delfi and MTE & Index[399]

The European Court of Human Rights (ECtHR) received a complaint from the online news portal Delfi, which was fined in Estonia for user comments that were posted in relation to one of its articles. While the article was found

[394] Code of Conduct on countering online hate speech – results of evaluation show important progress. 2017.06.01. http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=71674

[395] https://www.ic.gc.ca/eic/site/Oca-bc.nsf/eng/ca02920.html

[396] See more: Canadian Government Banning Settlement Demands in Copyright Notice-and-Notice System, 2018. Oct. 30. http://www.michaelgeist.ca/2018/10/noticesystemfix/

[397] Bradshaw, Samantha and Philip N. Howard: Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation. Computational Propaganda Project, University of Oxford. 2018. p. 12.

[398] C-324/09 L'Oréal and others v eBay, judgment of 12 July 2011

[399] Delfi v Estonia (App 64569/09) ECtHR 2015. and MTE and Index v Hungary (App 22947/13) ECtHR 2016.

to be lawful (and contained information of public interest) about 20 of the related 185 comments contained personal threats and offensive language directed against the owner of the company, including racist remarks. The various instances of national courts disagreed on whether the Information Society Services Act, based on the e-Commerce Directive, was applicable. The Supreme Court found that the narrow definitions of Articles 12-14 of the Directive do not apply to Delfi, as "publishing the comments [we]re not merely of a technical, automatic and passive nature".

It is worthwhile considering why eBay, Delfi, MTE and Index all believed that as online portals they could avail themselves of the exemption provided in Articles 14-15 of the e-Commerce Directive. They all regarded the questionable content as third-party content for which they bore no liability. At a minimum, it should be acknowledged that the law was not entirely clear on this: while the courts declared that they were not simple "hosting services",[400] they could not offer an alternative clause of liability.

The ECtHR even accepted without hesitation that the restriction was foreseeably laid down in law. It decided the case without regard to the issue of attribution of liability: simply on the basis of the content. In fact, the applicants should not have applied to the ECtHR, but rather should have urged their national courts to seek a preliminary judgement from the ECJ, as their main claim was not that their freedom of expression was violated, but that the content was not attributable to them. Although the applicants' main service was providing content, their ancillary service included providing a platform for comments. In this respect, their threshold of responsibility should be similar to those whose primary activity is providing platform services, such as eBay or Facebook.[401]
These cases signal a considerable level of uncertainty in the legal interpretation of the roles and responsibilities of online service providers for third-party content, which raises a concern for the **rule of law**.

In a case initiated by the Federation of German Consumer Organisations, the **Berlin Regional Court found that Facebook violated consumer law** by "hiding" its non-privacy friendly default settings in its privacy centre, and not providing sufficient information about this when users register. In addition, it found eight clauses in Facebook's terms of use to be invalid.[402]

As seen in this section, case law relating to platform providers is relatively limited and in some cases contradictory. Neither the governing law, nor jurisdiction or court competence is clear. No clear definition of principles can be observed. However, in other legal instruments of the European Union and the Council of Europe, the conditional immunity from liability for third-party content is clearly defined, with no obligation to monitor – the only exception being the Proposal for a Regulation on preventing the dissemination of terrorist content online.

3.1.4 Proposed responsibility

We have discussed above what should *not* be the responsibility of platform providers: they should not be liable for third-party content, and not be obliged to monitor third-party content. However, **they ought to be responsible for their own actions**[403] and not for their clients'.

They transmit and amplify certain content and suppress others, handle enormous amounts of personal data, and facilitate communication between advertisers, professional content providers and users. These activities clearly **interfere with human rights**, such as freedom of expression, including the right to receive information, privacy, human dignity (to be aware of reality, not to be deceived). In other branches of industry, when a few large

[400] In fact, "access services" would be closer to their activity, see Article 12. of e-Commerce Directive.

[401] In the Delfi case, the content represented hate speech, and the commenting section brought revenues to the portal, so the Court found that the moderate fine did not violate Article 10. In the MTE & Index case, the content itself did not amount beyond justified criticism, and at least MTE was a non-profit portal of public interest issues.

[402] Judgment of the Berlin Regional Court dated 16 January 2018, Case no. 16 O 341/15. The decision is not final.

https://www.vzbv.de/sites/default/files/downloads/2018/02/14/18-02-12_vzbv_pm_facebook-urteil_en.pdf

[403] Contracting with external partners should also count here, including advertisers, and app creators.

companies have a significant effect on masses of people, those industries are regulated in the interest of the public. They might even be defined as common carriers, or services of general interest, and subjected to even more regulation, such as the prohibition to deny contracting, or must-carry rules. In the context of social media, these rules could relate to giving preference to reliable sources (attested by public or non-profit institutes, fact-checkers, media-bias-checkers), to public service content (if it also fits the previous requirements).

Current recommendations and policy papers do not appear to step beyond the dichotomy of editorial responsibility vs. mediators' *distributory* (American term) responsibility. We recommend that the actual activity of platform providers is listed, and examined one-by-one as to whether it impacts human rights and democracy to such an extent that it needs regulation. Even though their activity has a strong effect on how content can be perceived, it still does not reach the level of editing, and therefore this analogy should no longer be used.

Their activity could be regarded **facilitation**, mediation or amplification – from which we recommend the term facilitation, which describes that this action is more than neutral mediation, but does not necessarily include amplification for all platform providers.

Some of the principles of these actions have already been expressed by several papers referred to in our study. The question remains how these principles can be incorporated in regulation so that they are enforceable beyond self-regulation, which raises justified scepticism at the moment.

Starting from the basis of the prohibition of discrimination, through respect for privacy, freedom of expression and the right to receive information, such regulation should also include the obligation of viewpoint neutrality.

3.1.4.1 Neutrality and diversity

Currently, the biggest platform providers do not represent any political agenda or public issue – or at least not one conspicuous from the North Atlantic perspective. But there is no legal obligation on them to keep to this. **At any time, a competing platform could emerge that would – either openly or covertly – represent a particular ideology.** In this daunting scenario, the hypothetical social media provider would tailor its algorithms to give preference to favoured views and people, without informing users of this bias. Although the owner of Facebook currently appears cooperative vis-à-vis US and EU parliamentary bodies, the company did not refrain from using disinformation and manipulation as marketing devices.[404]

It therefore appears necessary that this prohibition is included among the basic obligations of platform providers, at least above a certain size. A softer version of this could prescribe that such a viewpoint is clearly signalled to the potential and actual users, as do Catholic social network apps for example.

Dominant social media platforms have an influence on public opinion. This raises the issue of their responsibility in editing the content feed with their algorithms. **The algorithms of dominant platforms must not contain viewpoint-based discrimination, or use other discrimination on protected characteristics** (for example, race, or ethnic origin), and may be required to set their algorithms to promote diversity (see also below).

3.1.4.2 Algorithms

Algorithms are an important part of facilitating communication, but they have profound effect on how content is perceived and on many other aspects of the social media experience. Companies attest that they keep changing their algorithms, and that these form part of their intellectual property and therefore cannot be revealed to the public.

[404] Facebook policy chief admits hiring PR firm to attack George Soros. The Guardian. 22. Nov. 2018. https://www.theguardian.com/technology/2018/nov/21/facebook-admits-definers-pr-george-soros-critics-sandberg-zuckerberg

The frequent change of algorithms may be very beneficial for the design and development of new innovations, but it can appear as **a global human experiment without the consent of the participants.** While the software code does not need to be revealed to the public, it can be expected that the basic principles of algorithm usage are disclosed to users, and that they are offered options. In the context of this study, **users should be offered options** regarding their preferred level of diversity (which could be chosen on a slider),[405] or about being targeted by personalised advertisements (which is among the options now at some intermediaries, such as Google, whereas Facebook offers a more complicated setting scheme). Optional algorithm settings should also be offered to the user regarding whether to receive local news, political issues, etc.

3.1.4.3 Advertisements

Platform providers should be responsible for clearly identifying advertisements and sponsored content as such (principle of identification). This principle is horizontally applied in advertising regulations, including the AVMS Directive. Social media platforms should pause and approve all paid content before putting the advertisement into effect. As opposed to voluntary content, advertisements (including boosted posts) that ensure a revenue stream for the social media platform should be moderated – supervised by the platform provider, which **has a higher level of responsibility for these as opposed to voluntary content**.

Identification of **political advertising and public issue advertising** should also be obligatory. However, platform providers are able to label such content only if their publisher is explicit about the payment factor and the subject matter of the content. But, politicians and political parties can currently use social media on equal terms with private individuals, and professional media companies, which voluntarily publish politically motivated content as a legitimate exercise of their freedom of expression. Political parties and professional politicians should be regarded as '**influencers'**[406] in respect of political content, which is not independent information even if it is not sponsored.

Platform providers should not decide on the political nature of content. Political jokes, opinion articles, short but emotional tweets could be very difficult to identify properly. For example, **symbolic political references can be understood only in the local context** (for example soccer games, or featuring a public figure such as Beyoncé could have a political implication).

Therefore, labelling political content based on its topic would pose a risk to freedom of expression. Political speech – whether sponsored or voluntary – forms a key part of free public discourse, but it may be subject to restrictions.[407] A basic requirement should be to inform users about the sources of political content when it comes from committed actors such as politicians and political organisations.

Thus, for the purpose of identifying politically-motivated content from independent content (even if it carries political opinion) **politicians, political organisations and other actors that regularly publish information of public interest and attract a sizeable audience should be labelled as 'influencers'**. This would also add to the trustworthiness of their content: having this label would prove that the source is authentic. Such an identification is also helpful in the supervision of campaign financing. Making political and public issue campaigns transparent and at the same time protecting individual users' privacy would not be possible if they are not somehow distinguished.

3.1.4.4 Fake accounts

Platform providers could be responsible for making their best efforts to ensure that their accounts are registered by human individuals, or registered organisations – as opposed to robots and creators of pseudo accounts.

[405] Helberger, N., Karppinen, K., & D'Acunto, L. (2018). Exposure diversity as a design principle for recommender systems. *INFORMATION COMMUNICATION & SOCIETY*, *21*(2), 191–207. https://doi.org/10.1080/1369118X.2016.1271900

[406] The word "influencer" is used here to political parties, active politicians, governments and public authorities, NGOs and communication agencies and the new brand of influencers themselves.

[407] Animal Defenders International v. the United Kingdom. 48876/08. Judgement of 22.4.2013.

Facebook already requires verification through email or mobile number, and limits the number of accounts per email address to one. Although not explicitly stated that bots and AI personalities cannot register, the real name policy obviously rules this out.[408] The **real name policy was found unlawful under German law, saying that users should be allowed to use platform services anonymously, using pseudonyms**.[409] We agree that the use of pseudonyms should be allowed, as long as '**influencers' can be unambiguously identified as such**.[410] This is in accordance with the Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression[411] which, too, prefers instead the so-called **'impersonation policy**': "narrowly crafted impersonation rules that limit the ability of users to portray another person in a confusing or deceptive manner".

Artificial intelligence acting as an individual user is not without precedent: Lil Miquela, a fictitious model already has 1.5 million followers on Instagram, while Bermudaisbae – a devoted supporter of Donald Trump and white supremacy, denying climate change, is also present, although with only 736 followers on Twitter. Both are virtual models, and so is blawko22 (133 000 followers on Instagram). While Miquela started as an arts project, these models appear to be on the rise, and they claim 'sentience' – a first step before claiming their rights.[412] While nothing prevents an artist or anyone else to have one or two of their profiles featuring a virtual person, the owner of the profile should be disclosed and be identifiable. We should also count on the emergence of virtual politicians as well, after the first was launched in 2017.[413] The sound functioning of democracy, transparency and human dignity are strong arguments, while the virtual personalities do not have fundamental rights (yet).

Thus, in the context of administering their platforms, platform providers – or at least dominant platform providers – should be required by law to use some low level of verification, but not full user identification, which would amount to an unnecessary intrusion.

3.1.4.5 Protection of privacy and personal data

Social media acquires vast amounts of personal data, including special categories of personal data[414] about its users. The General Data Protection Regulation (GDPR) creates obligations on social media companies to process personal data lawfully, fairly, transparently, etc. Recent case law of the CJEU recognises social media companies as **joint controllers**, alongside the entities who use their services by, for example, creating social media pages.[415] However, the de facto implementation of these obligations is often insufficient as the recent scandals have shown.[416] As already mentioned before, the robust *de jure* protections offered by the GDPR will only be as effective as their de facto enforcement, led by the national data protection authorities, whose independence and resources must be secured by the Member States and rigorously monitored by the European Commission.

[408] Already envisaged in: European Group on Ethics in Science and New Technologies. Artificial Intelligence, Robotics and 'Autonomous' Systems. March 2018. http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf

[409] Judgment of the Berlin Regional Court dated 16 January 2018, Case no. 16 O 341/15. The decision is not final. See: Facebook's default settings and some of its terms of service and privacy policies are in breach of consumer law. https://www.vzbv.de/sites/default/files/downloads/2018/02/14/18-02-12_vzbv_pm_facebook-urteil_en.pdf

[410] See footnote 391.

[411] United Nations, Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 6 April 2018, paragraph 30. (https://undocs.org/A/HRC/38/35).

[412] Miquela had not originally disclosed that she was artificial - she admitted this only after she was pushed by Bermudaisbae, who hacked her account.

[413] Meet the world's first virtual politician. 2017. dec. 15. https://www.victoria.ac.nz/news/2017/12/meet-the-worlds-first-virtual-politician

[414] Article 9, GDPR - sometimes called sensitive data in common language.

[415] CJEU [GC] decision in the Case C-210/16 *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH,* 5 June 2018.

[416] See e.g. European Parliament resolution of 25 October 2018 on the use of Facebook users' data by Cambridge Analytica and the impact on data protection (2018/2855(RSP)), http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2018-0433+0+DOC+XML+V0//EN

Another 'leg' of the EU data protection law – the ePrivacy Directive[417] – also contains relevant obligations (e.g. with respect to confidentiality of communications, use of communications' data – i.e. content and metadata, direct marketing and cookies). However, it does not explicitly outlaw 'tracking walls' (conditioning access to websites upon the individual being forced to 'consent') or by default prevention of tracking individuals' digital footsteps. The new ePrivacy Regulation, which sufficiently addresses these issues, is urgently needed.[418]

Stakeholders may regard personal data as an asset, and a basic capital in their business activity, but internet legal scholar Jonathan Zittrain warns: doctors, lawyers, investment brokers also deal with huge masses of (sensitive) personal information, and they must eschew monetising these in their own interest.[419] All are obliged by special regulations to keep client data secret, and we could add journalists and their sources to this list. Given the considerable civic powers of social media platforms (see section 7.1.9. on Disruptive technologies and their implications for democracy), it can be argued that their data protection-related obligations should exceed the minimum requirements enshrined in EU and national data protection laws.

For instance, the lack of transparency in the ad tech industry makes it complicated to trace the original advertisers (they may intentionally conceal their identity or act through intermediaries) and the **limited transparency solutions offered by digital platforms are often insufficient**.[420] For example, although Facebook offers users some explanation about why they were targeted with a certain ad, it does not provide an access to a searchable repository of all ads targeted to an individual user and a repository of all ads purchased by specific commercial or non-commercial advertisers.

The problematic areas listed above in 3.1.4. outline the proposed scope of responsibility of platform providers, without affecting restriction of content.

3.2 The limits of national and international rules that set a limit to freedom of expression applying to the international environment of social media

3.2.1 Assessment of international legal standards for possible ways of tackling disinformation and propaganda

This section will examine the legal frameworks to counter disinformation and propaganda under international legal agreements and principles. The problem of disinformation and propaganda violates the moral values of societies, the principles of ethical journalism, several human rights, and harms democracies. However, most of the known cases escape the boundaries of legal categories. In the following subsections we will examine how existing laws can be related to this complex phenomenon. Due to the constraints of this study, the discussion of existing national laws is limited to illustrating our thesis with examples. The discussion of international legal

[417] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), Official Journal L 201, 31 July 2002.

[418] European Data Protection Supervisor, Opinion 6/2017 on the Proposal for a Regulation on Privacy and Electronic Communications (ePrivacy Regulation), https://edps.europa.eu/sites/edp/files/publication/17-04-24_eprivacy_en.pdf. Note that although under the current ePrivacy Directive confidentiality protections do not apply to the over-the-top (OTT) communications services such as Skype, Viber and instant messaging apps, its scope of application has been recently broadened by virtue of the European Electronic Communications Code. For more information, see: Buckwell, M., New European Electronic Communications Code means the application of the ePrivacy Directive to OTTs, 21 December 2018, https://iapp.org/news/a/new-european-electronic-communications-code-means-the-application-of-the-eprivacy-directive-to-otts/

[419] Zittrain, Jonathan, Engineering an Election (June 20, 2014). Harvard Law Review Forum, Vol. 127, p. 335, 2014; Harvard Public Law Working Paper No. 14-28. Available at SSRN: https://ssrn.com/abstract=2457502, at 340.

[420] See e.g. Transparent Referendum Initiative, Questions we're working on part 2: When is an ad political? 27 February 2018, https://medium.com/@TransparentRef/questions-were-working-on-part-2-what-counts-as-a-political-ad-d410209c5df6; Ghosh, S., Facebook approved fake political ads 'paid for' by Cambridge Analytica, 3 October 2018, https://nordic.businessinsider.com/facebook-approved-political-ads-paid-for-by-cambridge-analytica-2018-10?r=US&IR=T

instruments of human rights can be regarded as a more rounded, yet very concise illustration of how international human rights treaties relate to this issue.

Below we will assess the possibilities of future legislation to permit the restriction in future of dissemination of disinformation and propaganda using legal tools. While it may be tempting to introduce new legislation that would prohibit the dissemination of disinformation and propaganda, the core of these concepts cannot be defined sufficiently narrowly for such legislation to withstand constitutional scrutiny.

The principles of freedom of expression enjoy high esteem globally, primarily in the United States and in European countries where the European Court of Human Rights actively safeguards those human rights enshrined in the European Convention on Human Rights. The Charter of the European Union binds the European institutions and Member States when they apply European law, and can also be regarded as a limit to legislation.

Freedom of expression serves – as already mentioned in this study – deliberation of public issues in a democracy, and also the fulfilment of individual human potential. But this freedom protects not only the useful and favourably received ideas and information, but also those which 'offend, shock or disturb'. "Such are the demands of that pluralism, tolerance and broadmindedness without which there is no 'democratic society'. This means, among other things that every 'formality', 'condition', 'restriction' or 'penalty' imposed in this sphere must be proportionate to the legitimate aim pursued."[421]

Although freedom of expression does not enjoy absolute protection, the possibilities of restriction are limited. One of the basic tenets of the protection of freedom of expression is that content-based restrictions should be kept to a minimum. In American jurisprudence, any restriction based solely on the content (rather than the context, the speaker or the effect) requires evidence of a compelling state interest, and should be narrowly tailored. Acknowledging that this carries in itself the possibility of error, but "the people in our democracy are entrusted with the responsibility for judging and evaluating the relative merits of conflicting arguments".[422] Suppressing false information or harmful political ideas by the government "usurps the right of the people to make such decisions for themselves".[423] In European jurisprudence, the only content-based restriction without regard to the context and the effect, is the denial of the Holocaust (see below).

Of course, several items of disinformation could be held illegal under one or other of the existing legal rules: defamation and libel are restricted in all jurisdictions. However, several items do not reach the threshold of illegality, for various reasons: they do not hurt a certain person (therefore they are not libel); they do not reach the threshold of incitement towards a racial group, or minority; they do not feature prohibited symbols, or symbols of unconstitutional organisations; their material does not resemble any of the views of already banned unconstitutional organisations; they do not relate to public emergency situations (scaremongering); they do not call for sabotage; and do not defame the state as such, nor constitutional organs. This list is far from being exhaustive, but it clearly illustrates that even if some jurisdictions (e.g. German) have ample criminal measures of militant democracy in place to protect state stability from concerted communicative actions, they only apply to specific cases, sometimes under specific circumstances. These rules are sufficiently narrowly defined to be in line with international standards of freedom of expression.

Widening the frames of such definitions would conflict with the freedom of expression principles in Article 10 of the ECHR. Below we analyse the ECtHR's principles of assessment used to decide on whether a restriction (or a legal instrument in that case) is in accordance with the Convention. (1) The restriction should be laid down foreseeably in a law, (2) directed at a legitimate objective, in other words, there must be a connection between

[421] Handyside v. the United Kingdom judgment of 7 December 1976, § 49.

[422] First Nat'l Bank v. Bellotti, 435 U.S. 765, 792 (1978).

[423] Geoffrey R. Stone: Restriction of Speech Because of Its Content: The Peculiar Case of Subject-Matter Restrictions. University of Chicago Law School. Chicago Unbound. 1978. https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1530&context=journal_articles

the policy and the achievement of the goal and (3) be necessary in a democratic society. In addition, the applied measurement should be proportionate to achieving the desired objective. Should Member States envisage drafting new laws against disinformation and propaganda, they should keep in mind the above points. The restrictions should be as narrowly defined as possible, and the more closely the objective is associated with the right of an individual, the stronger is the justification. The case of disinformation and propaganda calls for a protection of the **public interest: national security, territorial integrity, public safety, the prevention of disorder or crime, the protection of health and morals**, ... preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.[424] These are equally legitimate goals as the protection of the rights of persons, but the correlation between the expression and the goal may be more difficult to draw. The legitimate aim should be realistically be protected with the restriction, and this goal should not be achievable with other, less restrictive tools.[425]

While the overall purpose of the fight against disinformation can be national security, territorial integrity and public safety, the connection of an individual piece of fake news item to these goals may be very limited. However, prohibiting concerted actions that are clearly aimed at inciting instability in society have a chance to withstand constitutional scrutiny. These considerations are reflected in our recommendation for a certain criminal rule in chapter 6 (Conclusions). A criminal prohibition of speech should be very narrowly and precisely formulated to be in line with human rights standards. For example, containing subjective elements such as the ill intent to mislead, and to negatively affect the society, or to gain financial or political advantages, and objective elements such **as originating from disguised sources or identities, using artificial intelligence to boost dissemination**. However, if the objective circumstances can also be made illegal individually, then content restriction may become unnecessary.

A regulation which restricts the technological possibilities (coding) can prevent such actions without the need to police human behaviour: for example, if disguising identities is not possible under normal circumstances, if political advertisements must designate their source, and be transparent, if artificial dissemination techniques can be used for political purposes only under transparent circumstances, then a safer online environment can be maintained without unnecessarily using the instrument of criminal law, which should only be the *ultima ratio* of enforcement.

The last important aspect is proportionality: thus, when legislating against disinformation, the applied sanction is also decisive from constitutional perspective. For example, the measure to add a warning label to suspicious content that informs the people about the possible nature of such content, is of a lighter interference with free speech rights, than the complete removal of the content.[426]
The ECtHR has found that a sentence of imprisonment did not violate Article 10 relating to the statement: "I support the PKK national liberation movement; on the other hand, I am not in favour of massacres. Anyone can make mistakes, and the PKK kill women and children by mistake…". The statement was published in *Cumhuriyet*, the largest national daily paper, in an interview with the former mayor of Diyarbakır, the most important city in south-east Turkey, in 1987.[427] The Court had to "ascertain whether a fair balance has been struck between the individual's fundamental right to freedom of expression and a democratic society's legitimate right to protect itself against the activities of terrorist organisations."[428] The Court carefully considered the context, given that the

[424] Article 10. (2), protection of the reputation or rights of others are omitted from the list, for the reason that it is already sufficiently regulated in all Member States, especially in the light of disinformation.

[425] Kai Möller; Proportionality: Challenging the critics, *International Journal of Constitutional Law*, Volume 10, Issue 3, 1 July 2012, Pages 709–731, https://doi.org/10.1093/icon/mos024

[426] See Times Newspapers Ltd v. United Kingdom (Nos. 1 and 2) (Application nos. 3002/03., 23676/03., March 10. 2009); Wegrzynowski and Smolczewski v. Poland (Application no. 33846/07., July 16. 2013.)

[427] Zana v. Turkey. Application no. 18594/91. 1997. Nov. 25.

[428] Ibid. at 55.

interview coincided with murderous attacks carried out by the PKK on civilians in south-east Turkey, where there was extreme tension at the time.

From the perspective of today's struggle against misleading and manipulative information, one of the most important factors for the Court was that the words" could be interpreted in several ways but, at all events, they are both contradictory and ambiguous".[429] Thus, instead of interpreting them as balanced or moderate, the Court sensed **a threat in the ambiguity**, one which could further raise tensions in an already bloody conflict. Therefore, the Court accepted that the governmental restriction responded to a pressing social need. Considering that only one fifth of the sentence had to be spent in prison, and the rest on parole, the Court did not find the sanction disproportionate.

On the other hand, in the case *Kommersant Moldovy* v *Moldova*,[430] the Court established violation of Article 10. The court of Moldova ordered termination of a journal which published a serial critically discussing the actions of Moldovan authorities against separatists in the Transnistria. The Court accepted the protection of national security and territorial integrity as legitimate reasons, but it held that the national courts did not define which elements of the articles were problematic or how exactly they threatened national security and integrity. Aside from the lack of reasoning, this case clearly had a deficit in proportionality as well.

3.2.1.1 Hate speech

The International Covenant on Civil and Political Rights (ICCPR) creates an obligation on States Parties to prohibit hate speech. Article 20(2) provides that: "Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law".

The United Nations' International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) obliges signatories to adopt measures to eradicate all incitement to discrimination:

> declaring an offence punishable by law all dissemination of **ideas based on racial superiority** or hatred, incitement to racial discrimination, as well as all acts of violence or incitement to such acts against any race or group of persons of another colour or ethnic origin, and also the provision of any assistance to racist activities, including the financing thereof;
> - Shall declare illegal and prohibit organizations, and also organized and all other propaganda activities, which promote and incite racial discrimination, and shall recognize participation in such organizations or activities as an offence punishable by law" (Article 4 a), and b))

With very few exceptions (Myanmar, South Sudan, Malaysia, North Korea), all states have signed and ratified this treaty. Based on this obligation**, even the United States, along with other European Member States should adopt criminal laws or offences against dissemination of ideas based on racial superiority, and prohibit such organisations**.

The ECHR does not mention hate speech literally, but its Article 17 provides for the prohibition of abuse of the rights listed. Based on Article 17, the Court systematically rejects those complaints that claim protection for racially-biased expressions. **Article 17 is an example to militant democracy: it was created specifically to resist the revival of totalitarian regimes, and to exclude protection for any anti-democratic activity.[431]** Critical commentators complain of the injury caused to the consistency of legal protection, because in these cases the Court applies a content-based restriction without reference to the context and other circumstances.[432]

[429] Ibid. at 58-59.

[430] Application no. 41827/02.

[431] Mark E. Villiger: "Article 17. ECHR and Freedom of Speech in Strasbourg Practice" in Josep Casadevall [et al.] (szerk.): *Freedom of Expression: Essays in Honour of Nicolas Bratza* (The Netherlands: Wolf Legal Publishers 2012) 321, 322.

[432] Antoine Buyse: "Dangerous Expressions: The Echr, Violence and Free Speech" *ICLQ* 2014/April. 491–503; Hannes Cannie – Dirk Voorhoof: "The Abuse Clause and Freedom of Expression in the European Human Rights Convention: An Added Value for Democracy and Human Rights Protection?" *Netherlands Quarterly of Human Rights* 2011/1. 54–83, 57.

However, in fact it is only Holocaust-denial that is consistently rejected in the admissibility phase based on this article. For example, in *Garaudy* v *France*, the Court rejected the argumentation of "quest for the truth, historical research":

> There could be no doubt that disputing the existence of clearly established historical events, such as the Holocaust, did not constitute historical research akin to a quest for the truth. The real purpose of such a work was to rehabilitate the National-Socialist regime and, as a consequence, to accuse the victims of the Holocaust of falsifying history.
>
> The Court found that, since the applicant's book, taken as a whole, displayed a marked tendency to revisionism, it **ran counter to the fundamental values of the Convention, namely justice and peace**. The applicant had sought to deflect Article 10 of the Convention from its intended purpose by using his right to freedom of expression to fulfil ends that were contrary to the Convention. [433]

Among the Court's values, the protection of the Convention is clearly the priority: "[T]here is no doubt that any remark directed against the Convention's underlying values would be removed from the protection of Article 10 ... by Article 17".[434]

Other expressions inspired by totalitarian doctrine or that express ideas representing a threat to the democratic order were rejected somewhat inconsistently based on this article or another.[435]
Similarly, those complaints where the expressions were liable to lead to the restoration of a totalitarian regime were rejected.[436] The Council of Europe provided in June 2018 a comprehensive collection of those categories of speech, commonly called "hate speech", which do not enjoy the protection of Article 10 of the Convention. This document, which we prefer not to reiterate here, gives several concise examples to how the Court handles such cases.[437]

In sum, **there is still room to introduce further restrictions on speech that is harmful to the public interest in the realm of hate speech**, the restriction of which is a more accepted practice in the international community, especially bearing in mind that the ICCPR and ICERD clearly and explicitly call for its prohibition. Even though the ECHR does not specifically do so, Article 17 and the practice of the ECtHR clearly convey that such content does not enjoy protection. The requirement for the legislative would be to make the law as precise and narrow as possible, and to keep the sanctions proportionate. While libertarian theory dislikes hate speech restrictions, they could be **useful instruments of militant democracy.**

3.2.2 Actions of the UN

While the Council of Europe has a large number of recommendations, centred around freedoms, human rights, as well as diversity and quality of the media, the UN operates on different terms. UNESCO has created a complex training course for journalists, tackling "Fake News' and Disinformation".[438]

[433] *Garaudy v France,* Decision of 7. July 2003, no. 65831/01.

[434] *Seurot v. France no. 57383/00), 18 May 2004 (decision)*

[435] Lásd *Glimmerveen and Hagenbeek v The Netherlands,* Judgment of 11 October 1979, nos. 8348/78, 8406/78; *Schimanek v Austria,* Judgment of 01 February 2000, nos. 32307/96, 12774/87; *H., W., P. and K. v Austria,* Judgment of 10 December 1989, no. 12774/87; *Norwood v UK,* Judgment of 16. November 2004, no. 23131/03.

[436] Communist Party of Germany v. the Federal Republic of Germany, decision of the European Commission on Human Rights of 20 July 1957; B.H, M.W, H.P and G.K. v. Austria (application no. 12774/87), decision of the Commission of 12 October 1989; Nachtmann v. Austria, decision of the Commission of 9 September 1998; Schimanek v. Austria, decision of the Court on the admissibility of 1 February 2000.

[437] Factsheet - Hate speech. 2018. https://www.echr.coe.int/Documents/FS_Hate_speech_ENG.pdf

[438] Cherilyn Ireton and Julie Posetti: Journalism, 'Fake News' & Disinformation. Handbook for Journalism Education and Training. 2018. UNESCO – a training course description.

The **UN Special Rapporteur on Freedom of opinion and expression**, David Kaye has repeatedly expressed his concerns, not so much about disinformation but rather the exaggerated restrictive responses envisaged by states. In a **Joint Declaration** of leading monitors of freedom of expression around the world, Kaye and representatives of the OSCE, OAS and ACHPR expressed concern that the efforts by states to counter disinformation could lead to censorship, the suppression of critical thinking and other approaches contrary to human rights law. The Joint Declaration reminds states to respect international human rights standards and to only impose restrictions in accordance with the test for such (see above). It also recalls that "to prohibit advocacy of hatred on protected grounds [such as race, ethnic origin, etc.] that constitutes incitement to violence, discrimination or hostility", is in accordance with Article 20(2) of the International Covenant on Civil and Political Rights.

It also states further basic tenets, which are also emphasised in this study, such as:

- intermediaries should never be liable for third-party content;
- it is not possible to prohibit content on the basis of its falsity or objectivity, as these are ambiguous concepts and this would be contrary to international standards of freedom of expression.

The **Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression** contains meaningful concerns and recommendations for states and for companies (including platform providers). It calls upon states to **only seek to restrict content pursuant to an order by an independent and impartial judicial authority, and in accordance with due process and standards of legality, necessity and legitimacy**, and to refrain from establishing laws or arrangements that would require the "proactive" monitoring or filtering of content, and also from adopting models of regulation where government agencies, rather than judicial authorities, become the arbiters of lawful expression. They should avoid delegating responsibility to companies as adjudicators of content, which empowers corporate judgment over human rights values to the detriment of users[439] (this underlines the **preference for notice-and-notice procedure).** Further it recommends ICT companies **recognise international human rights standards as their authoritative standards rather than their own private interests** or various national laws. It also emphasises that the impact of these companies on public sphere demands that they **open themselves up to public accountability.**

The UN also expressed its critical position in respect of the proposed Regulation on preventing the dissemination of terrorist content online in a **Joint Letter** and called attention to the risk of infringement to the right to access to information, freedom of opinion, and expression. This proposed Regulation **would interfere with the no-monitoring principle of service providers**, and proposes that Member State authorities could impose various undefined obligations on hosting providers[440] (see more above in 3.1.1.2.).

The responsibility of platform providers could also be shaped by the **UN Guiding Principles on business and human rights**, which provides for the corporate responsibility of all business enterprises to respect human rights: to avoid causing or contributing to adverse human rights impacts through their own activities, and address such impacts when they occur, as well as to seek to prevent or mitigate adverse human rights impact that are directly linked to their operations, products or services by their business relationships, even if they have not contributed to those impacts.[441]

[439] United Nations, Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 6 April 2018, paragraph 66-67-68, 70-72. (https://undocs.org/A/HRC/38/35).

[440] Mandates of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression; the Special Rapporteur on the Right to Privacy and the Special Rapporteur on the Promotion and Protection of Human Rights and Fundamental Freedoms while Countering Terrorism, OL OTH 71/2018, Geneva.

[441] UN Guiding Principles on Business and Human Rights Implementing the UN "Protect, Respect and Remedy Framework". 2011. OCHCHR. https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

4. DESCRIPTION AND ASSESSMENT OF POLICIES AND MEASURES DEVELOPED AT MEMBER STATE AND EUROPEAN LEVEL TO COUNTER PROPAGANDA AND DISINFORMATION THREATS

KEY FINDINGS

- State-run fact-checkers cannot sufficiently demonstrate their selection criteria, their due process and systematic methodology, which makes them vulnerable to criticism.
- National legislative attempts are divergent and lack impact analysis. Their operation needs to be assessed regularly.
- Germany's Network Enforcement Act does not sufficiently address disinformation and propaganda, because the criminal rules to which it refers only restrict false information in certain specific cases. However, it may exercise a chilling effect on freedom of expression.
- French law puts the burden of decision on courts, which carries the risk of judicial overload.
- French authorities have entered into a co-regulation agreement with Facebook. Its operation and impact need to be assessed by future research.
- Self-regulation of platform providers may bring positive changes, in particular efforts to increase exposure to more diverse content. The Code of Practice contains commendable principles which should be achieved – either through self-regulation, or by the force of law.
- The Action Plan against Disinformation has the potential to help Member States to operationalise its recommendations, but it contains no plan in case of division within the European Union.
- One of the primary drivers of disinformation and propaganda is the advertising-based business model of platform services. This needs to be changed before self-regulation can deliver effective results.
- Electoral laws in Member States are divergent. Fake news and disinformation are not explicitly mentioned in election acts and electoral codes.
- Registering campaign expenditures could ensure more transparency if it is coupled with relevant data on the advertiser, in whose interest it is published, and the platform where it is published.
- The European Parliament (EP) can monitor campaign expenditure of political parties during the EP election campaigns, but this monitoring currently does not ensure sufficient information.
- Sensitisation and awareness raising were key preventive instruments in recent elections, addressed at all segments of the population.
- Other promising practices were investment in investigative journalism, fact-checking and a credibility index to increase resilience in the media system, as well as information literacy.
- Campaign silence regulations have limited use in the context of transborder online media, as international sources and private messaging can still carry last-minute disinformation attacks.
- Political expressions during election campaigns enjoy the highest protection of freedom of expression. Self-regulation and engagement with the principles of fairness (including campaign silence) by political parties would be desirable.
- Tackling disinformation and propaganda requires the cooperation of all social actors and stakeholders, from business, media and political parties to educational institutions and NGOs.

4.1 The challenge of managing national and EU policies and measures to counter propaganda and disinformation threats

As national regulators promise to counter the proliferation of online disinformation across the bloc, a panoply of measures have been – or are about to be – adopted within and across European Member States and the EU. When designing any state intervention aimed at countering disinformation, some major issues must first be addressed. First, the definition of disinformation would be indispensable (see section 1.1. in this study). Second, a definition of platform providers would be necessary (see subsection 3.1.1. of this study).

Under the online model, the editor-producer maintains control over the content of articles but not over their distribution or curation. While traditional media editors seek to preserve and strengthen consumer trust in their newspaper brand, platform providers prioritise advertising revenue and traffic, with little regard for the quality of content and consumer trust in that quality. The separation of the roles of content provider and platform provider is at the core of the ongoing current debate on whether and how public authorities should step in.

The following sections examine the major approaches available to policy makers, both at the national and EU level. In particular, they focus on the following aspects.

4.1.1 State interventions

According to the most prescriptive and intrusive model, public authorities are expected to control the media environment by themselves (5.2.1.1) or require others to do so (5.2.1.2). This approach has been criticised insofar as it entails censorship – whether censorship by authorities, or censorship by private actors, both can stifle freedom of expression. This approach has been introduced on both sides of the Atlantic by two public-sponsored and state-run, dedicated entities.

4.1.1.1 State-run fact-checkers

In the US, this is the case of the recently-created Global Engagement Center,[442] which helps the US government ensure that streams of data are not contaminated by state-sponsored misinformation or falsehoods. The US Secretary of State established the Global Engagement Center (GEC) in April 2016 pursuant to Executive Order 13721. The GEC is charged with leading the U.S. government's efforts to counter propaganda and disinformation from international terrorist organisations and foreign countries. Its declared mission is to "**lead, synchronize, and coordinate efforts of the Federal Government** to recognize, understand, expose, and counter **foreign state and non-state propaganda and disinformation efforts** aimed at undermining United States national security interests".[443] The 2017 National Defense Authorization Act (NDAA) expanded the GEC's mission to include countering the adverse effects of state-sponsored propaganda and disinformation.

In Europe, the European Union entrusted a similar role to the East Stratcom Taskforce, which is also responsible for the *Disinformation Review*.[444] This new service was created as a conclusion of the European Council meeting on 19 and 20 March 2015 after identifying the need to challenge the ongoing disinformation campaigns by Russia and was set up within the European External Action Service (EEAS).

Today the *Disinformation Review* involves a network of 400-plus experts, journalists, officials, NGOs and Think Tanks in over 30 countries reporting disinformation articles to EU officials, and then to the public. Its declared mission is to debunk fake news and Russian propaganda. The *Disinformation Review* has been described as "the

[442] To know more, check the GEC website on the US State Department Website: https://www.state.gov/r/gec/.

[443] 2017 National Defense Authorization Act (NDAA)

[444] https://eeas.europa.eu/headquarters/headquarters-homepage_en/9443/Disinformation%20Review.

best weekly disinformation bulletin anywhere in the West"[445], but also contested for its **alleged lack of methodology and for not ensuring due process.**

In the Summer 2017, a request for documents to the EEAS was introduced to seek further information about the EEAS East Stratcom Team,[446] namely the criteria it uses to identify disinformation/fake news, and how it notifies/interacts with entities that are placed on the *Disinformation Review*.[447] Additionally, the same request for access to documents asked how the Task Force selects members for its network of academics and NGOs.

The response from the EEAS[448] conceded the absence of a clear set of criteria for labelling disinformation/fake news, and stated that the Task Force does not systematically communicate with any entity listed on the *Disinformation Review*. Furthermore, the EEAS did not clarify in that response how it selects its fact-checking partners and how they can join the stakeholder network. The concern put forward by the freedom of information request is that the criteria are vague and subjective, and the review violates due process in relation to enlisted sources of information.

A similar approach has been embraced in Sweden. The previous government announced the creation of a domestic authority: the new "psychological defence" (*psykologiskt försvar*) authority. The Swedish Civil Contingencies Agency has worked with the Swedish Election Authority, the security police and the national police to tackle foreign interference in the 2018 election. (See also in section 4.3.)

4.1.1.2 State-imposed Third-Party Liability: German, French and Italian laws

An alternative, equally prescriptive form of state intervention consists of imposing penalties on entities that engage, not just in content-creation but even mere circulation of 'illegal content'. A good example of this is the **German Network Enforcement Act** (NetzDG, German: Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken, also known as the Facebook Act),[449] which entered into force on 1 October 2017 and has been effective since January 2018. [450]

After heated discussions, it was accepted in a "watered down" version. It applies only to those social media providers with at least two million registered users within Germany.[451] The law did not set out new restrictions for content, but relied on existing criminal sections within the German Criminal Code, listing those sections and subsections relevant for the act. The relevance of this new Act was to introduce very short deadlines for removal of notified content, and to threaten meaningful fines in case this was not complied with. In addition, it required platform providers to maintain a transparent procedure to handle user complaints. In parallel, these platform providers such as Facebook, Twitter, and YouTube, are required to submit public reports detailing how many posts were flagged and how many reports were removed.[452]

Upon notice, removal or blocking of content which is "manifestly illegal" within 24 hours, and within 7 days in other cases is required. If the unlawfulness of the content depends on the falsity of a factual allegation, or on other factual circumstances, the social media provider can give the user an opportunity to respond to the

[445] Lecture by Edward Lucas: https://www.youtube.com/watch?v=sx_cqNR3bJA

[446] The East StratCom Team is a part of the administration of the European union, focused on proactive communication of EU policies and activities in the Eastern neighbourhood and beyond.

[447] https://www.asktheeu.org/en/request/eeas_east_stratcom_task_force_po

[448] Response to the freedom of information request by Alberto Alemanno. 27 Sep 2017. eeas.sg.affge.2 (2017) 4999047. https://www.asktheeu.org/en/request/4559/response/14630/attach/html/4/Reply%20to%20request%20under%20regulation%201049.pdf.html

[449] http://www.bmjv.de/DE/Themen/FokusThemen/NetzDG/NetzDG_EN_node.html.

[450] Modifying the Netzdurchsetzunggesetz of 2017, available at: https://germanlawarchive.iuscomp.org/?p=1245

[451] Exceptions apply to platforms offering journalistic or editorial content, for which the service provider is responsible, as well as to platforms designed to enable individual communication or the dissemination of specific content. It remains unclear the categorisation of services offering ratings, music or games that have therefore to be assessed on a case-by-case basis.

[452] The law also contains a half-yearly reporting obligation and a requirement for social media providers established outside of Germany to have an authorised legal person to receive service within Germany.

complaint before a decision is reached. In this case the assessment can be delayed beyond seven days. In addition, platform providers were obliged to create a report in the German language twice a year, and publish them in the Federal Gazette as well as on their own websites (this applies only to those platform providers in receipt of more than 100 complaints per calendar year about unlawful content). The content of the report is defined in the Act in painstaking detail. The fines for not complying with the above-mentioned obligations are considered high (probably not for Facebook): an initial fine of EUR 5 million, which could rise up to EUR 50 million, depending on which obligation was violated.

In comparison to the French bill, this law is **not restricted to electoral campaign period, and nor does it specifically target disinformation.** Rather, – as its name suggests – **it provides for the enforcement of already existing restrictions of free speech in the German Criminal Code**, such as hate speech, Holocaust-denial, defamation, threats of violent action, and **propaganda**, which comes closer to the subject of this study (section 86 of the German Criminal Code: "Dissemination of propaganda material of unconstitutional organisations"). However, the scope of the rule is narrowed down to certain actors: the offence can only be committed by banned political organisations, or a government, organisation or institution outside Germany which pursues similar objectives to one of the banned political organisations. The definition of propaganda material is "written materials ... the content of which is directed against the free, democratic constitutional order or the idea of the comity of nations".

The entity in charge of implementing the legislation is the Federal Office of Justice, which reports directly to the Minister of Justice. Although it does not, in principle, play a role in the day-to-day analysis of the truthfulness of the supposed disinformation content, it is not an independent institution as it is clearly linked to the government. Within the first month of application of the law, a large number of messages were deleted, including hate speech from far-right German party AfD (Alternative for Germany), but also from online satirical newspapers, such as Titanic.[453] The backlash effects on freedom of speech are already visible and critics call for the abrogation of the law "the desire for protecting political culture – plausible as it is – does not suffice as a basis to limit human rights; it burdens the NetzDG with the complexity of an act of moral regulation".[454]

According to much of the criticism, the pressure on platform providers led to large quantities of lawful content being deleted, because they have an incentive to err on the safe side.

The UN's Special Rapporteur on Freedom of Expression has written to the German government to warn about the potential consequences of its law. "With these 24 hour and seven day deadlines – if you are a company you are going to want avoid fines and bad public branding of your platform," he says. "If there is a complaint about a post you are just going to take it down. What is in it for you to leave it up? I think the result is likely to be greater censorship." This seems to be confirmed by empirical data collected since the entry into force of the law. The Federal Office of Justice has received a limited number of complaints with regard to social network providers failing to comply with the NetzDG. This suggests that social network providers are removing all "unlawful content", thus acting as arbiters of truth within their own platforms. **There is therefore collateral damage to freedom of speech and communication on social networks insofar as the network's legally-mandated action may result in lawful content also being blocked.** This seems problematic as it may effectively censor content that might form an important part of the social and cultural debate that must exist in any liberal democracy.

[453] See: https://www.theguardian.com/world/2018/jan/05/tough-new-german-law-puts-tech-firms-and-free-speech-in-spotlight

[454] Schulz, W. (2018), "Regulating Intermediaries to Protect Privacy Online – The Case of the German NetzDG", Albers, M. and Sarlet, I. (eds.), *Personality and Data Protection Rights on the Internet*, Forthcoming, p.9.

Table 10: Reported numbers by selected platforms, January - June 2018

	No. of reported items	Removed	Removed within 24 hs
Facebook	1,704	21.2 %	76.4 %
YouTube	241,827	27.1 %	93.0 %
Twitter	264,818	10.8 %	97.9 %
Google+	2,769	26.1 %	93.8 %

Source: Media Policy Project Blog.[455]

The table shows that when platforms removed content, it was removed in most cases within 24 hours. This could be a sign that in fact the reported content was manifestly illegal beyond doubt, but it can just as well signal over-censorship. One thing is sure: very little is known about how exactly the decisions are taken, what types of content are reported, and which of them are removed and why.

The German branch of the *Reporters Sans Frontières* (*Reporter ohne Grenzen*, ROG), called for the creation of an independent supervisory body to oversee the deletion practices of platform providers. This body should compose the representatives of various stakeholders, including users and NGOs. Such supervisory bodies could develop guidelines for dealing reported content, communicate to the public, and function as an appeal body. The body recommended by ROG appears similar to a Press Council, organised on a self-regulatory basis and comprising various stakeholders.

Similar to discussions around the exchange of electronic evidence with the US CLOUD Act,[456] the German example **shows a new tendency to shift legal analysis to private service providers**. In this legislation, they are the entity in charge of **assessing the 'legality' and truthfulness of online messages**. On the positive side, the regulated procedure and the reporting obligations of the platform providers are a step forward towards transparency and accountability for their actions, and can provide a starting ground for further research as well.

A similar approach was adopted in **Italy** ahead of its recent national parliamentary elections. In February 2017, a draft law was introduced to the Italian Parliament with the declared purpose of countering 'Fake News'. The law would criminalise the posting or sharing of "false, exaggerated or tendentious news", imposing fines of up to EUR 5 000 on those responsible. In addition, the law proposed imprisonment for the most serious forms of fake news such as those that might incite crime or violence, and imposed an obligation on social media platforms to monitor their services for such news. Moreover, the Italian government has created an online portal where people can **report hoaxes**. The portal prompts users to supply their email address, a link to the misinformation they are reporting and any social network they found it on. The requests are conveyed to authorities at the *Polizia Postale*, a unit of the state police that investigates cybercrime, who will fact-check them and, if laws were broken, pursue legal action. In cases where no laws were broken, the service will still draw upon official sources to deny false or misleading information.

French President Emmanuel Macron is the latest political leader to hop on the anti-fake news bandwagon. On 3 January 2018, his parliamentary majority proposed a law – a so-called "emergency legal action". The set of two new proposals on disinformation (*Proposition de loi relative à la lutte contre la manipulation de l'information*) aimed at tackling the dissemination of false rumours during electoral campaigns. The text was formally adopted on the 20 November 2018, after 10 months of vibrant debates. Before being endorsed, it went back and forth twice from the Parliament to the Senate between June and November,[457] and was rejected twice by the high

[455] http://blogs.lse.ac.uk/mediapolicyproject/2018/08/16/removals-of-online-hate-speech-in-numbers/

[456] In judicial cooperation, the CLOUD Act aims at improving cross-border electronic data request by granting the service providers the prerogative to assess the proportionality, necessity and legality of a foreign law enforcement request to access online content stored in the US.

[457] See: http://www.assemblee-nationale.fr/15/ta/ta0180.asp

chamber.[458] The first time the text entered the Senate, the members rejected it without even discussing it in detail. The Law Commission Rapporteur explained this rejection stating that the chamber had doubts on the efficiency of these proposals and feared possible **further risks to the freedom of communication**.[459] A group of senators and MPs and the Prime Minister then consulted the **Constitutional Council** on the text. The Council reviewed the proposals and determined they were in conformity with the Constitution. It declared that "it is the prerogative of the legislators to reconcile the constitutional principle of the fairness of the ballot with the constitutional freedom of expression and communication".[460] Despite these institutional debates, the law is expected to be operational for the European Parliament elections of May 2019.

This bill deals specifically with disinformation circulated on the internet. It empowers users of **social media platforms**, such as Facebook and Twitter to notify disinformation, as it is currently the case for nudity, violence, harassment, suicide or self-harm, forbidden sales, terrorism and hate speech. The bill also allows the Superior Council of Audiovisual Media (*Conseil Supérieur de l'Audiovisuel*)[461] to suspend the broadcasting of programmes by a foreign broadcast company. This power was also extended to the internet, despite the Council's prerogative being limited until then to broadcasting.

The judicial authorities are granted the power to delete content within 48 hours as part of an emergency procedure, requiring judges to decide on the **truthfulness** of an information and the **intent** of the author to manipulate public opinion, all within in a very short timeframe. The inherent risk of **judicial overload and issues related to the definition of disinformation** have been outlined by many commentators and were also present in the **numerous amendments** to the proposals discussed during the summer in the Parliament. In addition, calling on an interim relief judge to act — the fastest form of serving justice — will always be too little too late. Evidence suggests it could even backfire: qualifying a piece of news as fake and thereby giving it greater publicity gives the news piece a boost and spreads its reach even further.

Under this course of action, the legislator and ultimately the courts can either decide what constitutes fake news, or outsource this responsibility immediately to social media. Additionally, it seems unrealistic to expect social media platforms based overseas, and mainly in the US, to implement French judicial decisions as there is no bilateral framework of cooperation in this matter. As a result, companies like Facebook Germany have hired – and continue to do so – more human curators and partnered with fact-check organisations in an attempt to keep misinformation out of people's feeds. As regards the French solution, there seems to be a clear risk that an incumbent government constrains the freedom of expression of its opponents, be they citizens writing on their blogs or accredited journalists writing for major publications. Moreover, both systems contain one major flaw: when fake news stories do get denounced as potentially false, or the interim judge is ready to take action, it is already too late and the story has gone viral.

The most interesting point of criticism was that the effective press law of France (the law of 29 July 1881 on the freedom of the press) already stated that "the publication, broadcasting or reproduction via any mean of false information is punishable with a EUR 45 000 fine if they are likely to disturb the peace" (article 27).[462] Various commentators therefore pointed out the uselessness and redundancy of this new law and its possible negative effects on freedom of the press and freedom of speech. The new law's aim is to tackle disinformation on the

[458] See: https://www.legifrance.gouv.fr/affichLoiPreparation.do;jsessionid=C5AC40849E4538F34F1D1D56A61A7A6C.tplgfr37s_3?idDocument=JORFDOLE000037151987&type=general&typeLoi=prop&legislature=15

[459] Rapporteur Christophe-André Frassa, see: https://www.francetvinfo.fr/internet/reseaux-sociaux/facebook/le-senat-rejette-les-propositions-de-loi-sur-les-fake-news-sans-meme-discuter-du-texte_2869295.html. The general discussions are available, in French, at: http://www.senat.fr/cra/s20180726/s20180726_4.html#par_437

[460] Decision n° 2018-774 DC, available in French at: https://www.conseil-constitutionnel.fr/actualites/communique/decision-n-2018-774-dc-du-20-decembre-2018-communique-de-presse

[461] See: https://www.csa.fr/

[462] See: https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006070722

internet specifically, although the 1881 law already provides a framework related to disinformation published "via any means" and is not limited to electoral campaigns. However, the main argument for the new law is that during election campaigns, disinformation could be removed swiftly (within 48 hours) but, by a judicial decision, rather than by the platform provider – thereby respecting the rule of law.

The French Criminal Code also defines the dissemination of false information as an offence in its article 322-14. This is not limited to any particular type of media, and it is more similar **to defamation and scaremongering**. Both are punishable by a sentence of two years imprisonment and a EUR 30 000 fine.[463]

To tackle disinformation during electoral campaigns, the Electoral Code already defines the practice as "any information or accusation on a fact deprived of verified elements in order to make it credible" (article 97). Since the introduction of the law for trust in the digital economy of 21 June 2004 (*Loi sur la confiance dans l'économie numérique*), there is a possibility for judicial authorities to **delete online content if they are related to violent statements, hate speech and discrimination** (article 6).[464]

The jurisprudence of the Court of Cassation and the Court of Appeal defined the concepts of "fake news and trouble to public order" and already ruled on these cases on different occasions. Disinformation (*fausses nouvelles*) is defined as a "specific and circumstantial fact" (*CA Paris, 11e ch., sect. A, 18 mai 1988*). Additionally, "the fact should be false, meaning misleading, erroneous or inaccurate in the materiality of the fact and the circumstances" (*CA Paris, 11e ch., 7 janv. 1998*). As for disturbing the peace, or public order, the Court of Cassation decided that the disinformation does not have to have actually disturbed the peace, but only that it was likely to do so (*CA., 26 juin 1968*).

In this already established framework, the 2018 law took a step further in considering the urgency of deleting online disinformation content during elections, while leaving aside similar problems related to other types of manipulative information online, such as hate speech and violent content, on the basis that disinformation is more urgent and involves a greater risk of disturbing the peace. The Government acknowledged in the proposals that the current legal apparatus is able to condemn authors of disinformation, but it relied on "recent developments" (*actualité récente*) to justify the need for a new law **aiming at rapid deletion** of online disinformation content.[465] The bill took the viewpoint that the matter of disinformation is primarily transmitted on the internet and by foreign broadcasting channels and therefore only applied to online platforms, leaving out newspapers, national broadcasting channels and radio stations. Authors pointed out that despite the fact that some media have developed "fact checking" tools, they can still be circulating disinformation themselves.[466]

In Spain, the former Government of the Popular Party proposed a new legislation to counter disinformation in December 2017 (*Ley relativa al impulso de las medidas necesarias para garantizar la veracidad de las informaciones que circulan por servicios conectados a Internet y evitar injerencias que pongan en peligro la estabilidad institucional en España*).[467] Whilst the rationale is similar to what was proposed in France, Spain focuses more on the "threat to the institutional stability of the country".[468] The proposal aimed at finding mechanisms to "seal" the truthfulness of online content, increasing the prerogatives of law enforcement authorities and strengthening

[463] See: http://www.codes-et-lois.fr/code-penal/article-322-14

[464] See: https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000801164

[465] Argumentation available at: http://www.assemblee-nationale.fr/15/propositions/pion0799.asp

[466] See for instance: Bigot, L. (2018), "Rétablir la vérité via le fact-checking : l'ambivalence des médias face aux fausses informations", *Le Temps des médias*, n° 30.

[467] See: http://www.gppopular.es/wp-content/uploads/2017/12/171219-PNL-Noticias-falsas.pdf?_ga=2.226110172.738521628.1541868032-2076439625.1541868032

[468] Ibid, p.1.

international collaboration in the matter. The Spanish Parliament rejected the proposals on the 21 March 2018,[469] but the topic has stayed on the political agenda.

When it comes to assessing the effect of third-party liability regimes as tested in Germany, France and Italy, one has to consider the tension existing between Article 14 of the e-Commerce Directive[470] – which entrusts law enforcement agencies (and not service providers) to enforce criminal laws and these newly-emerging approaches (see more above in subsection 3.1.2.). In the meantime, pushed by the public salience of the issue and the regulatory efforts threats by governments, **tech companies have hired thousands of employees** — and invested in new machine-learning tools — to spot and remove racism, anti-Semitism and other forms of hate speech from their services. Yet the **scale and nature of these efforts remain largely voluntary and therefore difficult to scrutinise not only by governments but also by users and journalists**. Between 1 January and 31 March 2018, Facebook reported it had taken action against roughly 2.5 million posts, photos and other kinds of content for violating its rules against hate speech.[471]

The above-mentioned examples outline fundamental questions related to the freedom of speech and press and potential rule of law issues regarding the understanding of the concept of disinformation:

- Are the new powers given to public entities, whether or not institutionally linked to Governments, a threat to freedom of speech? And are these entities the best placed to make a judgement on the truthfulness of an information? Are judicial authorities best-placed for urgent removals of content published online? Are platform providers best-placed to legally assess the truthfulness of online content?
- Is it really the disinformation, or the manipulative dissemination techniques that cause the harm to human rights and democracy?
- What are the impact and risks of such measures on online press platforms?
- What could be the potential negative effects of a misuse of these new legislations? What is the real impact on freedom of press and free speech?

From this brief analysis of these Member State case studies, it seems these questions **have not been sufficiently discussed** for the preparation of such legislation**.**

Online disinformation is a complex phenomenon that regulators have yet to really master. Therefore, if it appears too soon to create regulation that can be effective, there is something that can be done: co- and self-regulatory approaches seem worth experimenting.

4.2 Co-regulation and state-regulation

4.2.1 Co-regulation

The first, pioneering illustration of co-regulation in the disinformation space is offered by the partnership announced by France and Facebook on 12 November 2018 at the Internet Governance Forum organised by UNESCO in Paris. Under a six-month partnership, French authorities will gain access to and monitor Facebook's policies and tools for stopping posts and photos that attack people on the basis of race, ethnicity, religion, sexuality or gender. The genesis of this partnership builds on the previously discussed French so-called anti-fake news law.

[469] See: http://www.congreso.es/portal/page/portal/Congreso/Congreso/Iniciativas?_piref73_2148295_73_1335437_1335437.next_page=/wc/servidorCGI&CMD=VERLST&BASE=IW12&FMT=INITXDSS.fmt&DOCS=1-1&DOCORDER=FIFO&OPDEF=ADJ&QUERY=%28162%2F000550*.NDOC.%29

[470] 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')

[471] https://transparency.facebook.com/community-standards-enforcement#hate-speech

From this unprecedented experience of co-regulation between public authorities and a social media platform, advisors to the French President hope to determine "the necessary regulatory and legislative developments" to fight online hate speech. Moreover, on 12 November 2018, also at the Internet Governance Forum, President Emmanuel Macron launched the Paris Call for Trust and Security in Cyberspace.[472] This high-level declaration on developing common principles for securing cyberspace has already received the backing of many states, as well as private companies and civil society organisations. These high-level principles – which have not been signed by the US[473] – call for greater trust and safety online, including countering malicious actors who are trying to undermine "electoral processes through malicious cyber activities".

At the time of writing, the terms of the partnership between the French government and Facebook remain unknown as well as its potential impact. It is worth highlighting the unprecedented nature of this form of co-operation between a social media platform and a public authority in addressing a major societal, informational and economic challenge.

4.2.2 Self-regulatory mechanisms (market self-correction)

The most prominent approach, favoured by the industry, to counter disinformation has been – and continues to be – self-regulation. Given the inherent dominance of social network providers, they would be better placed to act and to address the challenge posed by disinformation – provided that they are motivated to do so. The advertising-run (pay-as-you-go) business model, in which advertisers are only charged when a page is viewed or clicked on, ensures that social media companies have no incentive to play the role of arbiters of truth. It is against this backdrop that the industry has been embracing several voluntary schemes over the last couple of years, well before public authorities stepped in. It is yet to be seen how consequential adherence to the agreed principles will be.

4.2.2.1 Pure self-regulation

Facebook, for example, is testing an innovative approach whereby it alters the environment in which a disputed, or outright fake, story is presented, rather than removing it entirely from the site.[474] It now features "Related Articles" beneath the story in question and invites readers to access additional information, including pieces that have been greenlighted by third-party fact checkers. "Related Articles" are designed to give more context, which research has shown is a more effective way to help people get to the facts. This boils down to an empirical question whether exposure to alternative viewpoints plays a role in combatting (or reinforcing) misperceptions. Indeed, Facebook has found that when it shows "Related Articles" next to a false news story, it leads to fewer shares than when the "Disputed Flag" is shown.[475] **Academic research[476] confirms this conclusion, suggesting that this design-centred approach could make a real difference in readers' perceptions**.

Unlike the prescriptive approach pioneered by Germany and now embraced by France through the anti-fake news law, providing links to related articles does not necessarily imply any editorial judgment about their truthfulness. But it does push readers to encounter facts and other points of view more serendipitously, in a way that mirrors the disparity of views in real life. It also invites the reader to form their own opinion. To be sure, this raises the question of algorithmic accountability — exactly how are those related articles and alternative views

[472] Full Text of the Paris Call for Trust and Security in Cyberspace https://www.diplomatie.gouv.fr/IMG/pdf/paris_call_text_-_en_cle06f918.pdf

[473] List of Supporters of the Paris Call for Trust and Security in Cyberspace https://www.diplomatie.gouv.fr/IMG/pdf/soutien_appel_paris_cle8e5e31-2.pdf

[474] To know more, https://newsroom.fb.com/news/2017/12/news-feed-fyi-updates-in-our-fight-against-misinformation/

[475] See supra.

[476] Leticia Bode and Emily K. Vrada, "In Related News, That Was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media" 4(65) Journal of Communication (2015), pp. 619-638.

chosen? But it is a worthy experiment. New research suggests exposure to alternative viewpoints has a tangible effect on readers.[477]

The emergence of this feature underlines the ability of a platform such as Facebook or Twitter to engage seriously with the thorny problem of disinformation. It also suggests their readiness to set aside — at least for a while — an obsessive business model based on increasing user engagement and monetising their data. The implementation of such an approach **across social networks** would set an important precedent. It could help close the gap between what is best for users and the dominant advertising business model.

4.2.2.2 Induced self-regulation

The first and only example of induced self-regulation is the one triggered by the EU. In April 2017, European Commission Vice-President Andrus Ansip, in charge of the completion of the Digital Single Market, qualified fake news as a major threat to European democracies. At the same time, he highlighted the need to protect freedom of speech and trust people's common sense. As a result, in its 2018 Communication "Tackling online disinformation", the Commission put forward an action plan and self-regulatory tools to tackle the spread and impact of online disinformation in Europe and ensure the protection of European values and democratic systems. One of the actions was to convene a Multistakeholder Forum on disinformation that was tasked with developing a Code of Practice to tackle online disinformation. The forum saw the participation of representatives of online platforms, leading social networks, advertisers and the advertising industry, who agreed on a self-regulatory Code of Practice to address the spread of online disinformation and fake news. Four principles guided its action:

1. Improve transparency regarding the way information is produced or sponsored;
2. Diversity of information;
3. Credibility of information;
4. Inclusive solutions with broad stakeholder involvement.

This Code of Practice, presented on 26 September 2018, marks the first time worldwide that the industry agreed, on a voluntary basis, to self-regulatory standards for combatting disinformation. The Code pursues the objectives set out in the 2018 Communication "Tackling online disinformation" by setting a wide range of commitments, from transparency in political advertising to the closure of fake accounts and demonetisation of purveyors of disinformation.

In an annex, the Code also identifies best practices that signatories pledged to apply to implement the Code's commitments. Signatories include some of the largest tech companies, such as Facebook, Google, Twitter, and Mozilla.

The Code of Practice divides the commitments into six sections:

1. Better scrutiny of advert placements and avoiding the promotion of websites or adverts that spread disinformation;
2. Ensuring that political advertising and issue-based advertising is clearly distinguished from editorial content and news and to improve transparency on its sponsors;
3. Tackling fake accounts and improving transparency around the use of bots;
4. Empowering consumers by making it easier to find trustworthy and diverse sources of news;
5. Empowering the research community by encouraging independent efforts to track disinformation and to support research into disinformation and political advertising.

[477] See e.g. Carnahan, Dustin, A Deliberative Mindset? Considering the Role of Motivation in Assessing the Attitudinal Consequences of Selective Exposure (2014). APSA 2014 Annual Meeting Paper. Available at SSRN: https://ssrn.com/abstract=2453390

6. Progress made against the commitments will be evaluated through annual reports published by the Code's signatories and reviewed by a third-party organisation.

As originally foreseen in the Communication, the Commission was supposed to follow the progress made closely and analyse the first results of the Code of Practice by the end of 2018. Should the results prove unsatisfactory, the Commission may propose further actions, including actions of a regulatory nature.

The Code set forth by the Commission offers the backbone of the EU approach in ensuring transparent, fair and trustworthy online campaign activities ahead of the European elections in spring 2019 – if possible, by self-regulation. By restricting the typology of content being addressed to that which is objectively false or misleading, and only where it is done for financial gain or would threaten a legitimate public interest such as security, public health or democratic processes, the definition appears to **be consistent with international standards on where particular forms of expression can be permissibly restricted**, such as those set out in Article 19(2) of the International Covenant on Civil and Political Rights.

Yet it has been criticised insofar as **there is nothing in the Code which would enable a producer or publisher to be made aware that their content had been identified as "disinformation**" and thus deprioritised or labelled in a particular way. Nor has any provision been made for users to appeal against such decisions, and obtain remedies where appropriate. Without these important due process guarantees, there is little transparency and accountability, as in the existing operation of the EU *Disinformation Review*.

The first reports on the implementation of the Code of Practice were submitted to the Commission in December 2018.[478] Some meaningful measures have been applied, for example, by Facebook, such as the content-agnostic approach to aggressive dissemination techniques, especially coordinated inauthentic behaviour (CIB). Further, Facebook reports applying a strict registration policy by disabling the possibility to create multiple accounts and to create inauthentic accounts, including creation by bots.[479] On the other hand, no mention is made of data protection and micro-targeting, which are regarded as key cornerstones of the fight against online manipulation.

4.2.2.3 The Action Plan Against Disinformation ahead of the European Parliament elections

In light of the upcoming European Parliament elections, the European Commission published its Action Plan Against Disinformation in December 2018, aimed at protecting the EU's "democratic systems and public debates". Following the 2015 European Council Decision to "challenge Russia's ongoing disinformation campaign" and the 2016 Joint Framework on countering hybrid threats, the Action Plan targets foreign agents who might be suspected of increasingly "deploying disinformation strategies to influence societal debates, create divisions and interfere in democratic decision-making".[480]

The European External Action Service has set up specific strategic communication task forces consisting of experts with relevant language and knowledge skills, to address the issue and develop response strategies in order to implement such an action plan.

The coordinated response to disinformation presented in this Action Plan is based on four pillars:

(i) improving the capabilities of Union institutions to detect, analyse and expose disinformation;

(ii) strengthening coordinated and joint responses to disinformation;

[478] Commission press release: Code of Practice against disinformation: Commission calls on signatories to intensify their efforts. Brussels, 29 January 2019. http://europa.eu/rapid/press-release_IP-19-746_en.htm

[479] Facebook Baseline Report on Implementation of the Code of Practice on Disinformation. Published on 29. Jan. 2019. http://ec.europa.eu/information_society/newsroom/image/document/2019-5/facebook_baseline_report_on_implementation_of_the_code_of_practice_on_disinformation_CF161D11-9A54-3E27-65D58168CAC40050_56991.pdf

[480] European Council conclusions, 18 October 2018

(iii) mobilising the private sector to tackle disinformation;

(iv) raising awareness and improving societal resilience.

The Action Plan has narrowed the definition of disinformation, and is concrete and practical. It has the potential of helping Members States operationalise the recommendations put forward in the Code and agreed by online platforms and the advertising industry. As such, its impact could go well beyond the next European elections in May 2019. However, it focuses primarily on the Eastern Neighbourhood and less on EU Member States themselves. The coordinated and joint responses, for example the Rapid Alert System can be understood in the context of a joint enemy. However, **the issue of division within the European Union between Member States** is not addressed in the Action Plan (for example, when the source of disinformation is in one Member State, targeting either its own society or another society).

4.2.3 Lessons learned

Given the public salience and unprecedented severity of the disinformation phenomenon, to fight fake news by law may seem an attractive option. **Yet limiting news output to "true" — essentially state-sanctioned — information could pose an even greater threat to democracy than disinformation itself**. Moreover, evidence suggests that to counter false information by law could even backfire: qualifying a piece of news as fake and thereby giving it greater publicity gives the news piece a boost and spreads its reach even further. Disinformation and propaganda are symptoms of deeper structural problems in our societies and media environments. To counter it, we need to take a step back so as **to examine the vulnerabilities these fake news narratives exploit.** In particular, we must unpack the underlying, self-reinforcing mechanisms that make this old phenomenon so pervasive today.

Chapter 2 analyses the structure of the new public sphere: algorithmic settings optimised for advertisers, monetising of personal data – these make social media platforms interested in maintaining the current business model. For this to change, a transformation of the social norm governing how content is presented and consumed online must occur. This can only be achieved through a sudden, individual and collective awakening – realising the pervasive effects of today's dominant business model governing how we inform ourselves and how critically we consume content online.

The jury is still out on how anti-disinformation measures can and should operate. More experimentation is needed and the French/Facebook partnership is an initiative worth studying.

4.3 Exploring existing national and European rules relating to election campaigns – how far they could be applied to counter disinformation and propaganda?

4.3.1 Rules that could counter disinformation and propaganda

Election procedures, campaign regulation and their supervision methods are crucial instruments for ensuring fair elections. Provisions for countering fake news, propaganda and disinformation and regulations concerning the use of campaign finances in electoral codes and election acts vary from Member State to Member State in the European Union. Interestingly, certain regulations aimed at more transparent campaign financing and political advertising can be found mainly (but not exclusively) in the Central Eastern European Member States that joined the Union in 2004 and 2007 (Bulgaria, Hungary, Latvia, Poland, Romania, Spain, Sweden).

Transparency in campaign financing is a vital goal in almost all election laws in the 15 Member States. Strict rules in financing advertising activities could also lead to better control on online media. The terms 'propaganda' and 'campaign' are used almost as synonyms in several electoral laws (for example in the Spanish electoral law). Propaganda has a more negative connotation while campaign tends to be more neutral.

Campaign silence is mentioned only in a few election laws, and even in those only implicitly. Usually "24 hours before the voting day" (Poland) or "on the day of the election or on the day before it" (Portugal) campaign activities are banned. But, for example, electoral campaigning is prohibited in Spain "once the campaign is legally finished", which is quite a vague reference to campaign silence. In 2018, Hungary removed the previously 48-hour campaign silence regulation, and retained it in a limited form: only on ballot day and only for audiovisual political advertising.

Several campaign regulations concerning political ads or campaign silence apply to **offline media only** (radio, television, printed materials like posters and leaflets), and therefore their potential to impede the negative processes of disinformation and propaganda during elections is limited. However, there are Member States where campaign silence has been extended to **online media as well**. For instance, in Spain, prohibition of electoral propaganda includes online media.[481]

4.3.2 Notable examples

There are three noteworthy examples in the Member States examined.

- In Poland, the election act refers to "information that is untrue", its advertisers van be obliged to pay up to 100 000 zlotys (EUR 23 000) to an organisation of public benefit. Also required is that the disinformation shall be corrected, replied to or apologised for at the latest within 48 hours.
- In Spain, there is a competent central authority that is responsible for the allotment of free campaigning spaces provided by local governments. This body is the Central Electoral Commission, which also has a Radio and Television Committee.
- In Portugal, there are considerably strict legal rules concerning election campaigns. Campaign expenses per candidate cannot be more than fifteen times the monthly national minimum wage. Also, the campaign period is remarkably short compared to that in other Member States, as it does not exceed two weeks. Article 58 contains an important safeguard of freedom of speech and expression when it claims that "during election campaigns no limitation may be imposed on the expression of political, economic and social principles, without prejudice to any civil or criminal liability" – however, this could prevent measures that would counter disinformation.

Disinformation and propaganda as such are not explicitly mentioned in the above election laws or electoral codes, however, in many cases there are implicit references to the fight for fair and equal campaign activities and transparent expenses that do address the topic of this present research. Also, it is vital that in many Member States, free and fair access to media for campaigning political parties is ensured by law.

4.3.3 Latest legislative initiatives in countering disinformation

When selecting the states to be examined for this chapter, special attention was given to those states which held elections in the past few years. Several of them also attempted to counter disinformation by legislative initiatives, albeit most remained unsuccessful. The most explicit and targeted initiative took place in the United States in 2017, where three senators introduced the Honest Ads Act. The main aim was to "help prevent foreign interference in future elections and improve the transparency of online political advertisements".[482] The bill directly targeted Facebook, Google and Twitter after the Russian interference in the US presidential elections in 2016.

In the summer of 2018, some EU Member States initiated amendments in their election act and these amendments have already been passed by their legislative bodies. In June 2018, the Polish Lower House repealed an amendment passed in January the same year concerning the Electoral Code, and which obliged polling

[481] Cappello M. (ed.), Media coverage of elections: the legal framework in Europe, IRIS Special, European Audiovisual Observatory, Strasbourg, 2017, p 114

[482] Warner, Mark R. (2017): The Honest Ads Act, Source: https://www.warner.senate.gov/public/index.cfm/the-honest-ads-act

stations to record and transmit the work of the electoral committees. The purpose of this was to "increase the transparency of elections", but it was found to be in conflict with the GDPR and the amendment had to be repealed – stirring political controversies.[483] This illustrates how difficult it may be to find the right balance between the interests of accountability and human rights, in this case between transparency and privacy.
The Romanian Senate also adopted a change in the electoral code in the summer of 2018. The changes concern the electoral campaign for presidential and European Parliamentary elections and extends the scope of political campaign to schools, and also to the day of voting.[484]

Not only did things change in the CEE region: the French Parliament put the topic of restricting disinformation on the table in 2018. This is explained in more detail in subsection 4.2.3.[485]

A specific example is that of Sweden, which, rather than reacting subsequently to the effects of fake news, adopted a proactive approach before the general elections in September 2018 and set up an authority in January to counter disinformation and foreign influence campaigns.[486] The authority's main purpose was to **strengthen resilience against disinformation, and provide "psychological defence" (*psykologiskt försvar*) for the population**.[487] The Swedish Civil Contingencies Agency has worked with the Swedish Election Authority, the security police and the national police to tackle foreign interference in the 2018 election. It is noteworthy that the authority is responsible (among others) for identifying, analysing and confronting influencing operations.[488]

As illustrated in subsection 3.2.3 above, existing legal rules cannot be fully applied to counter disinformation and propaganda – the same applies to elections rules. In the above-mentioned examples (Spain, Poland, Portugal), strict rules may help authorities to prevent and fight 'untrue information', but these endeavours can restrict the freedom of thought and speech.

The ultimate dilemma is to find the right balance between freedom of speech and the fight against disinformation and propaganda. This should also be interpreted in light of international agreements on fundamental rights.[489]

Some practices for countering disinformation in the selected 15 EU Member States are discussed in section 4.3.

Various actions initiated or finalised by Member States to tackle disinformation (a report in the United Kingdom, laws in Ireland, France, Russia and Croatia, a media literacy campaign in Belgium, law enforcement measures in Germany and Italy, an investigation in Turkey, and task forces in Sweden, Denmark and Spain)[490] could possibly restrict freedom of speech and the press. These actions are extremely heterogeneous and do not necessarily concern the electoral system and the functioning of the rule of law.

[483] Poland in English (2018): Parliament votes through controversial electoral code amendment (15.06.2018), Source: https://polandinenglish.info/37663489/parliament-votes-through-controversial-electoral-code-amendment

[484] Marica, Irina (2018): Romanian Senate adopts electoral law changes, in: Romania-Insider, Source: https://www.romania-insider.com/senate-adopts-electoral-law-changes/

[485] Fraser, Matthew (2018): The legal fight against 'fake news' must not veer into censorship, in: The Conversation, Source: http://theconversation.com/debate-the-legal-fight-against-fake-news-must-not-veer-into-censorship-98049

[486] Funke (2018)

[487] The Local (2018): Sweden to create new authority tasked with countering disinformation, Source: https://www.thelocal.se/20180115/sweden-to-create-new-authority-tasked-with-countering-disinformation

[488] The Local (2018)

[489] Renda, Andrea (2018): The legal framework to address "fake news": possible policy actions at the EU level, Policy Department for Economic, Scientific and Quality of Life Policies (CEPS - Centre for European Policy Studies and College of Europe) Directorate-General for Internal Policies PE 619.013-June 2018, p. 20.

[490] Funke, Daniel (2018): A guide to anti-misinformation actions around the world, in: Poynter, Source: https://www.poynter.org/news/guide-anti-misinformation-actions-around-world

4.3.4 The Communication on securing free and fair European elections

The Commission's Election Package[491] notes that Member States are responsible for the organisation and the monitoring of the election process. It also notes that citizens should be able to discern who is speaking to them online through advertising and political messages, as well as who is paying for these, with other recommendations for platform providers with the aim of minimising disinformation accessible through their platforms. The Recommendation encourages Member States to establish and support a national elections network, and for them to cooperate with relevant other authorities, and that they appoint contact points to take part in a European cooperation network for the European Parliament (EP) elections, with the view to serve as a real-time European alert process and exchange of information and practices. All recommendations in the Election Package are commendable, but they **might be insufficient**. Respecting the national competences of Member States, **cooperation alone may not ensure the required effect**, especially if mutual trust between Member States, and the common understanding of the rule of law shows signs of cracks.

4.3.5 Possibility of monitoring campaign expenditures during the EP elections

Before the EP elections some of the national parties organise and undertake campaign and propaganda activities in Member States. There are national authorities[492] in every Member State that are responsible for keeping a transparent record of all campaign expenditures from every participating political party. These records are open to the public and in this way the EP can also monitor campaign financing.

However, the current frames of campaign financing supervision suffer from substantial limitations. First, the data neither indicate the nature of the expenditure, nor the contracting partner. For example, political advertising as a budget item does not necessarily refer to disinformation activities. Second, in most countries the agencies and controlling institutions receive this information from the political movements, parties or candidates themselves. Third, most of this information and the financial reports only become available some months after the elections.

Therefore, these instruments cannot guarantee a control mechanism on expenditures concerning disinformation, and are not sufficient for monitoring campaign finances during the EP election campaigns. In the event of a breach of any party financing regulations, several European countries have financial, criminal or administrative legal sanctions. Criminal sanctions are very rarely applied, although they would be available in Estonia, Belgium, Cyprus, Estonia, Finland, France, Greece, Slovakia, Denmark and the United Kingdom.

A new amendment of the Regulation on the statute and funding of European political parties and European political foundations was passed by the European Parliament, and the European Council[493] states in Recital (30a) that the "European Public Prosecutor's Office (EPPO) has the task of investigating alleged criminal offences in the context of the funding of European political parties and European political foundations which affect the financial interests of the Union." This competence may provide the EU with a better tool to monitor the expenses of political parties – mainly in those Member States that participate in the European Public Prosecutor's Office.

[491] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Securing free and fair European elections. Brussels, 12.9.2018. COM (2018) 637 final.

[492] Directorate General for Internal Policies: Party financing and referendum campaigns in EU Member States, European Union, Brussels, 2015, pp. 39-44. Source: http://www.europarl.europa.eu/committees/en/supporting-analyses-search.html

[493] Regulation (EU, Euratom) 2018/673 of the European Parliament and of the Council of 3 May 2018 amending Regulation (EU, Euratom) No 1141/2014 on the statute and funding of European political parties and European political foundations. Source: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0673&from=EN

4.4 Promising practices implemented during recent national or local elections, or referenda

4.4.1 Introduction

Parts of disinformation and propaganda campaigns aim directly at influencing national elections or referenda (see section 1.3). In recent years, almost all of the elections in EU Member States (elections in the Netherlands in 2017; the French presidential election in 2017; the Swedish general elections in 2018;[494] the Hungarian parliamentary elections in 2018)[495] showed signs of foreign intervention to a greater or lesser extent.

The number of recommendations is fairly high and there are some key points highlighted by numerous experts, but the number of promising examples is considerably lower. This section will summarise these latest practices from those states that had elections in the past few years.

4.4.2 Media and information literacy

Education in critical thinking, development of a critical perception of reality and the ideal of the well-informed citizen are key elements of resilience against fake news and disinformation. Making society more sensitive to disinformation and alternative facts was key before the elections in the **Netherlands, in France or in Sweden.** Raising awareness of the possibility and possible consequences of interference in the election process was a crucial step in these societies. Education of not just youth but the elderly and the public at large is an essential way of strengthening resilience among citizens and voters.

According to the Special Eurobarometer on democracy and elections,[496] when respondents were asked about their concerns about the use of the internet in the pre-election period during even local, national or European elections, 73 % of the internet users were concerned about disinformation or misinformation online, which is a clear indication that most internet users have already heard about disinformation phenomena in Europe. However, a vast majority of them, 58 % of the total sample, agreed their country is doing what is needed to prevent illegal and fraudulent activities during elections, whether at local, national or European level.

The new proposed text of the AVMS Directive includes a definition of ‘media literacy’ as one which should “aim to equip citizens with the critical thinking skills required to exercise judgment, analyse complex realities and recognise the difference between opinion and fact”. Article 28.b.2.j requires Member States to provide for effective media literacy measures and tools, and raise user awareness of them, and Articles 30b and 33 reiterate this goal with similar content.

Media literacy programmes should focus on how citizens can distinguish between real and fake news, how they can reduce the effect of disinformation campaigns and conspiracy theories and how they can be more critical and doubting about propaganda and disinformation. One of these examples is from the Nordic countries, where as **part of the official school education** in Sweden, lessons, human and material resources are dedicated to the education of future voters and citizens by means of developing their critical thinking, and critical perception of propaganda and disinformation. **Gamification** could also be used as a possible tool to raise awareness and literacy skills to combat disinformation.[497]

[494] News and Political Information Consumption in Sweden: Mapping the 2018 Swedish General Election on Twitter? https://comprop.oii.ox.ac.uk/research/sweden-election/?

[495] Political communication in the Hungarian election campaign? https://univiennamedialab.wordpress.com/2018/04/04/political-communication-in-the-hungarian-election-campaign/?

[496] See more: Special Eurobarometer 477: Democracy and elections. http://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/surveyKy/2198

[497] For example, see the NATO Stratcom's game to differentiate between real information and falsehood: https://www.stratcomcoe.org/facebook-game-teaches-how-spot-disinformation;

and the game built by Cambridge University to build your own fake news thread: https://getbadnews.com/#intro

We can find a number of other countries where media and information literacy programmes are part of the **official curriculum** (e.g. Italy), but there are also some other countries where grassroots movements and civil society organisations are developing such training programmes (e.g. CEE countries). As the AVMS Directive obliges Member States to promote and take measures for the development of media literacy skills (Article 33), this should not only remain at the level of civil society.

4.4.3 The crisis of journalism and how to overcome it: building a stronger future for journalists

A long-term harmful effect of the disinformation crisis is how it has undermined trust in the press. The already dominant relativism has grown with a feeling that "everyone is lying". This could also be a strategic objective, to undermine confidence in democratic institutions and processes, including the media. At the same time journalism is facing a general crisis.[498] The industry is underfinanced, leading traditional press organs to close or cut down on staff, often losing the most talented and creative journalists. This general crisis is further exacerbated by pseudo and troll news makers.[499]

An assessment of current media ownership relations and the indicates that, mostly in Central and Eastern European countries, and especially in those countries where public service media broadcasters are unable to provide balanced and unbiased information (e.g. Hungary), **smaller but independent media houses and editorial offices play an increasingly important role in political and economic debates**.[500] These editors can contribute to the emergence of democratic social relations, democratic discourse and public dialogue, with credible information based on facts from verified sources, and with unbiased, impartial and high quality content. These editorial boards – some of which also operate as an NGO – **need financial support to maintain their independence from economic and political pressures** and in order to secure their operation in a long-term and balanced way. In several instances in the CEE region, these groups also take up the fight against Russian propaganda and disinformation campaigns (e.g. 444 in Hungary) with credible information and fact-checked contributions to the public sphere.

As it was noted above, media should also take responsibility to deliver trustworthy, quality and reliable information to news consumers, thereby boosting trust in traditional print, broadcast or online media. In order to reach this objective, quality content should be based on impartial and reliable information, pluralistic views and on the basic principle of promotion of democratic values, including diversity, social cohesion and cultural diversity, helping to overcome the adverse effects of the disinformation war and contributing to public trust and citizens' resistance (see more in chapter 7: Recommendations).

4.4.4 Campaign silence[501]

The legitimacy of campaign silence rules has triggered many controversies recently. In the context of an online information space that crosses national boundaries, and horizontal communication by citizens, these regulations can only have a limited effect. But this limited effect – for example, preventing at least dominant companies from disseminating game-changing disinformation with aggressive automated methods just hours before the voting – may have an effect that should not be underestimated. On the one hand, the same could occur just before the start of campaign silence, leaving no time for the political opponent to fight back. On the other hand, in the latter case, voters would still have some time left for reflection to consider the veracity of the information. Also, an election committee might lawfully supervise and allow an exceptional lifting of the campaign silence in case something extraordinary is published just before it begins.

[498] See more: Alex T. Williams: Measuring the Journalism Crisis: Developing New Approaches That Help the Public Connect to the Issue. International Journal of Communication 11(2017), Feature 4731–4743.

[499] Xymena Kurowska, Anatoly Reshetnikov: Neutrollization: Industrialized trolling as a pro-Kremlin strategy of desecuritization. Security Dialogue, 49(5), 345–363. Source: http://journals.sagepub.com/doi/10.1177/0967010618785102?

[500] See more the Soft Censorship Reports on Hungary published by Mérték Media Monitor year by year. http://mertek.eu/en/our-works/press-freedom/

[501] See more the 4.2. point of this Report.

The Eurobarometer survey on Democracy and elections found that a majority of respondents who use the internet were in favour of **introducing on online social networks the same strict silence period that is required for other media**.[502]

Member States have different provisions regarding the institution of campaign silence. Usually campaign activities are banned "on the day of the ballot" (Hungary) or "24 hours before the voting day" (Poland, France) or "on the day of the election or on the day before it" (Portugal). But, for example, in Spain electoral campaigning is prohibited "once the campaign is legally finished" – quite a vague reference to campaign silence. An interesting example is the French election rule that prevents media from quoting presidential candidates or their supporters within 24 hours of the vote, and this prevented most French voters obtaining information about the so-called "Macron Leaks".

4.4.5 Credibility index

Creating and maintaining a list of all media outlets and their so-called credibility indices can be a form of raising trust in the (trustworthy) media, increasing awareness and information literacy, and altogether combatting disinformation campaigns, in both online and offline media environments. For example ***Le Monde***[503] **published a list with hundreds of websites and their level of reliability before the French presidential elections**. A similar list with disinformation and propaganda portals was published in Hungary in 2018. One alternative could be that search engines take into account these reliability indices and rank the results accordingly.

A significant example is Microsoft News, which has partnered with more than 1 000 publishers and 3 000 brands in 140 countries in order to provide a credibility index of English-speaking media. On average, these partners put out more than 100 000 pieces of unique content per day. A further alternative is for social media platforms to cooperate with credibility indices or fact-checkers, and take this feature as a factor in algorithmic operation. Some even argue that end users should have the **option to see only trusted** (certified) news on their social media sites. It must be noted that several competitive credibility indices may be available, and that the system could be exploited in the same way as other technological innovations. Creating credibility indices through broad cooperation – e.g. with the help of journalistic associations – would raise the chances of a widely-accepted tool being created.

Regaining trust in media may have a significance from the perspective of democratic public discourse. In those EU Member States where trust in traditional media, including public service media, is higher, for example in France or Belgium, the proportion of **social media consumption is significantly lower**. In Belgium (in the Flemish as well as in the French region) respondents highlighted the significant role of traditional media outlets and their trust in them.[504] The long-standing trust in traditional media was a key element, for example, during the French presidential election in 2017.
Considering that social media platforms are the main channels of disinformation, it could be concluded that increased trust in traditional media could mitigate the harmful effects of disinformation.

[502] Proportions range from 87% in Croatia, 83% in Ireland and 80% in Greece and Hungary to 58% in Sweden, 64% in Finland and 67% in Austria. In Croatia (54%) and Denmark (52%) at least half were strongly in favour of this measure. Eurobarometer Special. 477. Democracy and elections. September 2018. at 72.

[503] Le Décodex, un premier pas vers la vérification de masse de l'information

https://www.lemonde.fr/les-decodeurs/article/2017/02/02/le-decodex-un-premier-pas-vers-la-verification-de-masse-de-l-information_5073130_4355770.html

[504] Eugénie Coche: 'Fake news' and online disinformation Case study – Belgium. Source: https://www.ivir.nl/publicaties/download/Case-study-Fake-News-Belgium.pdf

Access: 2 January 2019.

4.4.6 Fact-checking initiatives and media innovation projects

Fact-checking has become a trendy buzzword, and several initiatives emerged even beyond the United States, for example in Sweden. During the Swedish general elections campaign, four leading news outlets began a joint fact-checking initiative in order to combat disinformation (otherwise, in Sweden[505] and in the Czech Republic[506] separate government agencies are dealing with the fight against Russian propaganda). The European Union is operating the 'EU versus Disinformation' campaign run by the European External Action Service East Stratcom Task Force. Such services are very effective tools for journalists, researchers, policy-makers and every mindful media consumer. When they are created with the cooperation of various stakeholders, such as state, non-state, civil, academic and technology-relevant actors, experts and specialists, their efficiency and trustworthiness can be enhanced.

The limits of fact-checking sites are in their slowness: by the time they examine 'suspicious' content, it is likely to have rapidly multiplied and been distributed to many users. Furthermore, if troll industries or manipulating organs simply increase the mass of disinformation material, the efforts can turn into a 'Sisyphean task' and expend all investment of human and material resources in a futile effort. Creating blacklists and whitelists of websites and sources that have a tendency to provide disinformation / trustworthy information may be more practical, as is already the case. Besides providing information to interested users, this could draw the attention of potential advertisers to the quality of the site (naming and shaming).

Some of these sites have a **cooperation agreement with Facebook** that helps to avoid the further spread of completely fake news posts. Automatic algorithms can also be used to reduce the visibility of these posts. These measures and solutions can be **in line with international legal obligations on freedom of expression and the press, as it does not limit the basic background action of the act** (the free expression). Political candidates or parties can bypass traditional media outlets and use the social media page directly, as the US presidential election campaign in 2016 demonstrated: Donald Trump used his Twitter page as a main communication channel. Even though profiles are currently not usually listed among fact-checkers, this could be added to the services.

4.4.7 Involved parties / institutionalised protection

With the cooperation of a diversity of various stakeholders, better results can be achieved in the fight for fair and transparent elections. Stakeholders need to create an institutionalised way of protecting free, clean and fair elections and campaigns while safeguarding democratic institutions and processes.

Within this framework, one of the most important challenges is **to engage the widest possible range of stakeholders**. To achieve this goal, it is necessary, for example, **to break the ice between political parties** as was the case in France during the presidential election campaign: former President Francois Hollande warned the people of the threat of fake news and the role of disinformation in the campaign; and **all parties involved cooperated to overcome the dangers** (except for Le Pen's Front National Party). Campaigns must include a wide range of stakeholders (as we have also seen in France – a high level of cooperation among the state, political parties and the media), with the broadest possible geographical coverage (federal/state/ national/regional/local organs) and **an interdisciplinary approach as witnessed in Germany where ethical hackers and software engineers were involved in order to secure a pre-compliance examination** in connection with the weaknesses and vulnerabilities of the German electoral system and infrastructure. Broad cooperation may add the necessary perspective and long-term engagement that is key to a successful fight against disinformation and propaganda.

Affected stakeholders are: legislative bodies and authorities; political parties and campaign staff; election infrastructure, election software company experts; traditional media providers; social media providers; educational organisations and staff; representatives of the academic sphere and researchers.

[505] Sweden raises alarm on election meddling. https://euobserver.com/foreign/140542

[506] Czech Republic to fight 'fake news' with specialist unit. https://www.theguardian.com/media/2016/dec/28/czech-republic-to-fight-fake-news-with-specialist-unit?CMP=share_btn_tw

In sum, the promising practices applied in various states during or before elections were those that are already known and recommended by other instruments, for example by the Commission Action Plan: increasing awareness and media literacy, improving the quality of journalism, including fact-checking and credibility indices, and the cooperation of all stakeholders within society. Besides platform providers and media outlets, self-regulation by political parties and their engagement to respect ethical campaign principles would also be necessary.

5. FUTURE PROSPECTS

KEY FINDINGS

- Profiling individuals by drawing rich inferences about them will be one of the key trends in the future development and application of the new technologies.
- In the future, falsification technology will become ever more perfect, and indistinguishable from genuine information (e.g. 'deep fakes'), as well as more accessible.
- Virtual Reality, Augmented Reality and Artificial Intelligence will be more commonly used. Distributed Ledger Technologies (DLT) and Blockchain pose new challenges and opportunities for policymaking.
- Blockchain is currently being examined as to whether it could be used to trace the authenticity of content.
- New technologies rely on continuous, pervasive and often unacknowledged and invisible tracking of individuals online and offline (e.g. real-time face recognition).
- With private messaging services on the rise, as opposed to public social media platforms, there is a risk that disinformation becomes submerged and is less apparent to researchers and policy-makers.
- The exact correlation between disinformation and the political opinion and voting behaviour of individuals is not yet scientifically proven.
- The correlation between the lack of diversity (whether supply-side, such as with an autocratic media system, or demand-side, such as with filter bubbles) and susceptibility to manipulation would be worthy of further research.

5.1 Development of disruptive technologies and their impact on human rights and democracy[507]

In this section, the visible trends of future technologies, how they may impact human rights and democracy, as well as the global political implications of the disinformation and propaganda actions are assessed.

5.1.1 Overview of technological trends

Given that the technologies discussed in this section are still developing, analysis of their possible application in the civic domain is a highly speculative exercise. Nevertheless, it is reasonable to assume that the disruptive technologies currently penetrating or about to penetrate the commercial market may very soon find their application in the political sphere, including for the purposes of informational manipulation. This section will discuss the practices emerging across different markets, without attempting to assess the likely effectiveness of these techniques.[508] Finally, the section will look into the challenges these technologies pose to the functioning of the rule of law.

5.1.2 Machine learning and 'deep fakes'

According to the researchers from Stanford University, the modified videos and imagery disseminated nowadays still exhibit many artefacts, which makes most forgeries easy to spot. But Gartner predicts that by 2022, the

[507] The authors of this report are grateful for valuable consultations and critical feedback on this section provided by the experts in data analytics, distributed ledger technologies and software engineering: Alex Comunian, Gabrielle Pellegrinetti and Giulia del Gamba and during the preparation of this section. Responsibility for any errors in the resulting work remains with the authors.

[508] Same reservations were expressed by the authors of the report commissioned by the UK Information Commissioner's Office, see Bartlett J et al., The Future of Political Campaigning. DEMOS: 2018, p. 26, https://www.demos.co.uk/wp-content/uploads/2018/07/The-Future-of-Political-Campaigning.pdf

majority of people residing in mature economies will consume more false information than true information.[509] Artificial Intelligence (AI) is named as a primary driver of the future 'counterfeit reality', where telling the difference between the original and manipulated content will become close to impossible for people and progressively difficult for machines.[510]

Deep learning, a subfield of machine learning,[511] has been increasingly experimented with to create realistic manipulations of video and imagery ('deep fakes').[512] A quick overview of the 'deep fakes' techniques,[513] suggests some examples of its future application.

Table 11: Application of 'deep fakes' technology

Technique	Example of application
Changing how real people appear to behave	Manipulating the gaze of someone in a photograph to change their apparent reaction Changing the expressions of someone in a video by making it appear they were saying something that was not actually said Creating a fake video of a person speaking using his or her own voice to change the context of where comments were made Manipulating the words of someone using samples of their own voice to create a controversial statement of a person
Creating fake people	Creating realistic photographs of non-existent people to mask a fraudulent social media account Creating realistic speaking voices for non-existent people, for example, to leave a voicemail to someone
Changing the appearance of where something took place	Adding a virtual environment to live video to include new subjects in the video Changing the time of day or weather in a photo or video to make it seem that the event has happened at a certain place at a certain time

Source: Washington Post.[514]

Whilst these techniques were previously used in the film-making industry, now they are becoming "as commonly available as today's meme-generating apps" and can potentially be used for various purposes, including entertainment, consumer deception and, as the experts predict, political disinformation.[515] **The key trend with**

[509] Panetta, K., Gartner Top Gartner Top Strategic Predictions for 2018 and Beyond. Gartner, 3 October 2017, https://www.gartner.com/smarterwithgartner/gartner-top-strategic-predictions-for-2018-and-beyond/

[510] Ibid., see also Kim, H. et al., M., Deep Video Portraits. Journal ACM Transactions on Graphics 37(4): 2018.

[511] On the relationship between deep learning, machine learning and artificial intelligence, see the Future of Privacy Forum, The Privacy Expert's Guide to Artificial Intelligence and Machine Learning, 2018, https://fpf.org/wp-content/uploads/2018/10/FPF_Artificial-Intelligence_Digital.pdf

[512] Deep fake technology relies on the 'generative adversarial networks' approach. As explained, '[o]ne network learns to identify the patterns in images or videos to recreate, say, a particular celebrity's face as its output. The second network acts as the discriminating viewer by trying to figure out whether a given image or video frame is authentic or a synthetic fake. That second network then provides feedback to reinforce and strengthen the believability of the first network's output.' (Hsu, J., Experts Bet on First Deepfakes Political Scandal, 22 June 2018, https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/experts-bet-on-first-deepfakes-political-scandal)

[513] Ibid.

[514] Bump, P., Here are the tools that could be used to create the fake news of the future, 12 February 2018, https://www.washingtonpost.com/news/politics/wp/2018/02/12/here-are-the-tools-that-could-be-used-to-create-the-fake-news-of-the-future/?noredirect=on&utm_term=.19673e753072; also see Witness and First Draft, Mal-uses of AI-generated Synthetic Media and Deepfakes: Pragmatic Solutions Discovery Convening, 11 June 2018, http://witness.mediafire.com/file/q5juw7dc3a2w8p7/Deepfakes_Final.pdf/file

[515] Hsu, J., Experts Bet on First Deepfakes Political Scandal, 22 June 2018, https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/experts-bet-on-first-deepfakes-political-scandal.

respect to the application of machine learning to video content production is its increased accessibility and quality.

It can be argued that 'deep fakes' present an even more difficult problem than manipulated textual content, as they are **more likely to trigger strong emotions** than simple text,[516] and are less likely to be critically assessed before being 'consumed'. According to Sundar, people process audiovisual content based on a 'realism heuristic' as they assume that audiovisual content has a higher resemblance to the real world than textual and verbal content.[517] Arguably, if an infamous 'Pizzagate' story was accompanied by video 'evidence' of children being held in captivity, it might have taken more time and effort to debunk it than the original false story.

5.1.3 Advanced demographic analytics

Tracking technologies are growing progressively in sophistication and capabilities to monitor people's behaviour across different platforms.[518] The Internet of Things (IoT) presents new possibilities to gain real-time hyper-personal insights into peoples' behaviour. In contrast with the data about an individual's online behaviour, information coming from IoT is higher in volume, less situational and therefore more complete. In some instances, an individual's behaviour is observed, and data collected 24/7.[519] Facial recognition technology provides another lucrative biometric data source for private and state actors. For example, technology marketed by Microsoft can detect faces, facial features and emotions and match them against the existing repository of 'up to one million people'.[520]

Ability to profile individuals by drawing rich inferences about them is recognised as one of the key trends in the future development and application of these new technologies. There exists a myriad of traits that can be measured, and used to infer new data about individuals, such as their marital status and sexual orientation. More advanced profiling practices allow scoring or assessing people against benchmarks of "predefined patterns of normal behaviour". One example is an analysis of typing patterns on a computer keyboard to predict a person's confidence, nervousness, sadness, and tiredness.[521] A model used for psychometric profiling – OCEAN or 'Big Five' model – assesses individuals across five personality traits (openness, contentiousness, extraversion, agreeableness and neuroticism) and promises to reveal "the basic structure underlying the variations in human behaviour and preferences".[522] Other models supplement this assessment with additional characteristics, including values and needs.[523] Sentiment analysis applied to the datasets makes it possible to infer a person's position, attitude or opinion towards a specific topic, and recent advances in the science, including facial recognition technology, promise to make application of this technique even more effective.[524]

[516] Lin, H., The Danger of Deep Fakes: Responding to Bobby Chesney and Danielle Citron, 27 February 2018, https://www.lawfareblog.com/danger-deep-fakec-responding-bobby-chesney-and-danielle-citron.

[517] Tucker, J. et al., Social media, political polarization and political disinformation: a review of the scientific literature. Hewlett Foundation, March 2018, p. 22, https://hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf

[518] For example, until recently, device fingerprinting worked only if people continued to use the same browser—once they switched to another browser, the fingerprint was no longer very useful. In 2017, a method was published allowing to track a person across multiple browsers on the same device.518

[519] Helberger, N., Profiling and Targeting Consumers in the Internet of Things – A New Challenge for Consumer Law in "Digital Revolution" (pp. 135-161), Baden-Baden: Nomos Verlag, 2017.

[520] Singer, N., Microsoft Urges Congress to Regulate Use of Facial Recognition, 13 July 2018, https://www.nytimes.com/2018/07/13/technology/microsoft-facial-recognition.html

[521] Kaltheuner, F. and Bietti, E., Data is power: Towards additional guidance on profiling and automated decision-making in the GDPR, IRP&P, p. 4, https://www.winchesteruniversitypress.org/site/journalsindex.php/jirpp/article/view/45

[522] OCEAN, an acronym for openness, conscientiousness, extroversion, agreeableness, neuroticism. See: Grassegger, H. and Krogerus, M., The Data That Turned the World Upside Down, 28 January 2017, https://publicpolicy.stanford.edu/news/data-turned-world-upside-down

[523] For example: IBM, The science behind the service, 13 May 2018, https://console.bluemix.net/docs/services/personality-insights/science.html#science

[524] Bartlett J et al., The Future of Political Campaigning. DEMOS: 2018, pp. 18-20.

Profiling based on the OCEAN model gained particular prominence after media reported it being employed by campaigners during the 2016 US presidential elections.[525] **Whilst there is no conclusive evidence of its use in national elections or referenda in EU Member States,[526] national campaigns were reported to cooperate with agencies offering advanced data analytics services.[527]** There is also a growing industry of voter management systems that integrate voter data, profiling and content automation capabilities.[528] On the one hand, these commercially available applications can be used to understand the voters better and to improve the quality of engagement with them. At the same time, like any other technology, they can be exploited for malicious purposes – to gain access to voter data to exploit people's fears and vulnerabilities in order to ultimately manipulate their views and decisions.

5.1.4 Personalised targeting and content automation

Marketers are increasingly seeking to target consumers on an individual and personalised basis, which is made possible thanks to the pervasive tracking and advanced demographic analytics described in the previous section. **Combined with the increasingly sophisticated tools to monitor and measure the response to messages, content personalisation capabilities may come in very handy in designing viral messages.** As explained by Moore and Tambini, in this process, "messages are selected on the basis of their resonance rather than ideological or political selection".[529]

Another related trend in content delivery is its automatic generation. Natural Language Generation tools could be used together with micro-targeting to automatically generate content for unique users based on the insights about their personal, psychological and other characteristics. This technology is applied in the use of commercial chatbots, such as personal shopping assistants and polling chatbots. Chatbots were also used to guide voters during political campaigns in the US.[530] According to DiResta, in the context of informational manipulation, this development could enable the adversaries to run even more automated accounts and to make them "virtually indistinguishable from real people".[531]

5.1.5 Virtual reality and other media

Virtual reality (VR) technology aims to create an entirely immersive experience that fully transports the user away from reality and into a virtual world.[532] While VR equipment has existed since the 1960s, they are fairly recent arrivals on the commercial market: the first headset with fully-realised VR capabilities became commercially

[525] Vogels, R., Trump, Micro Targeting And The Mechanisms Of Data Capitalism, 17 December 2016, https://www.huffingtonpost.com/entry/trump-micro-targeting-and-the-mechanisms-of-data_us_585433c0e4b0d5f48e164efc

[526] Bodó, B. et al., Political micro-targeting: a Manchurian candidate or just a dark horse? Internet Policy Review 6(4):2017, p.8

[527] European Data Protection Supervisor (EDPS), Opinion on online manipulation and personal data, 3/2018, p. 11, https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf; Bashyakarla, V., Psychometric Profiling: Persuasion by Personality in Elections, https://ourdataourselves.tacticaltech.org/posts/psychometric-profiling/; Bennett, C. J., Voter databases, micro-targeting, and data protection law: can political parties campaign in Europe as they do in North America? International Data Privacy Law, 6(4), 261–275.

[528] Political Campaign Tools - Running a Digital Campaign, https://s3-eu-west-1.amazonaws.com/ecanvasser.com/blueprint/Political_Campaign_Tools-Running_A_Digital_Campaign.pdf; Bartlett J et al., The Future of Political Campaigning. DEMOS: 2018, p. 30.

[529] Bartlett J et al., The Future of Political Campaigning. DEMOS: 2018, p. 33.

[530] Ibid., p. 34.

[531] National Endowment for Democracy, The Big Question: how will 'deepfakes' and emerging technology transform disinformation, https://www.ned.org/the-big-question-how-will-deepfakes-and-emerging-technology-transform-disinformation/

[532] Adams, D. et al., Ethics Emerging: the Story of Privacy and Security Perceptions in Virtual Reality. Conference paper for Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018), August 2018, p. 1

available in 2016.[533] The VR market has been growing ever since, with VR revenue projected to grow exponentially in the upcoming years.[534]

VR technology is becoming an important medium in the world of commercial advertising.[535] In addition, there is an observed rise in the use of the VR in non-commercial settings, aimed at assisting in the treatment of health conditions and addressing pressing social and political issues.[536] Civil society organisations, such as Planned Parenthood (US), have launched VR campaigns to support their cause.[537]

There is evidence of VR's limited use in political campaigns.[538] Whilst there are no publicly documented use cases of VR in informational manipulation actions, experts in the field consider this as a future trend meriting special attention.[539] The key risks to users associated with VR fall broadly into three categories: data collection and inferences; physical harm and manipulation, and violation of immersive experiences.[540] In VR, a user immerses themselves into a world created by another person, which enables the creator to control the participant to a far greater extent than any other technology. In addition, as pointed out by O'Brolcháin et al., VR social networks may create a 'global village' with stronger in-group discourse than the one available in current social networks,[541] and, potentially, may reinforce group polarisation. **Given the participatory nature of disinformation dissemination discussed in this study, VR may further amplify its effects and increase the persuasiveness of the manipulated message.**

Although VR attracts a lot of attention due to its immersive character and media 'buzz' around it, there is an even stronger case to be made for augmented reality becoming a new medium of information. In comparison with VR, augmented reality does not imply a complete immersion into the digital world experience, but rather adds digital elements to a live view (e.g. Snapchat lenses, 'Pokemon Go' game). Also, augmented reality does not require additional equipment and may be enabled by downloading an application on a smartphone. Augmented reality, being more affordable than VR is currently being tested to deliver ads by digital platforms.[542] The risks associated with augmented reality technology are similar to those with VR.[543]

[533] Ibid.

[534] Statista, Virtual reality software and hardware market size worldwide from 2016 to 2020, by platform (in billion U.S. dollars), https://www.statista.com/statistics/528793/virtual-reality-market-size-worldwide-by-platform/

[535] Collections: VR in Advertising, https://www.adsoftheworld.com/collection/vr_in_advertising

[536] Think Facebook can manipulate you? Look out for virtual reality, 21 March 2018, https://theconversation.com/think-facebook-can-manipulate-you-look-out-for-virtual-reality-93118; Ugolik, K., Can Virtual Reality Change Minds on Social Issues? 11 January 2017, https://narratively.com/can-virtual-reality-change-minds-social-issues/

[537] Ugolik, K., Can Virtual Reality Change Minds on Social Issues? 11 January 2017, https://narratively.com/can-virtual-reality-change-minds-social-issues/

[538] Ingham, L., Holograms, VR and in-game ads: meet the future of the political campaign, 7 September 2015, https://www.factor-tech.com/feature/holograms-vr-and-in-game-ads-meet-the-future-of-the-political-campaign/

[539] NiemanLab, Disinformation gets worse, 2018, http://www.niemanlab.org/2017/12/disinformation-gets-worse/. Also note that the number of the virtual reality users worldwide is growing, see: Statista, Number of active virtual reality users worldwide from 2014 to 2018 (in millions), https://www.statista.com/statistics/426469/active-virtual-reality-users-worldwide/

[540] O'Brolcháin F. et al., The convergence of virtual reality and social networks - threats to privacy and autonomy, Science and Engineering Ethics, February 2016, 22(1), p. 6, https://link.springer.com/article/10.1007/s11948-014-9621-1

[541] Adams, D. et al., Ethics Emerging: the Story of Privacy and Security Perceptions in Virtual Reality. Conference paper for Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018), August 2018, p. 2; also see Virtual Reality: Issues and Challenges, http://web.tecnico.ulisboa.pt/ist188480/cmul/issues.html

[542] Facebook brings augmented reality ads to the news feed, 11 July 2018, https://www.proactiveinvestors.com/companies/news/200609/facebook-brings-augmented-reality-ads-to-the-news-feed-200609.html

[543] Adams, D. et al., Ethics Emerging: the Story of Privacy and Security Perceptions in Virtual Reality. Conference paper for Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018), August 2018, p. 2

5.1.6 SEO manipulation and voice-activated search

Analysis of former disinformation campaigns suggests attempts have been made to manipulate search algorithms in order for the desired content to appear on the top of the search results. Since then, the industry of search engine optimisation (SEO), used for both legitimate and dubious purposes, has undergone some transformations. For example, search engine results pages are changing from the long lists of web pages to direct responses to the query in question (**'rich answers'**).[544] Alongside 'rich answers', users of Google Search may now see several recommended questions under a 'people also ask' heading. It is possible that these functionalities will be subject to manipulation by malicious actors in the future.

As more and more people are using **voice-activated search**[545] and **virtual assistants**, there is increased demand for **voice search data** and greater efforts to include these data in the personalisation of search results.[546] The results delivered by virtual assistants can be further **manipulated by exploiting device vulnerabilities**.[547] The growth in voice device interactions may also increase the volume of non-screen-based content, such as podcasts – a channel that can be further exploited for disinformation purposes.[548]

5.1.7 Distributed ledger technologies

Distributed ledger technologies (DLT) pose a separate kind of challenge for policymakers. According to the research report commissioned by the British Standards Institution:

> A distributed ledger is a digital ledger that is different from centralized networks and ledger systems in two ways. First, information is stored on a network of machines, with changes to the ledger reflected simultaneously for all holders of the ledger. Second, the information is authenticated by a cryptographic signature. Together, these systems provide a transparent and verifiable record of transactions. Blockchain technology is one of the most well-known uses of DLT, in which the ledger comprises 'blocks' of transactions, and it is the technology that underlies the cryptocurrency Bitcoin.[549]

There are many regulatory issues, including taxation, jurisdiction and application of data protection rules, which are pending resolution when it comes to DLT. For instance, data protection experts point to the inherent tensions between Blockchain (a type of DLT) and the GDPR when it comes to determining the role of the controller, the

[544] For example, 'maybe you asked a question like "when was the Declaration of Independence signed?" and got an answer like "July 4, 1776"'. Or maybe you typed a general query like "Wizard of Oz" and saw a breakdown with a description of the film (DeMers, J. What Are 'Rich Answers' And How Do They Affect SEO?14 August 2016, https://www.forbes.com/sites/jaysondemers/2016/08/14/what-are-rich-answers-and-how-do-they-affect-seo/#73165da534df)

[545] In 2020, Gartner predicts that voice-activated searches will account for 30% of web-browsing sessions (May, B., Hey Google, How Do I Optimize For Voice Search? 20 August 2018, https://www.forbes.com/sites/forbesagencycouncil/2018/08/20/hey-google-how-do-i-optimize-for-voice-search/#7621acbf3800)

[546] Smik, D., Three Major Developments That Will Shape SEO In 2018, 14 February 2018, https://www.forbes.com/sites/forbesagencycouncil/2018/02/14/three-major-developments-that-will-shape-seo-in-2018/#1abc104f1b8e

[547] Carlini, N., Audio Adversarial Examples, https://nicholas.carlini.com/code/audio_adversarial_examples; Mamiit, A., Alexa, Siri, And Google Assistant Follow Malicious Voice Commands Hidden In Music, 11 May 2018, https://www.techtimes.com/articles/227465/20180511/alexa-siri-and-google-assistant-follow-malicious-voice-commands-hidden-in-music.htm

[548] DeMers, J., 7 Predictions For The Shape Of Content Marketing In 2020, 13 April 2017, https://www.forbes.com/sites/jaysondemers/2017/04/13/7-predictions-for-the-shape-of-content-marketing-in-2020/#2278fa86177d

[549] Deshpande, A. et al., Distributed Ledger Technologies/Blockchain: Challenges, opportunities and the prospects for standards, May 2017, p. 1., https://www.bsigroup.com/LocalFiles/zh-tw/InfoSec-newsletter/No201706/download/BSI_Blockchain_DLT_Web.pdf

feasibility of data anonymisation and the facilitation of the exercise of data subject rights.[550]
These challenges extend beyond the data protection domain into any type of content moderation. Conventional content moderation procedures (including ‘flagging’ and the notice-and-take-down systems) assume the presence of a centralised regulator and a technical possibility to remove the content swiftly. This is not necessarily the case in Blockchain, especially in the permission-less networks, where anyone is allowed to become a participating or validating ‘node’,[551] and there is no central authority to address.

On the other hand, **Blockchain** developers are looking into the application of this technology **to tackle a disinformation problem.** Initiatives currently under development include “an open protocol for tracking the credibility of news”, incentivising “discovery of truth” and “quality fact-checking”, the creation of social media platforms that use **digital identities, reputation systems and technology to indicate the authenticity** of digital media.[552] It seems, however, that these efforts mostly centre around one type of informational manipulation – verifiably false content – and do not look into other practices, such as dissemination of misleading content (for example, genuine information presented in the wrong context) or aggressive informational practices (for example, divisive advertising), also discussed in this report.

5.1.8 Algorithmic detection of disinformation

In 2017, Wendling reported on a considerable number of worldwide initiatives aimed at countering disinformation. These include both **human intervention to verify the veracity of information (fact-checking) and the use of algorithmic techniques and machine learning to identify and validate the content**. According to Figueira and Oliveira,[553] algorithmic detection methods can be broadly divided into three categories:

- Content-based, which focus on analysing the actual text of the informational piece;
- Based on the diffusion dynamics of the message (e.g. a number of tweets for a news headline of this type for a certain day);
- Hybrid, based on a weighted sum, or a group of features feeding a learning algorithm (e.g. both content-related and diffusion dynamics-related metrics).

For example, the linguistics-driven approach proposed by Perez-Rosas suggests differentiating between fake and genuine content by looking at the lexical, syntactic and semantic level of a news item in question. According to the authors, the developed system’s performance is comparable to that of humans in this task, with an accuracy up to 76 %.[554] Other works suggested extracting manually crafted features from news content such as the number of nouns, length of the article, fraction of positive/negative words, and more in order to discriminate

[550] Lyons, T. et al., GDPR and Blockchain. Thematic report prepared by the European Union Blockchain Observatory and Forum, 2018, https://www.eublockchainforum.eu/sites/default/files/reports/20181016_report_gdpr.pdf?width=1024&height=800&iframe=true

[551] Ibid. Also see : Meserole C. and Polyakova A., Disinformation Wars, 25 may 2018, https://foreignpolicy.com/2018/05/25/disinformation-wars/ and Greenspan, G., The Blockchain Immutability Myth, 4 May 2017.

[552] Greenup, S., Catalogue of all projects working to solve Misinformation and Disinformation, 9 June 2018, https://misinfocon.com/catalogue-of-all-projects-working-to-solve-misinformation-and-disinformation-f85324c6076c; Belise, G., Blockchain can eliminate fake news, 27 March 2018, https://the-blockchain-journal.com/2018/03/27/blockchain-can-eliminate-fake-news/; Huckle, S. and White, M., Catalogue of all projects working to solve Misinformation and Disinformation. Big Data 5(4) : 356-371, 2017

[553] Figueira, A., and Oliveira, L., The current state of fake news: challenges and opportunities, Procedia Computer Science 121 (2017): 817–825, pp. 820-822, https://ac.els-cdn.com/S1877050917323086/1-s2.0-S1877050917323086-main.pdf?_tid=59d836f5-89df-4fbc-a94b-19c1af0e1b36&acdnat=1545743311_63b257b853f17ff893a949210a7f4028; also see Shu, K., et al., Fake News Detection on Social Media: A Data Mining Perspective, https://arxiv.org/pdf/1708.01967.pdf

[554] Perez-Rosas, V. et al., Automatic Detection of Fake News, Proceedings of the 27th International Conference on Computational Linguistics, pages 3391–3401, Santa Fe, New Mexico, USA, August 20-26, 2018, http://aclweb.org/anthology/C18-1287

false articles.[555] Advances of image forensics algorithms were leveraged to create a browser verification plugin 'InVid', which is already used to debunk fabricated videos.[556]

There are **several limitations to the application of algorithmic detection techniques in the real world**, which are not yet fully overcome. The first stems from the nature of disinformation. Although some success has been achieved training algorithms to distinguish between false and true facts, more nuanced forms of disinformation (genuine information presented in the wrong context or aggressive informational practices) have not yet been sufficiently addressed. Another obstacle for the development of these solutions is a limited availability of data processed by the digital platforms,[557] as well a lack of transparency around the underlying social network's infrastructure. According to Humberman et al., "[w]hen explicit information on the social network is not available, the strength of the social links is hardly known and their importance cannot be deemed uniform across the network".[558] Lastly, some verification mechanisms (both algorithmic and curated by humans) may lead to unintended consequences. For example, a short-lived practice employed by Facebook to tag unverified stories as "disputed by 3rd party fact-checker", was found by the Yale University scientists to create an "implied truth" effect whereby untagged headlines were perceived by the users as more accurate.[559]

5.1.9 Disruptive technologies and their implications for democracy

As Kranzberg's first law states, "technology is neither good nor bad, nor it is neutral".[560] It can be argued that the same technology applied by different actors with varying motivations can lead to fundamentally opposite individual and societal outcomes. The risk level of a technology's particular deployment also depends on a variety of factors, including its selected features and functionalities.

This being said, the outlook for market trends identifies three key processes underlying the development of the disruptive technologies, which already have profound implications for democracy and fundamental rights.

The majority of the technologies discussed in this section rely on continuous, pervasive and often unacknowledged and invisible tracking of individuals online and offline (e.g. real-time face recognition). The data points collected about each member of a society are growing exponentially and now include biometric and genetic data. The industry is further capitalising on personal data, by 'repurposing' data collected in one context (e.g. loyalty cards) and using it in another, unrelated context (e.g. voter databases). Even more alarming is that applying machine learning algorithms to seemingly 'innocent' datasets (e.g. pages *liked* on Facebook) allows for sensitive and invasive inferences, such as the person's state of mental health. Independently of the truthfulness of such predictions, they are used to place individuals in clusters and to make decisions significantly affecting their lives, including, what information to serve to them. These decisions are often aimed at 'nudging' individual behaviour and are expected to transfer into the world of virtual and augmented reality.

[555] Bort, J. It took only 36 hours for these students to solve Facebook's fake-news problem. Tech Insider 2016, http://www.businessinsider.com/students-solve-facebooks-fake-news-problem-in-36-hours-2016-11; Silverman, C. and Alexander, L. How Teens In The Balkans Are Duping Trump Supporters With Fake News. 2016,https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trumpmisinfo?utm_term=.jxX7xRvNr0#.pnj8qZRxW; Zimdars, M. False, Misleading, Clickbait-y, and/or Satirical "News" Sources, https://docs.google.com/document/d/10eA5-mCZLSS4MQY5QGb5ewC3VAL6pLkT53V_81ZyitM/preview

[556] Teyssou, D. et al., The InVID Plug-in: Web Video Verification on the Browser, 2017, https://www.researchgate.net/publication/320570485_The_InVID_Plug-in_Web_Video_Verification_on_the_Browser

[557] E.g. Constine, J.,

Facebook restricts APIs, axes old Instagram platform amidst scandals, 4 April 2018, https://techcrunch.com/2018/04/04/facebook-instagram-api-shut-down/

[558] 7. Huberman, B.A., Romero, D.M., and Wu, F., Social networks that matter: Twitter under the microscope. 2008, https://arxiv.org/pdf/0812.1045.pdf

559 Pennycook, Gordon and Rand, David G., Assessing the Effect of 'Disputed' Warnings and Source Salience on Perceptions of Fake News Accuracy (September 15, 2017). Available at SSRN: https://ssrn.com/abstract=3035384, p. 9

[560] Kranzberg, M., Technology and History: "Kranzberg's Laws". Technology and Culture 27(3) : 544-560, July 1986.

Powered by advertising and the data brokerage industry, data-driven business models appear to be flourishing, with their products being indiscriminately available to everyone who has resources to purchase them – from car manufacturers to authoritarian governments, and disinformation campaigners. **And while there is a variety of actors on the demand side of technological innovations, the supply side is characterised by tremendous power concentration in the hands of the largest digital platforms.** They are "dominating the development and systems integration into usable AI services"[561] and experimenting with the application of machine learning to the gargantuan personal databases they control. **This, in turn, even further increases not only their market power but, even more importantly, the civic powers they already hold.**[562] As comprehensively explained by Moore, their civic powers have a direct effect on an individuals' ability to exercise their fundamental rights and on the functioning of the rule of law.

Table 12: Relationship between civic powers and democratic values

Civic powers of digital platforms		Rights, freedoms and values at stake
The power to command attention	→	Freedom of thought and right to privacy and data protection
The power to communicate news		Freedom of information and freedom of media
The power to enable collective action		Freedom of assembly and association
The power to give people a voice		Freedom of expression
The power to influence people's vote		Freedom and fairness of elections
The power to hold power to account		Transparency and accountability

Source: Martin Moore (*Tech Giants and Civic Power, 2016*)[563] and authors.

The extent of these civic powers has also been illustrated in the European Data Protection Supervisor's Opinion on online manipulation:

> Social media has been used to encourage people to vote, to vote for a particular candidate, and to discourage from voting altogether ('digital gerrymandering'). The major social media provider itself has encouraged voters to exercise their vote, and there is nothing to preclude them from doing the opposite. In comparison with the mainstream media outlet covering a news story, there is no trace or record of an editorial decision, only the results of filtering performed by an algorithm. Online intermediaries could, in theory, make it easier for a political party which their business or ideological interests align with to reach their supporters or vice versa, with former social media employees recently claiming to have been involved in keeping conservative issues from trending on the site. Whether allegedly dominant online platforms may (deliberately or not) use their power to influence voting or not is less the point than the fact that they – in principle – have the ability to influence political decision-making processes.[564]

The same considerations apply to informational manipulation, including dissemination of false and misleading information and aggressive informational practices discussed in this report. The challenge of disruptive technologies is not only that they can increase the quality of doctored information, but that their development

[561] Nemitz, P., Constitutional Democracy and Technology in the age of Artificial Intelligence, 18 August 2018, p. 7, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3234336

[562] Moore, M., Tech Giants and Civic Power. King's College London, April 2016, https://www.kcl.ac.uk/sspp/policy-institute/CMCP/Tech-Giants-and-Civic-Power.pdf

[563] Ibid.

[564] Internal references omitted. European Data Protection Supervisor (EDPS), Opinion on online manipulation and personal data, 3/2018, p. 13, https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf

will also **further strengthen the dominance of the intermediaries of the automated public sphere** "in all areas of activities they are already in, and extend it to others".[565]

And whereas the public powers vested with state institutions are conditioned on rule of law principles, including a responsibility to act within the constraints set out by law, the comparable civic powers are not grounded in similar frameworks. **Thus, the third process underlying the development of disruptive technologies is a regulatory and enforcement limbo.** This limbo occurs partially out of the fear that any regulation will 'stifle innovation' and of industry pressure, partially because of regulators working in silos,[566] and partially because of the ungrounded belief in the power of self-regulation. Ultimately, it results in the situation where great power does not come with great responsibility, and where it comes, it often happens *post factum* with the irreversible damage already caused, as the case of informational manipulation proves. Where the **right to privacy, freedom of expression and freedom and fairness of elections** are at stake, there is an evident need for the coherent enforcement of existing laws, including **competition and data protection laws**, and design of new enforceable rules setting out the boundaries and **accountability structures** for the development of the new technologies and the exercise of the civic powers.

5.2 How are international relations affected by disinformation and propaganda?

5.2.1 Interpretation under public international law

Dissemination of disinformation is sometimes conducted by a state, purposefully targeting the society of a foreign state. There is evidence that at least a few countries have purposefully manipulated public discourse, intending **to sow distrust** in the targeted societies (see more in section 1.3).[567] The Russian Defence Ministry openly defined such a strategy as 'information war': "to destabilise a society and a state through massive psychological conditioning of the population, and also to pressure a state to make decisions that are in the interest of the opponent".[568]

Apparently, an international struggle is taking place for geopolitical power, and borderless cyberspace offers a new opportunity to gain influence.[569] While targeting and attacking hard objects, such as the infrastructure of a foreign state, would undoubtedly amount to an aggression, targeting societies is a softer method. International legal jurisprudence has not yet established clear definitions of international cyber-activities. A highly esteemed collection of principles and prospective rules on international law applicable to cyber operations is the Tallinn Manual. It is, however, a product of jurisprudence – authored by nineteen international law experts – and has no legally binding effect.[570]

Neither the Tallinn Manual, nor other expert reports or the EU instruments (below) clearly distinguish between different incidents, such as: (1) cyber-attacks against the physical infrastructure of a society as well as its informational network; (2) cyber-attacks against the security of networks and services, hacking and manipulating algorithms or governmental networks; (3) online attacks of disinformation and propaganda targeted at manipulating the electorate. While the first type of incident is clearly acknowledged by the Tallinn Manual and

[565] Nemitz, P., Constitutional Democracy and Technology in the age of Artificial Intelligence, 18 August 2018, p. 7, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3234336

[566] European Data Protection Supervisor (EDPS), Opinion on coherent enforcement of fundamental rights in the age of big data, 8/2016, https://edps.europa.eu/sites/edp/files/publication/16-09-23_bigdata_opinion_en.pdf

[567] The Digital Hydra: https://www.stratcomcoe.org/digital-hydra-security-implications-false-information-online, at p. 23.

[568] Russell, Martin: Russia's information war: Propaganda or counter-propaganda? EPRS European Parliamentary Research Service. Members' Research Service PE 589.810. at p.2. http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/589810/EPRS_BRI(2016)589810_EN.pdf

[569] Although the exact definition can be debated by scholars, the analogy is crystal clear: this is *"a conflict over ideological differences carried on by methods short of sustained overt military action* Merriam-Webster Dictionary of Cold War.

[570] Jensen, Eric Talbot: The Tallinn Manual 2.0: Highlights and Insights. Georgetown Journal of International Law. Vol. 48. 2017. 735-778. Pages. 740-744.

also the EU instruments (see below), the second is a grey zone, and the third, information warfare, is not acknowledged by legal instruments. 'Information war' is mainly used in literature in relation to the disinformation campaigns organised by or from Russia. In sum, 'information war' is not a legal term and should not be regarded as such.

In each of the cases, the question of 'attribution' is one of the most problematic areas: if a state denies responsibility in a cyber activity, attribution of the attack to the (suspected) state is close to impossible. The EU takes the position that "attribution to a state or a non-state actor remains a sovereign political decision based on all-source intelligence and should be established in accordance with international law of state responsibility".[571]

The United Nations has taken note of the "disturbing trends that create risks to international peace and security". A specialised group of governmental experts was created to address the issues of cyberwarfare (GGE). In its report, the GGE made a recommendation for consideration by states, for voluntary, non-binding norms, rules or principles for responsible behaviour of states.[572] After five productive meetings, the sixth meeting did not produce a draft on the interpretation of an 'armed attack'.[573] The new definition involves high stakes: if malicious cyber-operations against another country would justify self-defence under the UN Charter, including collective self-defence,[574] and the much debated concept of pre-emptive strikes, this could lead to an open war. This frightening possibility appears to freeze negotiations. One compromise appears to offer itself: that aggressive cyber-operations would only justify self-defence limited to cyber-operations (*lex talionis)*. But even in that case, an openly fought cyberwar would certainly be more devastating than the clandestine, secretive cyber-attacks of today. Therefore, although it may be tempting to clarify the legal situation, it is only worth doing so if there is a plan for afterwards.

5.2.2 EU possible actions and responses

With regard to propaganda and disinformation around the EP elections, the potential intrusions of third countries into EU and Member State affairs could be interpreted in the light of public international law. The Commission's Communication on Elections forecasts that, together with the High Representative, the Commission will be supporting the preparation of common European responses addressing any foreign involvement in elections in the European Union.

International law also allows the application of '**retorsions':** lawful, but unfriendly measures towards another state. Being lawful measures, they can be applied any time, without regard to the legal qualification of the preceding measure by the other state, or the question of attribution. Retorsions are mainly limited to diplomatic responses, where the EU should act at least jointly. For this purpose, the EU has created "A blueprint for coordinated response to large-scale cross-border cybersecurity incidents and crises".[575]

The Tallinn Manual 2.0 recognises the possibility of **countermeasures**, with limited applicability. These are actions that would otherwise be unlawful, but for the fact that they are a response to an internationally wrongful act attributable to another state. They should not reach the level of force, which makes them perfect for unfriendly cyber operations, but their application is limited. First, the original malicious cyber activity has to be attributed to a state, not merely to a non-state actor operating from the state's territory. Second, no collective self-defence applies, only the state affected by the malicious cyber activity has the right to resort to

[571] Cyber Diplomacy Toolbox: to coordinate response of EU MSs to malicious cyber activities. https://data.consilium.europa.eu/doc/document/ST-7925-2018-INIT/en/pdf

[572] Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security. UN GA A/70/174, 22 July 2015. at p. 7.

[573] Schmitt, Michael - Liis Vihul: International Cyber Law Politicized: The UN GGE's Failure to Advance Cyber Norms. Just Security Blog. 2017. https://www.justsecurity.org/42768/international-cyber-law-politicized-gges-failure-advance-cyber-norms/

[574] Article 51. and 52. UN Charter. Such self-defense is lawful until the Security Council takes action.

[575] Commission Recommendation (EU) 2017/1584, and its Annex.

countermeasures, therefore the EU or its Member States cannot help an affected Member State. Third, they can be used only if the original malicious cyber activity presents a "grave and imminent peril" to an "essential interest" of the state, and the responsive measure is the "sole means of safeguarding it".[576]

The EU's existing policy strategies for cyber-attacks apply only to those cyber-attacks against infrastructure that cause material harm, or disruption of services (although, **according to the Tallinn Manual 2.0, neither physical damage nor injury is required for a cyber act to be an internationally wrongful act**). The EU has passed a Directive on the security of network and information systems (NIS Directive), and a cybersecurity package in 2017. Online disinformation is not discussed extensively, but is part of the competences of the **EU Hybrid Fusion Cell (HFC).** An important and effective measure could be to protect online information systems, along with user data, from malicious foreign cyber activity.[577]

Further possible actions could be organised along the lines of already existing EU-level defence and security operations (e.g. Europol, ECTC (European Counter Terrorism Centre), ENISA, TFTP, INTCEN, the EU Cyber Defence Policy Framework, CSDP, EEAS, EDA, the Emergency Response Coordination Centre (ERCC), the Computer Emergency Response Team of the EU Institutions (CERT-EU), the Hybrid Fusion Cell (HFC).[578]

5.3 Avenues for future research

The exact correlation between disinformation and the political opinion and voting behaviour of individuals is not scientifically proven.[579] US-related studies have shown that people have difficulties determining when a particular piece of news is false,[580] and many of those who see false stories actually believe them.[581] People are more likely to be affected by inaccurate information if they see more and more recent messages reporting facts, irrespective of whether they are true.[582] There is also reason to believe that audio-visual messages can be both more persuasive and more easily spread than textual messages, but we do not know nearly enough about these dynamics – most research to date has focused on textual rather than visual and audio-visual misinformation.[583] At the same time, people often provide multiple rationales for their opinions and do not strictly base them on facts.[584] Therefore, even if they are exposed to true or false information, it does not necessarily mean that they are going to act on it.

Researchers are also using empirical methods to study the functioning and effectiveness of digital amplification mechanisms. Some of the studies in this field suggest that manipulation of people's newsfeed or search results

[576] Schmitt, Michael N.: Tallinn Manual 2.0. Cambridge University Press, 2017. Rule 26.

[577] Network and Information Security (NIS) Directive (2016/1148); Joint Communication: Resilience, Deterrence and Defence: Building strong cybersecurity for the EU (JOIN(2017) 450 final).

[578] See more at: Commission Recommendation (EU) 2017/1584. 13. Sept. 2017. on coordinated response to large-scale cybersecurity incidents and crises. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017H1584&from=EN

[579] Roozenbeek, Jon and van der Linden, Sander, The Fake News Game: Actively Inoculating Against the Risk of Misinformation, From: https://www.cam.ac.uk/sites/www.cam.ac.uk/files/fakenews_latest_jrr_aaas.pdf , pp. 3-4: "Although extensive research exists on political misinformation (for a recent review, see Flynn, Nyhan, & Reifler, 2017), there is some debate about the extent to which fake news influences public opinion (Shao et al., 2017; van der Linden, 2017), including social media "echo chambers" and "filter bubbles" (Bakshy, Messing,

& Adamic, 2015; Flaxman, Goel, & Rao, 2016; Fletcher & Nielsen, 2017; Guess et al., 2018)".

[580] European Parliament, Fake news' and the EU's response (April 2017), http://www.europarl.europa.eu/RegData/etudes/ATAG/2017/599384/EPRS_ATA%282017%29599384_EN.pdf

[581] Allcott H. and Gentzkow M., Social Media and Fake News in the 2016 Election (Spring 2017). Stanford University, Journal of Economic Perspectives, Vol. 31, No. 2, pp. 211-236. https://web.stanford.edu/~gentzkow/research/fakenews.pdf, p. 212

[582] Tucker, J. et al., Social media, political polarization and political disinformation: a review of the scientific literature. Hewlett Foundation, March 2018, pp. 40-48. https://hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf

[583] Ibid.

[584] Ibid., p. 51.

could influence their voting behaviour: for example, when social platform users were told how their friends had said they had voted, this prompted a statistically significant increase in the segment of the population (0.14 % of the voting age population or about 340 000 voters) to vote in the congressional mid-term elections in 2010.[585] In another study, researchers claimed that differences in Google Search results were capable of shifting the voting preferences of undecided voters by 20 %.[586] In the commercial context, empirical research has shown how adapting messages to the psychological characteristics of individuals can directly influence their (purchase) behaviour.[587] Other researchers challenged the conventional wisdom of the realities of political micro-targeting. The paper, published in August 2018, found evidence that political "messages [during the 2017 UK general election campaign] adhere closely to national campaign narratives" and "did not appear to be greatly more negative than other traditional modes of communication".

Empirical studies about the effect of propaganda on public opinion are rare, but recent independent empirical (longitudinal) academic research studies have found that Hungarian governmental communication putting migration in a negative light resulted **in a growth of xenophobia** especially among people living in the countryside, the less educated, and the elderly,[588] i.e. that part of the population whose informational environment was limited to government-friendly media because of the regional or social or educational circumstances.[589] The correlation between lack of diversity (whether supply-side such as with an autocratic media system, or demand-side, such as with filter bubbles) and susceptibility to manipulation would be worthy of further research.[590]

Cause and effect relationships had also been very much contested in traditional media theory.[591] The correlation between violent audiovisual content and the harm caused among children and youth, as well as the effect of content on the formation of political opinions has been examined repeatedly by experiments, and the results were often contradictory. This did not prevent legislators around the globe from restricting violent content and hate speech on mass media.

The regulation of mass media has been justified by three theoretical arguments: (1) The "pervasive effect"[592] of audiovisual media. While the early internet was regarded as a 'pull' type medium demanding a more conscious

[585]; Allcott H. and Gentzkow M., Social Media and Fake News in the 2016 Election (Spring 2017), Stanford University, Journal of Economic Perspectives, Vol. 31, No. 2, pp. 211-236., p.219)

[586] Zuiderveen Borgesius, F. & Trilling, D. & Möller, J. & Bodó, B. & de Vreese, C. & Helberger, N. (2016). Should we worry about filter bubbles?. Internet Policy Review, 5(1). DOI: 10.14763/2016.1.401, p. 9. Regarding search engine manipulation, also see: Epstein R and Robertson RE, 'The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections.' (2015) 112 Proceedings of the National Academy of Sciences of the United States of America E4512. http://www.pnas.org/content/112/33/E4512.abstract?tab=author-info

[587] Matz, S.C. et al., Psychological targeting as an effective approach to digital mass persuasion, PNAS November 28, 2017 114 (48) 12714-12719, https://www.pnas.org/content/114/48/12714

[588] Kolozsi Ádám (2016): Sosem látott mértékű a magyarországi idegenellenesség, https://index.hu/tudomany/2016/11/17/soha_nem_latott_merteku_az_idegenellenesseg_magyarorszagon/ (letöltés: 2018. XI. 1.).

[589] Mérték Médiaelemző Műhely (2016–2018): Szúrópróba, http://mertek.eu/wp-content/uploads/2018/07/Sz%C3%BAr%C3%B3pr%C3%B3ba-25.pdf (letöltés: 2018. XI. 4.).

[590] See also the results that social media usage that involves participation in several networks reduces mass political polarization and echo chambers. Martens, Bertin, Luis Aguiar, Estrella Gomez-Herrera Frank Mueller-Langer (2018), "The digital transformation of news media and the rise of disinformation and fake news. An economic perspective", JRC Digital Economy Working Paper 2018-02. https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf 27.

[591] See the contesting theories of Harold Lasswell (bullet, 1927), Paul Lazarsfeld (two-step influence, 1948), Joseph Klapper (selective perception, 1949), George Gerbner (cultivation, 1969), McCombs and Shaw (agenda-setting, 1972), Herman and Chomsky (framing, 1988), Dayan and Katz (performative effect, 1992) - to name a few.

[592] Federal Communications Commission v. Pacifica Foundation, 438 U.S. 726 (1978)

consumption attitude from users, as opposed to the 'push' type of television,[593] web 2.0 design, streaming video and especially handheld devices have radically changed this. Content selection algorithms also aim at maximising user-engagement and making the service addictive.[594] Today's social media encounters can be addictive and intrusive. (2) Reaching masses of people, including children. While the information landscape today is scattered, the market of social media platforms is significantly more concentrated than that of traditional media, with Facebook having 2 234 million users in 2018.[595] (3) Scarcity of resources. While the scarcity of material resources is not any longer a hindrance, the scarcity of attention is becoming a significant obstacle in access to a diversity of content.[596]

Further research is needed on the demographic characteristics of those people who are most susceptible to manipulation. According to a recent study, a **significant generational divide** can be observed: people over 65 share seven times more fake news than young users.[597]

The generational divide can also be observed in the rapidly changing trends: the young generation prefers messaging services such as Snapchat,[598] and WhatsApp, and more picture-based platforms like Pinterest or Instagram, whereas Facebook's popularity (used by their parents and grandparents) has been slowly sinking (although still very dominant). The popularity of private messaging platforms and apps carries the risk that harmful content becomes submerged and is less apparent to researchers and policy-makers.

Overall, the use of social media in social and political communication and the effects of exposure to information and disinformation on individual beliefs and behaviour is one of the key areas that needs to be addressed in future research.

[593] Ashcroft v. American Civil Liberties Union, 535 U.S. 564 (2002) Lessig, Lawrence: What Things Regulate Speech: CDA 2.0 vs. Filtering. https://cyber.harvard.edu/works/lessig/what_things.pdf. See also: Two eras of the internet: pull and push. 21.12.2014. http://cdixon.org/2014/12/21/two-eras-of-the-internet-pull-and-push/

[594] European Data Protection Supervisor (EDPS), Opinion on online manipulation and personal data, 3/2018, p. 13 , https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf

[595] Instagram at the 6th place had 1 billion users. [595] See numbers at: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/. See also: Social Media Use in 2018. http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/

[596] Helberger, N., Katharina Kleinen-von Königslöw and Rob van der Noll (2015), "Regulating the new information intermediaries as gatekeepers of information diversity", info, Vol. 17, No. 6, p.50-71, (https://doi.org/10.1108/info-05-2015-0034)

[597] Andrew Guess, Jonathan Nagler and Joshua Tucker: Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Sci Adv* 5 (1), eaau4586. DOI: 10.1126/sciadv.aau4586

[598] Kantar Media: News in social media and messaging apps. Qualitative research report Prepared for the Reuters Institute for the Study of Journalism, University of Oxford with the support of the Google News Initiative. Sept. 2018.

6. CONCLUSIONS

KEY FINDINGS

1) The threat to democracy

- Global trends show that the phenomena of populism and authoritarianism are not isolated flaws in the system, but have a chance of becoming systems themselves.
- Therefore, tools for strengthening democratic resilience must be defined, protected and preserved: a higher degree of rigidity is needed in the constitutional system so that changes in the type of governance should not be able to touch the cornerstones.
- Research shows that the sources of disinformation or propaganda in EU Member States are a mixture of government agencies, politicians and parties, private contractors and civil society organisations. Either way, several national governments seem to have a vested interest in not tackling the issue via legal or policy measures.
- All measures that empower Member States can be used to reach the opposite goals in the hands of captured states or authoritarian states. In the project to defend liberal democracy, no cooperation can be expected from states that do not support liberalism.

2) Political communication

- Social media has given voice to the underprivileged classes of society whose opinions were less visible before, but who tend to be weaponised to share manipulative content. Their voices have been taken over by opportunistic political groups to realise particular political interests.
- The rules of political and public issue advertising should be re-regulated and harmonised in the Member States.
- Political parties should initiate self-regulation and commit to ethical principles, aimed at fair, transparent, constructive and reason-based campaigns.
- The financing of political parties and political campaigns should be re-regulated to ensure complete transparency.

3) The media system

- The EU should invest in developing a European identity, through the concept of citizenship (*Bürgerschaft*). Some basic values expressed in Article 2 TEU, like human rights and the rule of law must be commonly shared in the EU area. The communication should be user-friendly enough to replace the populistic narratives with new ones.
- The EU should invest in supporting a diverse and high-quality media system, enhanced through fact-checking, credibility indices and investigative journalism. The creation and support of a European supranational news media should be considered.
- Public service broadcasters should be relied on; however, the Commission should supervise their operation and funding in states where doubts are raised relating to their ethical operation, or their abuse for propaganda purposes.
- Platform providers' legal rights and obligations should be clearly regulated. They are not responsible for third-party content but responsible for their algorithms, for respect for human rights including data protection and for administering their platforms (discussed in more detail in chapter 7 on recommendations).
- Algorithms must be regulated: their principles should foster diversity, prioritise trustworthy content, be transparent and give users options on which principles they use or reject.

- The monopoly situation of certain social media platforms should be exhaustively analysed; reducing concentration, introducing interoperability or imposing 'quasi-public service' obligations on dominant actors should be considered.
- The act of micro-targeting based on sensitive information violates human dignity, the right to freedom of (truthful) information and distorts public discourse – its minimum requirement should be opt-in consent by the user.
- New technology, and social media as such, makes individuals vulnerable to having their personal data exploited. This calls for imposing a higher level of responsibility on social media service providers. The GDPR and the anticipated ePrivacy Regulation, as well as their enforcement, are of crucial importance.
- Future technologies (including AI) may have devastating effects on human autonomy, if privacy rules are not consistently enforced. Therefore, a change in the attitude of advertisers and service providers towards personal data should be encouraged, similar to the confidentiality obligations of doctors or lawyers.
- The terms applied by dominant social media providers for verifying user profiles, and limiting user profiles to one per email address, should be consequently required from other platforms as well.
- Corporate and political actors should be subject to a higher level of verification than individuals.
- The criminalisation of the most grievous forms of (organised) disinformation actions may be considered.

The EU has already been active in tackling disinformation and propaganda, as the numerous soft and hard instruments show. More needs to be done to create a safe environment for fair political discourse, and to delineate the rights and responsibilities of platform services.

6.1 Definitions

The word 'disinformation' is recommended to designate deliberately false, distorted or misleading information, rather than the term 'fake news', which has been used in dubious political contexts.

To define content that is not subject to verification, such as biased or exaggerated opinions or manipulated content aimed at misleading the audience (especially content inciting negative emotions), spread with the intention to manipulate political views, we use the term 'propaganda'. It has been observed that the impact is largely influenced by unethical dissemination methods, which could also be taken as a distinctive factor.

The main elements of the definitions are the (i) type of information, (ii) falsity of information, (iii) intention of the author, and (iv) consequence of dissemination.

Such informational manipulation campaigns can originate from domestic or foreign sources, which can be either private or state actors. Their content could be soothing and intending to disrupt a political discourse and distract the participants, or inflammatory in order to inspire strong emotions like fear, disgust, surprise or anger. They can spread spontaneously through actions of the audience or they can be amplified by automated means: using political bots or social bots to spread masses of similar content, reaching the audience through micro-targeting or tricking the algorithms by false 'likes'.

6.2 Building democratic resilience

The tendencies observed in public discourse threaten democratic processes and the institution of democracy. Political figures and parties have taken advantage of citizens' wider and interactive participation in the online public discourse. It should be examined whether social cohesion, stability and **support for democracy could be won by tangible measures on the ground** – outside the realm of communication and media – such as measures **addressing real concerns**, aimed at **social inclusion, education and the empowerment of citizens**. Populistic politicians sometimes touch upon real fears and concerns of people and appear to represent them sensitively. In contrast, democratic decision-making sometimes appears to be distant, detached and beyond the reach of ordinary voters. In other instances, decision-making may seem too slow, hesitant or even corrupt, and incapable of solving 'real' problems (see section 2.2).

Deliberative democracies need to respond to public fears, take notice of the risk-based society and counteract. This responsiveness must remain committed "to deliberation, in the form of reflection and reason giving".[599]

This study does not have the ambition to reach beyond pointing at the possibility that the current crisis of democracy could signal a substantial **shake-up of trust in the way the democratic system operates**. **In this case, tackling the communication problems alone will not lead to long-term results**.

6.2.1 Defending the constitutional institutions of the democratic state and of the EU

Global trends show that the phenomena of populism and authoritarianism are not isolated flaws in the system, but have a chance of becoming systems themselves. Within the EU, the recipe appears to be passed on from state to state, from one 'charismatic' leader to the other.[600] The EU faces a specific challenge with respect to its multi-level structure: manipulations of people's access to information may come and stay within a Member State, may come from one Member State to another, or from third countries or non-state actors. Research shows that the sources of disinformation or propaganda in EU Member States are a mixture of government agencies, politicians and parties, private contractors, and civil society organisations.[601] Either way, several national governments have a vested interest in not tackling the issue via legal or policy measures.

However, the EU suffers a deficiency of its democratic legitimacy in all cases, even if the manipulation curbs participatory democracy only within one Member State. Either through the representative functions in the institutions or directly through the EP elections, the EU's democratic values are violated in both cases.

Beyond tackling disinformation and propaganda as communication, also the **basic tenets of the democratic operations must be nailed down, protected and preserved. A greater degree of rigidity in the constitutional system is required,** so that changes in governance should not be able to touch these cornerstones. Earlier research by some of the authors recommended building safeguards into the system and early warning procedures to signal democratic backslide, and reacting as soon as alarming signals can be seen.[602] Regular supervision of the constitutional structure, carried out by a network of experts who know the context and the circumstances, is necessary.

If disinformation and propaganda are disseminated by states, then in the first stage – shown as scenario C in Table 8 on attempts to manipulate democratic processes in the EU – domestic legal remedies could still be used. However, in scenario D, **when the goal of the government is to challenge democracy, then the national correction mechanism, including the media, the parliamentary opposition, the prosecution or the judicial system, might be captured already and cannot be utilised.** Between the two stages, international legal supervision could still be used – when domestic mechanisms are too weak already, but the rule of law is still

[599] Sunstein, *supra* note 250, 1.

[600] Finchelstein, Federico (2017) From fascism to populism in history, Oakland, University of California Press.

See also: Lochocki, T. (2018). The Rise of Populism in Western Europe. [electronic resource] : A Media Analysis on Failed Political Messaging. Cham : Springer International Publishing : Imprint: Springer, 2018.

[601] With the exception of Hungary, where the source for disinformation and propaganda was exclusively the government. See: Samantha Bradshaw, Philip N. Howard, Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation, 2018, http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct2018.pdf

[602] Carrera, S, Guild, E & Hernanz, N 2013, 'The Triangular Relationship between Fundamental Rights, Democracy and the Rule of Law in the EU, Towards an EU Copenhagen Mechanism', Study, CEPS, Brussels.

respected. But in the last stage, even international legal rules are not respected by the governing power any longer.[603] Thus, a rule of law backslide cannot be turned back by legal tools alone, therefore it should be prevented whenever possible. The EU has a vital interest and an unequivocal obligation to protect and enforce the values enshrined in Article 2 TEU, out of which democracy and the rule of law, including free and fair elections, stand out. It is therefore obliged to stand up against abuses whether procedural (scenario C) or substantive (scenario D).

All measures that empower nation states can be used to reach opposite goals in the hands of captured or authoritarian states. Whether it is possible to draft effective measures that cannot be used for harm, is a conundrum yet to be solved. Softer measures carry less risk, but also induce lesser positive changes – similar to medications. Nevertheless, **previous studies have recommended an early warning system, which can at least signal when a state gets into this trouble.**[604]

6.2.2 Civic education for a democratic Europe

The EU is more than just an economic union, but its further evolution is faltering. Strengthening **European identity** would be necessary, which could be advanced by a common European media service. This does not have to be a 'public service' in the traditional sense, but sponsorship of, for example, investigative journalism, is recommended and so is its supranational European perspective in both news and opinion articles. Beyond these, we also recommend social goals that can be realised partly by the media, and partly by other social programmes.

While cultural differences should be cherished, the **perceptions of state and political participation, human rights and values should be more commonly shared** in the EU area, also at the level of the population. We recommend social programmes **to increase and promote the common understanding of what the EU values are**.

Building resilience against empty populism requires additional awareness from policy makers. It is advisable to create a **package on how to** improve the relationship between citizens and the state or decision makers, primarily through **good governance, transparency and integrity,** as well as social dialogue, addressing the realistic needs of society. Research shows that for false information to be challenged effectively within the human brain it needs to be replaced with an **alternative narrative.**[605]

To create an educated citizenry, a flowing process of communicating research should be developed, supported and actively maintained. This should be regarded as a key public service function that can create a solid basis in society for **knowledge and 'trust in science'**. A responsible EU institution should be designated, for example the European Research Council, in cooperation with a supranational EU media service.

Liberal democracy has been regarded as solely based on reason, rather than emotional persuasion. But the human species is not limited to reason: emotions play a great role in decision-making, as so vigorously proven at

[603] 'It was observed that in some Signatory States the Court's judgments are surrounded by bitter political debates, their execution is boycotted, and "certain political leaders seek to discredit the Court and undermine its authority'. Recommendation 2110 (2017), 29 June 2017, The implementation of judgments of the European Court of Human Rights, at 7. See also: Michalopoulos, S: Orban attacks the European Court of Human Rights. Euractiv. 30 March, 2017.

[604] Bárd, Petra, Sergio Carrera, Elspeth Guild and Dimitry Kochenov. With thematic contribution by Wim Marneffe. European Parliament, European Parliamentary Research Service, 2016. An EU mechanism on Democracy, the Rule of Law and Fundamental Rights. http://www.europarl.europa.eu/EPRS/EPRS_STUD_579328_AnnexII_CEPS_EU_Scoreboard_12April.pdf Also published as CEPS Paper in Liberty and Security No. 91. CEPS. Brussels

Petra Bárd, Judit Bayer, A comparative analysis of media freedom and pluralism in the EU Member States, research paper for the Policy Department C: Citizens' Rights And Constitutional Affairs. http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571376/IPOL_STU(2016)571376_EN.pdf

[605] Wardle, Claire - Hossein Derakhshan: Information Disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe report DGI(2017)09. at 78.

each election. Democratic values, like **people's sovereignty, people's empowerment, human rights and faith in science** could indeed be advocated passionately. Anyone who feels that science and passion are not compatible should only remember the late Stephen Hawking (1942-2018), who gave so much to society by publishing science in forms accessible to laypersons.[606]

6.3 Elections and political campaigns

Although disinformation and propaganda are not limited to political and public issues, these are the areas where they have the most devastating effects on societies and democracies. Public communication deeply determines societies, hence democracy. Therefore, **public issue advertising should also be subject to the same regulation as political advertising, in order to cover potential loopholes for social media.** The general rules on labelling and dissemination of political and public issue advertising should be maintained continuously even outside the campaign period. Stricter rules may be introduced temporarily during periods of election campaigns or referenda.

The rules in the Member States are far from harmonised; a few of them refer to campaign silence alone, and many apply only to offline campaign activity. Poland prohibits the publishing of untrue information, and Portugal has strong restrictions on finances and campaign rules. Sweden has set up a psychological defence authority, and many other countries have initiatives for new regulations.

Micro-targeting voters with political advertisements allows a political party to target individuals with tailored messages, to persuade different types of voters with different messages. While this can be useful to those parties that have the resources to exploit this advantage, the practice fragments the public discourse, the shared information basis which would be indispensable to the formation of public opinion. **The harm of micro-targeting is partly caused to those who receive it and partly to those who do not and thus have no information of what content their fellow citizens are exposed to.** Beyond this, micro-targeting on political issues is often **based on sensitive data. We recommend some regulation of micro-targeting for the purposes of protecting democratic public discourse, the fairness of elections and protection of personal data.** As a minimum, micro-targeting should be based on explicit consent and should contain a reference to the legal basis, the possibility to unsubscribe, etc.[607] Even a full prohibition of political micro-targeting could be considered, or requiring more transparency from political parties, notably to disclose the amounts spent on online political micro-targeting.[608]

6.3.1 Ethical guidelines

Political campaigns are overwhelmingly passionate and critical, and criticism should not be stifled. Our proposals consciously do not relate to the *content* of political and public issue advertising. Content evaluation could lead to censorship and continual legal disputes; therefore, our regulatory proposals focus on the actors, the methods and the effect.

Still, self-regulation of political parties should be encouraged **and the principles of an ethical political campaign formulated** and spread. **The principles should focus on a fair, transparent, constructive campaign strategy, the guidelines for which are below.** While soft guidelines are not always sufficiently powerful, in the present situation it appears **necessary to set down basic principles and make them part of civic and media literacy.**

[606] In his last interview he warned against artificial intelligence. Stephen Hawking warns artificial intelligence could end mankind. 2014. https://www.bbc.com/news/technology-30290540

[607] Article 7. E-Commerce Directive (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)

[608] See more in: See: Zuiderveen Borgesius, F.J., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., Bodo, B. and de Vreese, C., 2018. Online Political Microtargeting: Promises and Threats for Democracy. *Utrecht Law Review*, 14(1), pp.82–96. DOI: http://doi.org/10.18352/ulr.420

6.3.2 Campaign financing

Campaign financing has long been a pressing, overlooked issue in democracies, and the problem just has grown more severe. First, the diverse horizontal media environment fostered by the proliferation of content and social media makes the follow-up of campaign communication a complex task. Second, political processes are watched with cynicism by many voters who think that politics is influenced by strong economic actors rather than individual voters. This belief is a fertile ground for conspiracy theories and needs to be addressed. For these reasons, the structure of campaign financing needs to be substantially reconstructed.[609]

While currently this is a Member State competence, a common EU policy could be developed in this regard, because it effects the status of democracy within the EU. The competence of the European Public Prosecutor to investigate criminal offences relating to the funding of European political parties is seen as a promising instrument.

One approach to increasing trust in the democratic process is promoting the transparency of campaign financing, more specifically publishing the details of spending in a searchable database, easily accessible to the public (for example, like OpenSecrets.org: Cost of Election).[610]

6.4 A complex media policy

New media technology has substantially altered the dynamics of public discourse. Citizens have direct access to one another's views and gather into groups; underprivileged classes of society who were much less visible before have gained a voice. The open array of public discourse is exploited by political opportunists, who weaponise users to share and amplify populistic political propaganda, and occupy the public sphere through automated dissemination methods.

The dramatic changes in the media market have led to a crisis in professional journalism and quality media. Yet, even a strong media system could not resist the organised, disguised, concerted manipulation that happens with industrial professionalism. **When a consistent system of websites, various social media profiles and media channels complement each other, use cross-references, feature a professionally designed and built campaign, including manipulation and falsified pictures, articles or even videos, and especially when these are backed by an influential public figure (e.g. a head of state) then media policy is powerless to withstand such pressure.** These fights must be combatted with stronger instruments like criminal law and international law (discussed below).

Nevertheless, in the long term, improving the quality of the media system and the **media literacy** of the audience would pay back its cost. For example, **disseminating false messages through private messaging services like WhatsApp cannot be supervised by state regulation**. This ought to be handled through the softer tools of increasing **user awareness**, by inducing users to double-check information before acting on it and educating users on the ethical principles of publishing and sharing. For example, since the historical example of the *War of the Worlds* radio programme, which generated a mass panic of people afraid of aliens, all users know now that radio programmes could also be fictitious. Today, given that **all internet users are potential journalists,** media literacy should include not only how to interpret media but also **how to publish content ethically.**

For these reasons, we recommend investing in the development of a high-quality and diverse media environment where fact-checking services, credibility indices and a supranational, quality media service help users to navigate in the content jungle. Ideally, none of these services should be performed by state authorities or public institutions. However, **subsidising and encouraging the launch** of such activities may be indispensable for success.

[609] One noteworthy initiative in this respect is the Mayday movement in the US. https://mayday.us

[610] OpenSecrets.org, Center for Responsive Politics, Cost of Election. www.opensecrets.org/overview/cost.php

6.4.1 Diversity, pluralism and concentration

The present media environment theoretically offers greater diversity of content than ever in human history (see more in section 2.1). Technology also offers tools to make a diverse selection of content available for each individual user. Algorithms are meant to increase user satisfaction, but in reality, they are optimised to increase user addiction and maximise profit for advertisers. They should be built to increase user exposure to diverse content. Besides diversity, access to truthful content should be made easy: pillars of trust need to be rebuilt, with the help of journalists and civil society. Increasing diversity is a shared responsibility of platform providers, media outlets, policy makers and users themselves (by making their own choices).

Public service broadcasting has a history of being such a pillar of trust. At the same time, **captured states and authoritarian states often exploit public service broadcasting for propaganda purposes.** The EU already has in place a Commission Communication on State Aid to Public Service Broadcasting, which should be applied.[611] **As public service media operates from public resources, it is crucial that it serves society as a whole** and not just the ruling government's political interests. Broadcasting propaganda and disinformation through any of these public channels severely **violates public service ethics.**

Diversity stands on several pillars, including ownership concentration, subsidising (state aid), the political independence of public service media and journalistic ethics.[612] These can be approached with the existing rules on merger control, competition, state subsidies to media outlets (including but not limited to state aid to public service broadcasting) and self-regulation. On the demand side, exposure to diversity should be increased, which can be improved through algorithmic design and media literacy (both addressed below).

Research suggests that social media usage that involves participation in several networks reduces mass political polarisation and echo chambers.[613] This points out **the necessity of having competition in the social media field.** Given the nature of the service, users would benefit from alternative services of the same kind only if **interoperability** is ensured, i.e. if connections and communication is facilitated between the platforms to enable users to keep contact with other users who are members of other social media platforms – similar to the interoperability of telecom services.

If this is not possible under current technological standards, then a platform providers' dominant position should entail certain obligations and responsibilities beyond those of non-dominant parties. These special duties would apply only to incumbent companies, similar to ‘common carriers’, or to the **regulation of public broadcasters** within Europe:[614] to ensure diversity; to give preference to high-quality, reliable sources (which are attested by fact-checkers and credibility indices) and public service content; to ensure neutrality; and to respect higher levels of privacy standards, transparency of algorithms and flexibility of settings for the convenience of users. As set out by the UN Guiding Principles, the "scale and complexity of the means through which enterprises meet [their corporate] responsibility may vary" according to their size, sector and the severity of the enterprise's adverse human rights impacts.[615]

The possibility of **merger control** among such similar services as Instagram, Facebook, WhatsApp, Google, YouTube and Snapchat should be considered.

[611] Communication from the Commission on the application of State aid rules to public service broadcasting 2009/C 257/1

[612] See Petra Bárd, Judit Bayer, A comparative analysis of media freedom and pluralism in the EU Member States, research paper for the Policy Department C: Citizens' Rights And Constitutional Affairs. http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571376/IPOL_STU(2016)571376_EN.pdf.

[613] https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf, 27. Oldal

[614] See also: Caplan, Robyn and Danah Boyd: Who Controls the Public Sphere in an Era of Algorithms? Mediation, Automation, Power. 05.13.2016. https://datasociety.net/pubs/ap/MediationAutomationPower_2016.pdf at page 5.

[615] Guiding Principles on Business and Human Rights**,** Implementing the United Nations “Protect, Respect and Remedy” Framework. New York and Geneva, 2011. https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf. Article 14.

6.4.2 Transparency and neutrality

The current transparency efforts of the digital platforms with respect to online advertising and targeting are inadequate. The strong financial incentives must be balanced by the force of law: to oblige platforms to provide users with meaningful transparency about political and issue-based advertising they are subject to across different devices (such as desktop, mobile and wearable devices).

Social media should remain 'social', thus it should not become dominated by bots, AI, professional influencers or campaigners, etc. The capability to **play in the same pool with equal rights, but much better opportunities, is injurious to the rights of the private individual users.** For this reason, the activities of robots and AI should be restricted, and **professional, commercial actors should disclose their real identities (as discussed below).**

6.5 Privacy and data protection

Social media is a type of service that, by definition, works with masses of personal data. Using such services equals sharing personal information through the platform. This makes users especially vulnerable, because they expose private information they would share only with a few selected friends, or would not even share it but from which it could be concluded (inferred) from their likes, shares and other actions on the network (for example, how often they check in, whose profiles they watch and for how long). The business model of online publishing is predominantly building on monetising the personal data of internet users. This practice violates the fundamental rights to the protection of personal data, privacy and autonomy. It damages trust in technological solutions, which **in the long run could stifle the success of innovations**. With a view to maximising the benefits of future technologies for the economy and society, consumer trust must be regained. **This business model is not sustainable even from an economic perspective; therefore, it needs to change.**

The rules and principles of EU data protection law, namely the GDPR, prohibit the collection, use, transfer and so on of personal information without ensuring that such processing meets data protection principles, including principles of lawfulness, transparency and purpose limitation. Moreover, the ePrivacy Directive sets forth the confidentiality of electronic communications and of the data (content and metadata) conveyed and used in the context of electronic communications. It also has specific provisions regarding the use of cookies and other mechanisms used to store information or to gain access to information stored in the terminal equipment of the user. These principles are often compromised where personal data are used for the purposes of commercial and non-commercial advertising, including micro-targeting. The act of micro-targeting – without the knowledge and understanding of the targeted individual, and informed and freely given consent – violates human dignity and the right to freedom of (truthful) information, and it destroys public discourse. To ensure the protection of individuals against such practices, **political micro-targeting should be recognised as solely automated decision-making that produces significant effects on individuals under Article 22 of the GDPR**. Where it is based on collected, observed or inferred special categories of personal data, it can only be allowed based on the informed, explicit and freely given consent of the individual or where significant public interest based on EU or national law merits so.

In general, the robust *de jure* protections offered by the data protection laws will only be as effective as their de facto enforcement against both the demand and the supply side of advertising led by the national data protection authorities, whose independence and resources must be secured by the Member States and rigorously **monitored by the European Commission.**

Given the considerable civic powers of the social media platforms, their data protection-related obligations **should be encouraged to go beyond the minimum requirements** enshrined in the EU and the national data protection laws. For example, to facilitate transparency and accountability obligations, they should **maintain a searchable repository of active and historical political and issue-based advertising targeting persons in the EU with detailed information about the criteria of targeting, buyers,** etc. Platform design choices should strictly adhere to the **data protection by design and by default obligations** enshrined in the GDPR, with a particular emphasis on the technical and organisational arrangements to ensure the **freely given consent** of the data subjects. Given the nature of processing performed by the social media platforms, **their data protection policies and data protection impact assessments should be particularly closely scrutinised by the data protection authorities.**

Tracking-based online advertising should be de-incentivised by adopting, at the soonest, a robust ePrivacy Regulation, which outlaws 'tracking walls' and includes other safeguards as advocated by the regulators in the field, e.g. the EDPS.

Personal data are often considered to be a sort of currency that keeps business and innovation moving. These are important advances of human society, but they ought to serve human beings, and not the reverse. Another argumentation is that collecting and combining data seek to provide better services to the users. Yet such 'better services' in the context of public discourse means **making choices in place of the users**, thereby limiting user autonomy.

The new disruptive technology services will be even more accessible, more automated, used more frequently and less recognisably. They will be based on using personal data. To avoid a frightening dystopia, the **approach to using personal data must be re-worked considerably**. Similar to doctors or lawyers, handling personal data should not entitle their processors to prey upon them.

6.6 Scope of responsibility of social media providers

The scope of responsibility of social media providers has been of particular interest and importance during policy debates. Here we summarise the cross-cutting issues that should be the responsibility of social media (and those that should not). In chapter 7 on recommendations, they are organised following a different logic, yet we also present them below to provide an overview.

While a **consensus appears to be formed that platform providers should not bear responsibility for third-party content,** their roles – what should be expected from them – should be defined, and they should be made responsible for those functions that they can best control. As a cross-cutting principle of their duties and responsibilities, platform providers should respect human rights and consumer protection rules. They should avoid infringing human rights and address the adverse human rights impacts with which they are involved.[616]

Social media providers should not have this responsibility:

- deciding on the legal or illegal quality of content, especially that of a political nature, without prejudice to making their best efforts to follow their own guidelines on the removal of manifestly illegal content, like child pornography or brutality.

Social media providers should have these responsibilities:

- ensuring full compliance with and enforcement, in relation to their partners, of the rules on privacy protection, including the GDPR and Directive 2002/58/EC (ePrivacy Directive), along with the recommended extensions of protection, namely the prohibition against profiling and targeting for political purposes and the prohibition against building on such inferred data;
- maintaining an ideologically neutral service – their algorithms should not favour any political or public interest, nor should they create any discrimination between users based on gender, race, ethnical origin or religious, philosophical or political belief;
- enforcing the envisaged rule on user verification, with special emphasis on protecting the privacy of the registered users and on the identification of political and public issue organisations;

[616] Guiding Principles on Business and Human Rights**,** Implementing the United Nations "Protect, Respect and Remedy" Framework. New York and Geneva, 2011. https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf

- enforcing the envisaged restrictions on micro-targeting and the use of AI methods for disseminating messages by political actors and public issue organisations;
- cooperating with transparency requirements regarding political and public issue advertising;
- formulating the principles of their content selection algorithms in easily understandable language and constructing options for the users to be able to oversee and approve or reject algorithmic principles, ensuring that such rejected algorithms are not applicable in relation to these user accounts;
- making best efforts to offer a diversity of content options to the users; and
- to provide safeguards against mistaken content removal, and setting these out in their policies.

6.7 Criminal rule to prohibit 'aggressive informational practices'

In some cases, organised campaigns featuring disinformation and propaganda are part of a dark game.[617] Although in many cases the audience and the platform providers can be blamed for amplifying this, the most harmful events are well-organised and well-financed. Several of the actions can be traced back to a few companies, such as Cambridge Analytica or the Internet Research Agency in Russia. Such activity would in most cases be financed from illegal resources, violating the rules of financing political parties. **States must not endure criminals systematically misusing intermediaries' services and their infrastructure. Such actions should be fought with the ultimate tools of the law, even criminal law; in cases where they originate in a foreign state, international law should be taken into consideration.**

Spreading disinformation is not a crime currently on the basis of its falsity alone (see more in section 3.2). Although finding conclusive evidence against the perpetrators can be a heavy duty for the investigative authorities, the significance of a criminal prohibition should not be discounted. As mentioned in section 1.1, disinformation and propaganda are not always literally false information, but can consist of misleading information, distractive or inciting content, or simply try to cause division. Therefore, **the core element of the prohibited behaviour should be the aggressive informational practice rather than the falsity of content.**

It is crucial that any envisaged criminal rule should build heavily upon the method rather than the content – the manipulative techniques, the automated dissemination methods, the misuse of personal data and the violation of political party financing. It should at least include the following criteria, pertaining to whoever disseminates to the public, or gives a mandate for such activity, information:

- that is false, falsified or misleading, or incites opinion;
- with the intention to mislead, polarise or destabilise society, or a substantial part of it;
- with the direct or indirect purpose of gaining political or geopolitical power, or financial gains;
- from disguised sources or identities, or using micro-targeting or using artificial intelligence to boost dissemination of such content; and
- with the potential to have a negative impact on social cohesion, public order or peace.

There could also be further criteria or aggravating circumstances:

- that it is organised perpetration of the same; and
- if it realises an impact.

It should be remembered that any prohibition that criminalises political expression could be misused by captured or authoritarian states, against political opposition and to stifle freedom of expression. Already an accusation of the crime could be used to threaten political opposition and have a negative effect on public discourse.

[617] Bradshaw, Samantha and Philip N. Howard: Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation. Computational Propaganda Project, University of Oxford. 2018. at page 7.

On the one hand, other recommended rules might also be sufficient to reach the desired goal **of preventing substantial damage to democratic societies without the unwanted chilling effect of a criminal rule on speech.** On the other hand, if the perpetrators of a disinformation and propaganda campaign have been found, they cannot be made accountable under the current legal rules, because spreading disinformation and propaganda are not illegal. These aspects must be contemplated before deciding on this policy recommendation.

7. RECOMMENDATIONS

A major part of democratic resilience building is to apply and further develop tools of militant democracy. This includes protecting the institutional checks and balances as well as election rules, and strengthening the protection of human rights such as data protection. It also extends to improving media pluralism, promoting civic education (including media literacy) and further programmes to safeguard the rule of law within the EU. Some of these aspects fall outside the scope of this study. Some of the authors have recommended measures in other papers.[618] All recommendations are directed towards one common objective: to safeguard democracy, the rule of law and respect for human rights. This should provide the basis for the competences of EU legislation. Given the cross-border dimension of the problems addressed, the recommended measures need to be adopted at the EU level in order to achieve the objectives. We have divided the recommended measures and policies into two main branches: (i) strengthening democratic resilience and (ii) media policy.

7.1 Strengthening democratic resilience

7.1.1 Imminent actions relating to the fairness of EP elections

1) The **Commission** should initiate at the OSCE to monitor the upcoming EP elections in Member States; in the longer term, it should check the available options for creating a European election observation body similar to the EEAS election observation service.

2) For the future, the EU should consider building a specific EU institutional capacity in the form of an EU electoral authority with powers to monitor and undertake field visits to Member States preceding the EP elections, also with the power to supervise political campaigns.

3) The European Court of Auditors and OLAF should pursue the investigation of campaign finances, including sponsorship of social media advertisements.

7.1.2 Regulation of political and public issue advertising

As a principle, all existing **rules on political and public issue advertising should be extended to any publication method**, with special regard to online publishing, including social media. Member States' rules on political and public issue advertising should **be harmonised by an EU directive.** All traditional principles should be applied, including the principle of separation (all advertisements should be clearly distinguishable from the editorial content and news), the principle of identification (all sponsored content should be identified as such) and so forth, and additional principles should be added (outlined below). These contain obligations for platform providers, yet other media providers that publish political advertisements may be affected.

1) **Ensuring transparency**
   - The digital platforms should proactively verify the identity of the advertisers (buyers and their clients) and allow users to access this information.
   - Each digital platform should maintain a searchable repository of active and historical political and issue-based advertising targeting persons in the EU. The repository should include information about the ad buyers (and their clients if applicable), amounts spent, targeting criteria (demographic, location, interests and other), the ad's reach (impressions), the period when it was active and other relevant information. Users should be allowed to filter information based on all these criteria. Each user should have access to an individual repository providing information about what political and issue-based ads that user was or is being targeted with, including the information listed above.

[618] Petra Bárd, Sergio Carrera, Elspeth Guild, Dimitry Kochenov (2016) An EU mechanism on Democracy, the Rule of Law and Fundamental Rights, Brussels: Center for European Policy Studies (CEPS). See also: Petra Bárd, Judit Bayer, A comparative analysis of media freedom and pluralism in the EU Member States. Study for the LIBE Committee, PE 571.376 EN, 2016, http://www.europarl.europa.eu/supporting-analyses.

- PR companies and platforms should be obliged to keep their contracts with political parties and candidates on file for supervision purposes to ensure election campaign transparency.
- Users regularly reaching large audiences with public issue content (for example, political parties, politicians, NGOs, communication agencies and other influencers) should be subject to closer scrutiny (e.g. verification of identity) and their 'influencer status' should be signalled on their profile.
- Each message conveyed to users should clearly display in its subject its nature of political advertisement.

2) **Dissemination methods**

- Bots, automated accounts and artificial intelligence should be ruled out of publishing and disseminating political and public issue advertisements.
- Political micro-targeting should be recognised as a decision based solely on automated processing, including profiling, which significantly affects data subjects and therefore Article 22 of the GDPR should fully apply to this practice.

3) **Campaign financing**

The EU should harmonise the rules on the financing of political campaigns for elections and referenda. Beyond the EP elections, these rules of the Member States deeply affect democracy within the EU, through the fairness of elections.

- A substantial reform towards a cleaner and more transparent structure for campaign financing should be outlined (its exact rules are beyond the scope of this study).
- Rules on party expenditures on political campaigns should be monitored more rigorously, along harmonised guidelines. Campaign expenses should be limited to an objective amount (for example, 15 times the monthly national minimum wage per candidate). Investigative authorities should be furnished with the appropriate power to carry out examinations to reveal connections between political parties or state budgets.[619]

4) **Self-regulation and civic information**

- Member States should create harmonised rules to induce political parties and other actors who take part in political campaigns to self-regulate – to lay down codes of conduct for an ethical and fair campaign, and to disseminate public information on these rules.
- The guidelines for a fair, transparent and constructive campaign strategy should include at least these attributes:
  - based on true information and real social needs;
  - not based on fear, social tensions or instincts;
  - avoids inciting hatred or hostility;
  - avoids ad hominem arguments (character assassination); and
  - is inclusive of all society, rather than targeted.

5) **Further research**

Aggressive informational practices should be studied for policy purposes.

- The EU should allocate sufficient resources to study aggressive informational practices deployed in non-commercial settings (e.g. civic events, elections and referenda) in order to better understand their real impact on individual views and behaviour. The research should appreciate and take into account a diversity of national contexts and practices across the EU.

[619] Like all governmental powers, this can also backfire in the hands of a captured, or authoritarian state. Authentic NGOs operate in a transparent way already.

- The feasibility of EU regulatory action outlawing aggressive informational practices in the non-commercial context should be explored. In particular, it should be determined whether the safeguards provided for consumers in the Unfair Commercial Practices Directive and other relevant consumer legislation should not be extended to other, civic domains.
- New technologies that may enable a proliferation of aggressive informational practices should be subject to ex ante fundamental rights impact assessments as an extension of the data protection impact assessment already mandated by the GDPR in high-risk data processing cases.

7.1.3 Data protection and privacy: Enforcement and development of the legal regulation

1) Compliance with the GDPR and current ePrivacy Directive, with particular regard to social media should urgently be enforced by the data protection authorities. The independence and resources of data protection authorities must be secured by the Member States and rigorously monitored by the European Commission.

2) In particular, social media platforms should properly comply with their accountability obligations and other obligations set forth by the GDPR and hence implement all the appropriate technical and organisational measures to ensure and be able to demonstrate compliance, but also carry out data protection impact assessments and prior notifications to the data protection authorities in accordance with the GDPR.

3) **Tracking-based online advertising should be de-incentivised by adopting a robust ePrivacy Regulation**, which outlaws 'tracking walls' (conditioning access to websites upon the individual being forced to 'consent') and includes other safeguards as advocated by the regulators in the field, e.g. the EDPS.

4) Platform providers should be obliged to prevent data mining on their platform and to prevent unlawful targeted advertising, especially based on sensitive information. Adequately trained staff should supervise the processing of lawfully collected personal data, with an obligation to prevent misuse.

5) **The ePrivacy Regulation should include a reference to 'platform providers', and define social media as a subcategory of them.**

6) **The ePrivacy Regulation should clarify that users can give their consent by virtue of their browser settings, and online websites should provide an interface to be able to read such electronic declarations. Interpretation of the GDPR allows this and it could also be clarified through guidelines.**

7.1.4 Civic education on democratic values, fostering dialogue and community building

The competent EU institution should identify territorial regions within the EU where it is necessary (based on social value surveys) to **launch local programmes, for example utilising travelling buses**, to include those parts of society otherwise difficult to reach: the adult and elderly generations, minorities and disadvantaged people. The goal would be to increase knowledge of the democratic values and institutions, including the election process and specificities of the political campaigns. The programmes should be organised in cooperation with local institutions, NGOs or social movements, but **with quality control maintained by the relevant EU institution.**

1) **The EU should provide support for complementing education programmes in all Member States:**
   - to cover European values of democracy and human rights **at all levels of education**;
   - to ensure appropriate education of teaching staff; and
   - to provide media literacy education – which has been set out in the proposed text of the AVMS Directive but its implementation needs to be supervised. Education on **ethical principles for publishing and sharing information needs to be included** in the programmes.

2) **The Commission should initiate and support supranational European programmes[620] for social cohesion and social inclusion, and sensitisation to disinformation and propaganda:**

[620] Examples can be: the Media Pluralism Monitor, http://cmpf.eui.eu/media-pluralism-monitor/, the European Centre for Press and Media Freedom, https://ecpmf.eu.

- to reduce polarisation and hostility and to mediate between the majority and social minorities as well as among minorities if necessary;
- to promote the idea of European identity (among others, through creating common EU narratives and the foundation of a common European identity, culture and values);
- to increase sensitisation to disinformation and propaganda and to impart information on fair political campaign principles before elections and referenda; and
- to close the socioeconomic gaps in media literacy education, addressing underprivileged parts of societies, minorities, and the adult and elderly generations, if such gaps exist, especially in the post-communist Member States.

7.1.5 Mainstreaming science in policy-making

1) Scientific efforts should be applied in planning policies. Considering that marketing and political campaigns are designed with the help of psychological science, the remedy against these could also be elaborated by science.

2) An interdisciplinary approach should be used for policy-making and cooperation, with the inclusion of software engineers, ethical hackers, psychologists, social researchers, etc.

3) Public resources should be devoted to research on how to armour citizens against manipulation.

4) **Communicating about science** should take place in a more systematic and generally accessible way, in cooperation with the media.

7.2 Media policy

7.2.1 Creating pillars of trust in the media

1) The European Commission should initiate and support supranational programmes to **organise platforms for and finance investigative journalism**, including solution journalism and data-driven journalism at the level of Member States and also the EU. It should encourage cross-border collaborations that bring together participants from more countries and editorial offices. While this branch performs precious public services for democratic states, it is often unwelcome because of the criticism. **The new EU budget period should extend** the availability of existing resources and allocate dedicated resources to these programmes.

2) Similar programmes should encourage and support independent journalistic organisations that provide **credibility indices** to create easily readable, frequently updated databases. Granting a ‘quality insurance label’ to credible news outlets could promptly inform users about the trustworthiness of a source.

3) Journalistic associations should be encouraged or supported to participate in fact-checking; fact-checkers are recommended to cooperate with social media sites, using a technological short-cut to stop the disinformation from spreading.

4) The European Commission should **supervise the operation of public service media providers**, whether they fulfil the criteria of prudent management and **task-based financing, and if their service fulfils the expectations of fact-based, fair and ethical journalism**. In Member States where it serves the interests of a captured state, the Commission **should not grant exemption under 86(2) of the Treaty**.[621]

[621] Art. 106. para. 2. and 107 para. 1 TFEU., See also the Communication from the Commission on the application of State aid rules to public service broadcasting 2009/C 257/1

5) The European Commission should initiate and support the creation of an **international, European, high-quality media service,**[622] which would provide EU-focused general news and opinion pieces in an engaging and diverse manner and is also disseminated through platform services.

6) The existing instruments against **ownership concentration** and **illegal state aid** should be applied to increase diversity in the media landscape. Regular check-ups of media freedom and pluralism within Member States as part of the rule of law framework should be carried out.[623]

7.2.2 Obligations of platform providers, including social media providers

The first set of recommendations should be legally regulated (*hard instruments*); the second part should be left for other instruments, such as self-regulation (*soft instruments*).

The set of obligations is not entirely independent from the market position of the media providers. While some of the minimum obligations should be required from all, the higher the market share, the greater the obligations that should be expected; these negotiable instruments are listed among the soft instruments.

7.2.2.1 Hard instruments

Platform providers have become dominant actors in business, entertainment and other sectors, but they are still not recognised by legal regulation as a specific type of service provider. Any regulation on the responsibilities and rights of platform providers, and among them social media providers, would need to rely on a legal definition of who the subjects of regulation are.

1) **'Platform providers'** should be **defined in the E-Commerce Directive** as a separate category of service providers, which provide the technological possibility for third parties to communicate their content to the public. Their activity includes organising and facilitating the transmission of content with the help of their algorithmic selection methods. These definitions should be added to the proposed counter-terrorism Regulation and the ePrivacy Regulation.

2) Platform providers should not be liable for third-party content, nor obliged to monitor third-party content. The proposed regulation against terrorist content should not derogate from the principle of no-monitoring.

3) Platform providers should be responsible for respecting human rights standards, expediently administering their platforms, protecting the personal data of their users and refraining from illegal discrimination between their users or their content (discussed below in detail). The E-Commerce Directive should contain a reference to their other obligations relating to data protection and other laws.

4) In dealing with illegal or objectionable content, the notice-and-notice procedure is recommended as being more respectful of freedom of expression than the notice-and-takedown procedure.

A specific regulation, a directive or amendment package should provide for obligations of platform providers as set out below.

1) **Data protection and privacy**

Platform providers should be responsible for proactively protecting the personal data of their users and enforcing data protection principles on their platforms:

- preventing hacking and data leaks;

[622] Some regard the Euronews as a common European channel; however, 53% of its shares is owned by Egyptian Media Globe Networks, 25 % by American NBC, and 22% only by European public service broadcasters, and it is not limited to the EU. Euractiv comes closest to such a function.

[623] See more inPetra Bárd, Judit Bayer, A comparative analysis of media freedom and pluralism in the EU Member States, research paper for the Policy Department C: Citizens' Rights And Constitutional Affairs. http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571376/IPOL_STU(2016)571376_EN.pdf.

- enabling targeted advertising on their platforms only after user opt-in;
- monitoring and preventing illegal activity that violates the protection of personal data on their platforms;
- informing users about which of their data are used for content selection or micro-targeting, and offering them the chance to exclude some personal data from this process; and
- conducting data protection impact assessments and respecting the personal data breach rules laid down in the GDPR.

2) **Neutrality**

Platform providers should be obliged to maintain ideologically neutral services:

- ensuring their algorithms do not systematically favour any political, ideological or religious opinion, or give preference to content that is their own or by an affiliated company; and
- avoid discriminating among users or their content based on protected characteristics such as race, gender or political opinion.

3) **Administering platforms**

- Platform providers should ensure – by technological means of supervision or verification – that the accounts are registered by human individuals rather than artificial intelligence or bots.
- Deleting fake accounts should not be regarded as a virtue but as an obligation.
- Bots, virtual personalities, trolls and influencers[624] should be identified as such.
- Users regularly reaching large audiences with public issue content should be regarded as 'influencers' (political parties, NGOs, communication agencies and other influencers) and, along with all organisations and commercial actors, should register only after identifying and disclosing their real identities online.

4) **Algorithms**

- Content selection algorithms should include the principle of diversity (offering different views).
- Platform providers should inform their users about the content-selecting principles of their algorithms.
- Users must have options on which principles they would like to use or reject, after receiving easily accessible information, using tools as simple as icons. One option should be to prioritise content that is found to be trustworthy by independent news organisations.
- Changes and experimenting with new algorithms should be transparent, providing easily accessible information to users.

5) **Advertisements**

- Platform providers should require identification from advertisers and enable users to access this information.
- Platform providers should identify advertisements and sponsored information on their platforms as such.
- Platform providers should maintain a searchable repository of political and issue-based advertising targeting persons in the EU (see also above under the title 'Elections').

[624] The word "influencer" is used here to political parties, active politicians, governments and public authorities, NGO's and communication agencies and the new brand of influencers themselves.

- Platform providers should grant their users access to an individual repository providing information about what political and issue-based ads they are or were targeted with.
- Platform providers should be obliged to keep their contracts with political parties and candidates on file for supervision purposes to ensure election campaign transparency.

7.2.2.2 Soft instruments

1) Platform providers should offer their users the option to choose the required level of diversity, and the extent to which they wish to see content that is different from what would normally be recommended to them by the default algorithms. The user should be able to make this choice easily and repeatedly.
2) Platform providers should include in their offered newsfeed information representing diverse views from what is otherwise recommended (e.g. 'different opinions', 'related news', 'other than this', as is already done by some platform providers).
3) Platform providers should give priority to news with certified credibility or even public service content.[625]
4) The notice-and-takedown procedure could be substituted with the notice-and-notice procedure, with the exception of absolutely clear cases for removal.[626] Thus, in accordance with the statement of Commissioner Věra Jourova,[627] platform providers **should not remove content in cases of ambiguity.**

7.2.3 The limits of self-regulation

1) On supervision, any self-regulation efforts by the digital platforms (e.g. content moderation or ad transparency initiatives) should be subject to **external scrutiny and impact assessment** in order to determine their efficiency and compliance with fundamental rights. The EU should allocate sufficient financial resources to the members of the research community in order for them to undertake such assessments in a **rigorous and timely** manner.
2) Self-regulation in the area of online advertising, and in particular political and issue-based advertising, cannot and **should not replace enforceable regulation**, including approved and enforceable codes of conduct.
3) Media self-regulation and journalistic ethical codes are to be complemented with best practice recommendations on how to avoid amplification of disinformation or manipulative content. However, the majority of such content is not distributed by quality journalism, which would adhere to self-regulation anyway.

[625] This can backfire in the case of a captured, or authoritarian state.

[626] See Petra Bárd, Judit Bayer, A comparative analysis of media freedom and pluralism in the EU Member States, research paper for the Policy Department C: Citizens' Rights And Constitutional Affairs. OLDALSZÁM. http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571376/IPOL_STU(2016)571376_EN.pdf.

[627] Commissioner Jourova emphasized in a public speech at CEPS, on 12. October 2018, Brussels, that platform providers, on the basis of the Code of Conduct countering illegal hate speech online, should not remove content in case of ambiguity.

Table 13: Recommended actions categorised by the type needed – application of existing instruments, amending existing instruments or new instruments

Type of instrument recommended	Recommended action
Existing instruments should be applied	1) The European Court of Auditors and OLAF should pursue the investigation of campaign finances, including sponsorship of social media advertisements. 2) Compliance with the GDPR and current ePrivacy Directive, with particular regard to social media should be urgently enforced by the data protection authorities. The independence and resources of data protection authorities must be secured by the Member States and rigorously monitored by the European Commission. 3) The European Commission should supervise the operation of public service media providers, whether they fulfil the criteria of prudent management and task-based financing, and if their service fulfils the expectations of fact-based, fair and ethical journalism. In Member States where it serves the interests of a captured state, the Commission should not grant exemption under Article 86(2) of the Treaty.[628] 4) The existing instruments against ownership concentration and illegal state aid should be applied to increase diversity in the media landscape. Regular check-ups of media freedom and pluralism within Member States as part of the rule of law framework should be carried out.[629]
Existing instruments should be amended	5) The Commission should empower the EEAS election observation service to monitor the upcoming EP elections in Member States, similar to third-country national elections. 6) The digital platforms should proactively verify the identity of the advertisers (buyers and their clients) and allow users to access this information. 7) It should be determined whether the safeguards provided for consumers in the Unfair Commercial Practices Directive and other relevant consumer legislation should not be extended to other, civic domains. 8) Tracking-based online advertising should be de-incentivised by adopting a robust ePrivacy Regulation, which outlaws 'tracking walls' (conditioning access to websites upon the individual being forced to 'consent') and includes other safeguards as advocated by the regulators in the field, e.g. the EDPS.

[628] Art. 106. para. 2. and 107 para. 1 TFEU; see also the Communication from the Commission on the application of State aid rules to public service broadcasting 2009/C 257/1.

[629] See Petra Bárd, Judit Bayer, A comparative analysis of media freedom and pluralism in the EU Member States (with Petra Bárd), research paper for the Policy Department C: Citizens' Rights And Constitutional Affairs. http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571376/IPOL_STU(2016)571376_EN.pdf.

	9) Platform providers should be obliged to prevent data mining on their platform and to prevent unlawful targeted advertising, especially based on sensitive information. Adequately trained staff should supervise the processing of lawfully collected personal data, with an obligation to prevent misuse. 10) The ePrivacy Regulation should include a reference to 'platform providers', and define social media as a subcategory of them. 11) The ePrivacy Regulation should clarify that users can give their consent by virtue of their browser settings, and online websites should provide an interface to be able to read such electronic declarations. Interpretation of the GDPR allows this and it could also be clarified through guidelines. 12) The EU should provide support for complementing education programmes in all Member States with media literacy education – which has been set out in the proposed text of the AVMS Directive but its implementation needs to be supervised. 13) An interdisciplinary approach should be used for policy-making and cooperation, with the inclusion of software engineers, ethical hackers, psychologists, social researchers, etc. 14) Communicating about science should take place in a more systematic and generally accessible way, in cooperation with the media. 15) 'Platform providers' should be defined in the e-Commerce Directive as a separate category of service providers, which provide the technological possibility for third parties to communicate their content to the public. Their activity includes organising and facilitating the transmission of content with the help of their algorithmic selection methods. These definitions should be added to the proposed counter-terrorism Regulation and the ePrivacy Regulation. 16) Platform providers should be responsible for respecting human rights standards, expediently administering their platforms, protecting the personal data of their users and refraining from illegal discrimination among their users or their content (discussed below in detail). The E-Commerce Directive should contain a reference to their other obligations relating to data protection and other laws. 17) The proposed regulation against terrorist content should not derogate from the principle of no-monitoring. 18) In dealing with illegal or objectionable content, the notice-and-notice procedure is recommended as being more respectful of freedom of expression than the notice-and-takedown procedure. The notice-and-takedown procedure could be substituted with the notice-and-notice procedure with the exception of absolutely clear cases for removal.
New instruments are needed	19) For the future, the EU should consider building a specific EU institutional capacity in the form of an EU electoral authority with powers to monitor and undertake field visits to Member States preceding the EP elections, also with the power to supervise political campaigns.

20) Each digital platform should maintain a searchable repository of active and historical political and issue-based advertising targeting persons in the EU. The repository should include information about the ad buyers (and their clients if applicable), amounts spent, targeting criteria (demographic, location, interests and other), the ad's reach (impressions), the period when it was active and other relevant information. The users should be allowed to filter information based on all these criteria. Each user should have access to an individual repository providing information about what political and issue-based ads the user was or is being targeted with, including the information listed above.

21) PR companies and platforms should be obliged to keep their contracts with political parties and candidates on file for supervision purposes to ensure election campaign transparency.

22) Users regularly reaching large audiences with public issue content (for example, political parties, politicians, NGOs, communication agencies and other influencers) should be subject to closer scrutiny (e.g. verification of identity), and their 'influencer status' be signalled on their profile.

23) A substantial reform towards a cleaner and more transparent structure for campaign financing should be outlined (the rules are beyond the scope of this study).

24) Rules on party expenditures on political campaigns should be monitored more rigorously, along harmonised guidelines. Investigative authorities should be furnished with the appropriate power to carry out examinations to reveal connections between political parties or state budgets.[630]

25) Member States should create harmonised rules to induce political parties and other actors who take part in political campaigns to self-regulate – to lay down codes of conduct for an ethical and fair campaign, and to disseminate public information on these rules.

26) Aggressive informational practices should be studied for policy purposes. The EU should allocate sufficient resources to study aggressive informational practices deployed in non-commercial settings (e.g. civic events, elections and referenda) in order to better understand their *real impact* on individual views and behaviour.

27) The feasibility of EU regulatory action outlawing aggressive informational practices in the non-commercial context should be explored.

28) New technologies that may enable a proliferation of aggressive informational practices should be subject to ex ante fundamental rights impact assessments as an extension of the data protection impact assessment already mandated by the GDPR in high-risk data processing cases.

[630] Like all governmental powers, this can also backfire in the hands of a captured, or authoritarian state. Authentic NGOs operate in a transparent way already.

29) The Commission should initiate and support supranational European programmes[631] for social cohesion and social inclusion, and sensitisation to disinformation and propaganda: (i) to reduce polarisation and hostility, and to mediate between the majority and social minorities as well as among minorities if necessary; (ii) to promote the idea of European identity (among others, through creating common EU narratives and the foundation of a common European identity, culture and values); (iii) to increase sensitisation to disinformation and propaganda and to impart information on fair political campaign principles before elections and referenda; and (iv) to close the socioeconomic gaps in media literacy education, addressing underprivileged parts of societies, minorities, and the adult and elderly generations, if such gaps exist, especially in the post-communist Member States.

30) The EU should provide support for complementing education programmes in all Member States on European values of democracy and human rights at all levels of education, with appropriate education of teaching staff.

31) Scientific efforts should be applied in planning policies. Considering that marketing and political campaigns are designed with the help of psychological science, the remedy against these could also be elaborated by science.

32) Public resources should be devoted to research on how to armour citizens against manipulation.

33) The European Commission should initiate and support supranational programmes to organise platforms for and finance investigative journalism, including solution journalism and data-driven journalism at the level of Member States and also the EU. It should encourage cross-border collaborations that bring together participants from more countries and editorial offices. While this branch performs precious public services for democratic states, it is often unwelcome because of the criticism. The new EU budget period should extend the availability of existing resources and allocate dedicated resources to these programmes.

34) Similar programmes should encourage and support independent journalistic organisations that provide credibility indices, to create easily readable, frequently updated databases. Granting a ‘quality insurance label’ to credible news outlets could promptly inform users about the trustworthiness of a source.

35) Journalistic associations should be encouraged or supported to participate in fact-checking; fact-checkers are recommended to cooperate with social media sites, using a technological short-cut to stop the disinformation from spreading.

[631] Examples can be: the Media Pluralism Monitor, http://cmpf.eui.eu/media-pluralism-monitor/, the European Centre for Press and Media Freedom, https://ecpmf.eu.

36) The European Commission should initiate and support the creation of an international, European, high-quality media service,[632] which would provide EU-focused general news and opinion pieces in an engaging and diverse manner and is also disseminated through platform services.

37) Platform providers should be responsible for proactively protecting the personal data of their users and enforcing data protection principles on their platforms.

38) Platform providers must not systematically favour any political, ideological or religious opinion, must not give preference for content that is their own or by an affiliated company, and must not discriminate among users or their content based on protected characteristics, such as race, gender or political opinion.

39) Platform providers should expediently administer their platforms.

40) Platform providers should create algorithms that foster diversity and empower users.

41) Platform providers should employ safeguarding rules for advertisements.

42) Platform providers should offer their users the option to choose the required level of diversity, including the extent to which they wish to see content that is different from what would normally be recommended to them by the default algorithms. The user should be able to make this choice easily and repeatedly.

43) Platform providers should include in their offered newsfeeds information representing diverse views from what is otherwise recommended (e.g. 'different opinions', 'related news', 'other than this', as is already done by some platform providers).

44) Platform providers should give priority to news with certified credibility or even public service content.[633]

Source: Authors.

[632] Some regard the Euronews as a common European channel; however, 53% of its shares is owned by Egyptian Media Globe Networks, 25 % by American NBC, and 22% only by European public service broadcasters, and it is not limited to the EU. Euractiv comes closest to such a function.

[633] This can backfire in the case of a captured, or authoritarian state.

REFERENCES

Academic papers, chapters and research papers:

- Adam, Barbara, Ulrich Beck and Joost Van Loon (eds.) (2000), *The risk society and beyond: critical issues for social theory*, London: Sage Publications, 3.
- Aday, Sean, Steven Livingston, and Maeve Hebert (2005), "Embedding the Truth: A Cross-Cultural Analysis of Objectivity and Television Coverage of the Iraq War." *Harvard International Journal of Press/Politics* 10, no. 1, January, p. 3–21. (https://doi.org/10.1177/1081180X05275727.)
- Alaphilippe, A., C. Ceccarelli, L. Charlet, and M. Mycielski (2018), "Disinformation Detection System: 2018 Italian Elections." Brussels: EU Disinfo Lab, June 1. (https://disinfo.eu/wp-content/uploads/2018/06/2018-Italian-elections-Case-report.pdf)
- Alaphilippe, A., C. Ceccarelli, L. Charlet, M. Mycielski (2018), "Developing a disinformation detection system and sourcing it live – The case study of the 2018 Italian elections", 17 May, Brussels. (http://disinfo.eu/wp-content/uploads/2018/05/20180517_Scientific_paper.pdf)
- Albers, M. and Sarlet, I. (eds.), Personality and Data Protection Rights on the Internet, Forthcoming
- Alemanno, Alberto (2018), "How to Counter Fake News? A taxonomy of anti-fake news approaches", European Journal of Risk Regulation, No. 1.
- Alemanno, Alberto and, Justine Brogi, Maxime Fischer-Zernin, and Paige Morrow (2018), "Is the EU Disinformation Review Compliant with EU Law? Complaint to the European Ombudsman About the EU Anti-Fake News Initiative", *HEC Paris Research Paper No. LAW-2018-1273,* 28 March 28. (https://ssrn.com/abstract=3151424 or http://dx.doi.org/10.2139/ssrn.3151424)
- Allcott H. and Gentzkow M., Social media and fake news in the 2016 election. Stanford University, Journal of Economic Perspectives 31(2): 211-236, 2017.
- Asmolov G., The Disconnective Power of Disinformation Campaigns. Journal of International Affairs 71(1.5): Columbia, 18 September 2018
- Bard, Mitchell T. (2017), "Propaganda, Persuasion, or Journalism?: Fox News' Prime-Time Coverage of Health-Care Reform in 2009 and 2014." Electronic News 11, no. 2, June 1, p. 100–118. (https://doi.org/10.1177/1931243117710278)
- Bárd, Petra and Judit Bayer (2016), "A comparative analysis of media freedom and pluralism in the EU Member States", Study for the LIBE Committee, PE 571.376 EN. (http://www.europarl.europa.eu/supporting-analyses)
- Bárd, Petra and Sergio Carrera (2017), "The Commission's Decision on 'Less EU' in Safeguarding the Rule of Law: A play in four acts", No. 08, p. 1-11, CEPS Policy Insight, Brussels. (https://www.ceps.eu/publications/commission's-decision-'less-eu'-safeguarding-rule-law-play-four-acts)
- Bárd, Petra and Wouter Van Ballegooij (2018), "Judicial independence as a precondition for mutual trust? The CJEU in Minister for Justice and Equality v. LM", *New Journal of European Criminal Law*, Vo. 9, No. 3.
- Bárd, Petra, Sergio Carrera, Elspeth Guild and Dimitry Kochenov (2016), "An EU mechanism on Democracy, the Rule of Law and Fundamental Rights", Centre for European Policy Studies (CEPS), Brussels, April.
- Bastos, Marco T., and Dan Mercea (2017), "The Brexit Botnet and User-Generated Hyperpartisan News." *Social Science Computer Review*, October 10.
- Bayer, Judit (2007), "Liability of Internet Service Providers for Third Party Content – A comparative analysis with policy recommendations", *VUW Law Report Special Edition*, Wellington, New Zealand, p. 1-109.
- Bayer, Judit (2017). "Media Pluralism in Third-Wave Democracies – A Potential of European Legislation to Improve Media Freedom and Pluralism". In Bajomi-Lázár, Péter (ed.) (2017), *Media in Third-Wave Democracies – Southern and Central/Eastern Europe in a Comparative Perspective*, Budapest: *L'Harmattan*, p. 19-44.
- Bayer, Judit (2018), "Media freedom and pluralism: legislation and enforcement at the European level", *ERA Forum Journal of the Academy of European Law*.
- Bayer, Judit (forthcoming), "The illusion of pluralism – Regulatory aspects of equality in the new media", the Euromedia Research Group.

- Bayer, Judit. (2010), "Public Service Television in a Changing Technological and Legal Environment", in Sapio B., L. Raycheva and A. Urban (Eds.) (2010), *Digital Television: Emerging Markets and Challenges for Policy Making*, *Special Issue of Communications, Politics and Culture*, Vol. 43, No. 2, pp. 6-23.
- Bennett, C. J., Voter databases, micro-targeting, and data protection law: can political parties campaign in Europe as they do in North America? International Data Privacy Law, 6(4), 261–275.
- Bertin Martens, Luis Aguiar, Estrella Gomez-Herrera and Frank Mueller-Langer, The digital transformation of news media and the rise of disinformation and fake news – An economic perspective; Digital Economy Working Paper 2018-02; JRC Technical Reports. https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf
- Bertin Martens, Luis Aguiar, Estrella Gomez-Herrera Frank Mueller-Langer: The digital transformation of news media and the rise of disinformation and fake news. An economic perspective. JRC Digital Economy Working Paper 2018-02. https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf
- Bessi a, Zollo F, Del Vicario M, Puliga M, Scala A, Caldarelli G, et al. (2016) Users Polarization on Facebook and Youtube PLoS ONE 11(8):a0159641. doi:10.1371/journal.pone.0159641.
- Best, James D.: Constitutional Speed Bumps, http://www.whatwouldthefoundersthink.com/constitutional-speed-bumps
- Bigot, L. (2018), "Rétablir la vérité via le fact-checking : l'ambivalence des médias face aux fausses informations", Le Temps des médias, n° 30.
- Bode, L. and Emily K. Vrada (2015), "In Related News, That Was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media", *Journal of Communication*, 4(65), p. 619-638.
- Bodó, B. et al., Political micro-targeting: a Manchurian candidate or just a dark horse? Internet Policy Review 6(4):2017
- Borgesius, F. et al., Online Political Microtargeting: Promises and Threats for Democracy. Utrecht Law Review, 14(1): 82–96
- Bossetta, M. (2018), The Digital Architectures of Social Media: Comparing Political Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election. Journalism & Mass Communication Quarterly I-26.
- Bradshaw, S. and P. Howard (2017) Troops, "Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation", Working paper no. 2017.12, University of Oxford. (http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/07/Troops-Trolls-and-Troublemakers.pdf)
- Bradshaw, Samantha, and Philip N Howard (2018), Online Supplement to Working Paper 2018.1, "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation.", Computational Propaganda Project, Oxford: Oxford Internet Institute (http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct_appendix.pdf.
- Brattberg, Erik, Tim Maurer (2018), "Russian Election Interference – Europe's Counter to Fake News and Cyber Attacks", Carnegie Endowment for International Peace, Washington, 2018.
- Buyse, Antoine: „Dangerous Expressions: The Echr, Violence and Free Speech" *ICLQ* 2014/April. 491–503;
- Cannie, Hannes – Dirk Voorhoof: „The Abuse Clause and Freedom of Expression in the European Human Rights Convention: An Added Value for Democracy and Human Rights Protection?" *Netherlands Quarterly of Human Rights* 2011/1. 54–83, 57.
- Caplan, R. and Danah Boyd (2016), "Who Controls the Public Sphere in an Era of Algorithms? Mediation, Automation, Power", 13 May. (https://datasociety.net/pubs/ap/MediationAutomationPower_2016.pdf)
- Cappello M. (ed.), Media coverage of elections: the legal framework in Europe, IRIS Special, European Audiovisual Observatory, Strasbourg, 2017, p 114
- Carrera, Sergio, Elspeth Guild and Nicholas Hernanz (2013), "The Triangular Relationship between Fundamental Rights, Democracy and the Rule of Law in the EU, Towards an EU Copenhagen Mechanism". European Parliament, DG IPOL, Brussels, October. (http://www.europarl.europa.eu/RegData/etudes/etudes/join/2013/493031/IPOL-LIBE_ET%282013%29493031_EN.pdf)
- Conway, Mike, Maria Elizabeth Grabe, and Kevin Grieves (2007), "Villains, Victims and the Virtuous in Bill O'Reilly's 'No Spin Zone.'" *Journalism Studies* 8, no. 2 (April): 197–223 (https://doi.org/10.1080/14616700601148820).
- Curran, J. (2011), Chapter 9 "Press as an agency of social control", in *Media and Democracy,* Routledge.

- Douglas, Mary and Aaron Wildavsky (1982), "Risk and culture: an essay on the selection of technical and environmental dangers", Berkeley: *University of California Press*.
- Dutton, William H., Bianca Reisdorf, Elizabeth Dubois and Grant Blank (2017), "Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States", *Quello Center* Working Paper No. 5-1-171, May. (https://ssrn.com/abstract=2960697 or http://dx.doi.org/10.2139/ssrn.2960697).
- Dworkin, R. (2013), "A New Philosophy of International Law'", *Philosophy and Public Affairs* 1, 41, p. 2–30.
- Engesser, Sven, Nayla Fawzi and Anders Olof Larsson (2017), "Populist online communication: introduction to the special issue", *Information, Communication & Society*, 24 May.
- Epstein R and Robertson RE, 'The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections.' (2015) 112 Proceedings of the National Academy of Sciences of the United States of America PNAS August 18, 2015 112 (33) E4512-E4521; https://doi.org/10.1073/pnas.1419828112
- Ernst, Nicole – Sven Engesser, Florin Büchel, Sina Blassnig and Frank Esser (2017) Extreme parties and populism: an analysis of Facebook and Twitter across six countries. Information Communication and Society. 20:9, 1347-1364.
- Eyler-Driscoll, S., A. Schechter, C. Patiño (2018), "Digital Platforms and Concentration", Second Annual antitrust and competition conference, Stigler center for the study pf the economy and the State University Booth School of Business, A ProMarket Production. (https://promarket.org/wp-content/uploads/2018/04/Digital-Platforms-and-Concentration.pdf)
- Fahlquist, Jessica Nihlén (2009), "Moral Responsibility for Environmental Problems—Individual or Institutional?", *Journal of Agricultural and Environmental Ethics*, April.
- Ferrara, Emilio (2017), „Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, June 30, 2017. (https://papers.ssrn.com/abstract=2995809).
- Figueira, A., and Oliveira, L., The current state of fake news: challenges and opportunities, Procedia Computer Science 121 (2017): 817–825, pp. 820-822, https://ac.els-cdn.com/S1877050917323086/1-s2.0-S1877050917323086-main.pdf?_tid=59d836f5-89df-4fbc-a94b-19c1af0e1b36&acdnat=1545743311_63b257b853f17ff893a949210a7f4028
- Finkel, J. et al. (2017), « Fake news and disinformation: The roles of the nation's digital newsstands, Facebook, Google, Twitter and Reddit ». *Stanford Law School*, (https://www-cdn.law.stanford.edu/wp-content/uploads/2017/10/Fake-News-Misinformation-FINAL-PDF.pdf)
- Fisher, Walter R. (1984) Narration as a human communication paradigm: The case of public moral argument, Communication Monographs, 51:1, 1-22, DOI: 10.1080/03637758409390180
- Fisher, Walter R.; The Narrative Paradigm: In the Beginning, Journal of Communication, Volume 35, Issue 4, 1 December 1985, Pages 74–89, https://doi.org/10.1111/j.1460-2466.1985.tb02974.x
- Fletcher, Richard, Alessio Cornia, Lucas Graves, and Rasmus Kleis Nielsen (2018), "Measuring the Reach of 'Fake News' and Online Disinformation in Europe." Reuters Institute for the Study of Journalism; University of Oxford, February (https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/Measuring%20the%20reach%20of%20fake%20news%20and%20online%20distribution%20in%20Europe%20CORRECT%20FLAG.pdf).
- Freedom House (2017), Freedom on the Net 2017 (https://freedomhouse.org/sites/default/files/FOTN_2017_Final.pdf)
- Friedman, George (2018), "Special Report. A History of Fake News", *Geopolitical Futures*, April.
- Future of Privacy Forum (2018), The Privacy Expert's Guide to Artificial Intelligence and Machine Learning (https://fpf.org/wp-content/uploads/2018/10/FPF_Artificial-Intelligence_Digital.pdf)
- Garland, David (2001), "*The culture of control"*, Oxford: Oxford University Press.
- Gelfert A., Fake news: a definition. Informal Logic 38 (1):84-117, 2018
- Ghosh D. and Scott B. (2018), #Digitaldeceit. The technologies behind precision propaganda and the Internet. Harvard Kennedy School, January 2018
- Giles, Keir (2016), "Handbook of Russian Information Warfare", Fellowship Monograph 9, Research Division, NATO Defense College, November.

- Ginsburg, Tom and Aziz Z. Huq, Mila Versteeg, The Coming Demise of Liberal Constitutionalism?, The University of Chicago Law Review, Volume 85, Issue 2 (March 2018) 239–255, 253.
- Gorwa R. and Guibeault D., Unpacking the Social Media Bot: A Typology to Guide Research and Policy. Policy and Internet, 2018. https://doi.org/10.1002/poi3.184
- Graves, Lucas and Federica Cherubini (2016), "The Rise of Fact-Checking Sites in Europe", Digital News Project, *Reuters Institute for the Study of Journalism*. (https://reutersinstitute.politics.ox.ac.uk/sites/default/files/research/files/The%2520Rise%2520of%2520Fact-Checking%2520Sites%2520in%2520Europe.pdf)
- Green, David A. (2006), "Public Opinion Versus Public Judgment About Crime: Correcting the 'Comedy of Errors'," 46 *British Journal of Criminology* 1, p. 131-154.
- Greenberg, Lawrence T., Seymour E. Goodman and Kevin J. Soo Hoo (1998), "Information Warfare and International Law", *National Defense University Press*. (https://www.researchgate.net/publication/235066729_Information_Warfare_and_International_Law)
- Hedman, Freja, Fabian Sivnert, Bence Kollanyi, Vidya Narayanan, Lisa-Maria Neudert, and Philip N. Howard (2018), "News and Political Information Consumption in Sweden: Mapping the 2018 Swedish General Election on Twitter.", Oxford, UK: Project on Computational Propaganda.
- Helberg, N., Profiling and Targeting Consumers in the Internet of Things – A New Challenge for Consumer Law in "Digital Revolution" (pp. 135-161), Baden-Baden: Nomos Verlag, 2017
- Helberger, N., k. Karppinen, K., & L. D'Acunto (2018), "Exposure diversity as a design principle for recommender systems", *Information Communication & Society*, Vo.*21*(2), p. 191–207. (https://doi.org/10.1080/1369118X.2016.1271900)
- Helberger, N., Katharina Kleinen-von Königslöw and Rob van der Noll (2015), "Regulating the new information intermediaries as gatekeepers of information diversity", info, Vol. 17, No. 6, p.50-71, (https://doi.org/10.1108/info-05-2015-0034)
- Howard, P. (2017), "Junk News and Bots during the French Presidential Election: What Are French Voters Sharing Over Twitter?", COMPROP data memo: *Oxford University*, 22 April 2017
- Howard, P. and Kollanyi, Bots (2016), "#StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum". COMPROP Research note 2016.1: *Oxford University* (https://arxiv.org/ftp/arxiv/papers/1606/1606.06356.pdf)
- Howard, P.(2017), "Social Media, News and Political Information during the US Election: Was Polarizing Content Concentrated in Swing States?", 28 September 2017 (https://www.recode.net/2017/9/28/16378186/twitter-fake-news-misinformation-russia-oxford-swing-states)
- Hsu, J. (2018), "Experts Bet on First Deepfakes Political Scandal", 22 June 2018, (https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/experts-bet-on-first-deepfakes-political-scandal)
- http://www.pnas.org/content/112/33/E4512.abstract?tab=author-info
- Huckle, S. and White, M., Catalogue of all projects working to solve Misinformation and Disinformation. Big Data 5(4) : 356-371, 2017
- Ignatieff, Michael and Stefan Roch (eds.) (2018) "Academic Freedom: The Global Challenge", Budapest: CEU Press.
- Ioannisyan, Daniel, Catherine Shanahan, Carien J. Touwen, Roman Dobrokhotov, Simas Čelutka, Gunnar Grimsson, Erdoğan İşcan, Conor Mcardle, "World Forum for Democracy Is Populism a problem – Lab 9 – Fake News: Does Fact Checking Work?" (https://rm.coe.int/wfd-2017-report-lab-9-fake-news-does-fact-checking-work-/168077075c)
- Ireton, Cherilyn and Julie Posetti: Journalism, 'Fake News' & Disinformation. Handbook for Journalism Education and Training. 2018. UNESCO – a training course description.
- Jensen, Eric Talbot (2017), "The Tallinn Manual 2.0: Highlights and Insights", *Georgetown Journal of International Law 735,* Vol. 48, p. 740-744.
- Jowett, Garth and O'Donnell V. (2011), *Propaganda & Persuasion*. SAGE Publications, 12 Apr 2011
- Jowett, Garth S. (2011), *Propaganda Through the Ages,* SAGE Publications, April.

- Kaltheuner, F. and Bietti, E. (2018), “Data is power: Towards additional guidance on profiling and automated decision-making” (https://privacyinternational.org/sites/default/files/2018-04/Data%20Is%20Power-Profiling%20and%20Automated%20Decision-Making%20in%20GDPR.pdf)
- Kandemir, Berfin, Alezander Brand (2017), "Social Media In Operations – A Counter-Terrorism Perspective", *NATO StatCom COE*, September. (https://www.stratcomcoe.org/social-media-operations-counter-terrorism-perspective)
- Kanuck, Sean (2010), “Sovereign Discourse on Cyber Conflict Under International Law”, *Texas Law Review,* Vol. 88, p.1571-1597, June.
- Karlova, N. and Lee, J. (2012), *Notes from the underground city of disinformation: A conceptual investigation*. Proceedings of the American Society for Information Science and Technology 48(1).
- Kim, H. et al., M. (2018), *Deep Video Portraits*. Journal ACM Transactions on Graphics 37(4)
- Kim, Y. (2018), Closing the Digital Loopholes that Pave the Way for Foreign Interference in U.S. Elections. Report based on study by (University of Wisconsin), 16 April 2018, https://campaignlegal.org/sites/default/files/04-16-18%20CLC-IO%20Issue%20Brief%20Young%20Mie%20Report%20FINAL.pdf
- King, Gary, Jennifer Pan, and Margaret E. Roberts (2017), “How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument.” American Political Science Review 111, no. 03, August: 484–501. (https://doi.org/10.1017/S0003055417000144).
- Klein, D. and Wueller J. (2017), *Fake news: a legal perspective*. Journal of Internet Law 20(10): 5-13
- Kranzberg, M. (1986), Technology and History: "Kranzberg's Laws". *Technology and Culture* 27(3) : 544-560, July.
- Kurowska, Xymena, Anatoly Reshetnikov (2018), “Neutrollization: Industrialized trolling as a pro-Kremlin strategy of desecuritization”, SAGE journals, Security Dialogue, 49(5), p.345–363.
- Larsson, Anders Olof (2018), “The News User on Social Media: A Comparative Study of Interacting with Media Organizations on Facebook and Instagram.” *Journalism Studies* 19, no. 15, November: 2225–42. (https://doi.org/10.1080/1461670X.2017.1332957).
- Lessenski, Marin (2018), “Common Sense Wanted. Resilience To ‘Post-Truth’ And Its Predictors In The New Media Literacy Index”, *Open Society Institute Sofia*, March. (http://osi.bg/downloads/File/2018/MediaLiteracyIndex2018_publishENG.pdf)
- Lessig, Lawrence: What Things Regulate Speech: CDA 2.0 vs. Filtering. https://cyber.harvard.edu/works/lessig/what_things.pdf
- Lisa-Maria Neudert, Bence Kollanyi, Philip N. Howard (2017), “Junk News and Bots during the German Parliamentary Election: What are German Voters Sharing over Twitter?” COMPROP DATA MEMO 2017.7, Oxford University, 19 September. (http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/09/ComProp_GermanElections_Sep2017v5.pdf)
- Loewenstein, K. (1937), “Militant Democracy and Fundamental Rights”, 31 *American Political Science Review.*
- Marchal, Nachema, Lisa-Maria Neudert, Bence Kollányi, and Philip N Howard (2018), “Polarization, Partisanship and Junk News Consumption on Social Media During the 2018 US Midterm Elections.” *COMPROP* DATA MEMO 2018. 5. Oxford: Oxford Internet Institute, November 1. (https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/11/marchal_et_al.pdf).
- Martin, J. (1958), Definition of propaganda in “International Propaganda: Its Legal and Diplomatic Control”. *University of Minnesota Press*.
- Marwick, Alice (2018), ‘Beyond the Magic Bullet Theory of Fake News: Disinformation as Identity Expression’. Keynote speech at the iCS Symposium on Challenges to Studying Disinformation (University of Copenhagen), 27-28 October 2018.
- Marwick, Alice and Lewis, Rebecca (2017), “Media, Manipulation and Disinformation Online”. *Data and Society* (https://centerformediajustice.org/wp-content/uploads/2017/07/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf)
- Matz, S.C. et al., Psychological targeting as an effective approach to digital mass persuasion, PNAS November 28, 2017 114 (48) 12714-12719, https://doi.org/10.1073/pnas.1710966114, https://www.pnas.org/content/114/48/12714

- Möller, Judith, Damian Trilling, Natali Helberger & Bram van Es (2018),"Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity" *Information, Communication & Society*, March.
- Möller, Kai; Proportionality: Challenging the critics, *International Journal of Constitutional Law*, Volume 10, Issue 3, 1 July 2012, Pages 709–731, https://doi.org/10.1093/icon/mos024
- Moore, M. (2016), Tech Giants and Civic Power. King's College London, April 2016, https://www.kcl.ac.uk/sspp/policy-institute/CMCP/Tech-Giants-and-Civic-Power.pdf
- Müller, Karsten, and Carlo Schwarz (2018), "Fanning the Flames of Hate: Social Media and Hate Crime." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, May 21. (https://papers.ssrn.com/abstract=3082972).
- Müller, Karsten, and Carlo Schwarz (2018), "Making America Hate Again? Twitter and Hate Crime Under Trump." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, March 30. (https://papers.ssrn.com/abstract=3149103).
- Natale, S. and Andrea Ballatore (2014), "The web will kill them all: new media, digital utopia, and political struggle in the Italian 5-Star Movement", *Media, Culture & Society*, 36(1), p105–121, January.(http://journals.sagepub.com/doi/abs/10.1177/0163443713511902)
- NATO (2017), "Digital Hydra: Security Implications of False Information Online", Anna Reynolds (eds), Riga, November. (https://www.stratcomcoe.org/digital-hydra-security-implications-false-information-online).
- Nemitz, Paul (2018), « Constitutional Democracy and Technology in the age of Artificial Intelligence", 18 August 2018 (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3234336)
- Neudert L-M. et al. (2017), "Junk News and Bots during the German Parliamentary Election: What are German Voters Sharing over Twitter?", COMPROP data memo: *Oxford University*, 19 September 2017.
- Neudert, Lisa-Maria (2017), "Computational Propaganda in Germany: A Cautionary Tale." Project on Computation Propaganda. Oxford, UK: Oxford Internet Institute. (http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-Germany.pdf).
- Newman, Nic, Richard Fletcher, Antonis Kalogeropoulos, David A. L. Levy and Rasmus Kleis Nielsen (2018), "Digital News Report 2018." *Reuters Institute, University of Oxford*. (http://media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf?x89475)
- Nielsen, R. and Graves, L. (2017), "News you don't believe": Audience perspectives on fake news. Factsheet: October 2017 (http://reutersinstitute.politics.ox.ac.uk/sites/default/files/2017-10/Nielsen%26Graves_factsheet_1710v3_FINAL_download.pdf)
- Nocetti, Julien (2015), "Contest and conquest: Russia and global internet governance", *International Affairs*, 91: 1, p. 111–130.
- Pfersmann, Otto, Shaping Militant Democracy: Legal Limits to Democratic Stability, in: András Sajó (ed.), Utrecht: Eleven International Publishing, 2004, 47-68.
- Rawls, John (1977), "The idea of public reason revisited", *The University of Chicago Law Review*, vol. 64, no. 3.
- Renda, Andrea (2018), "The legal framework to address "fake news": possible policy actions at the EU level", Policy Department for Economic, Scientific and Quality of Life Policies, CEPS, Directorate-General for Internal Policies, PE 619.013, June.
- Renno, R. (2018), « WhatsApp: The Widespread Use of WhatsApp in Political Campaigning in the Global South », 3 July 2018 (https://ourdataourselves.tacticaltech.org/posts/whatsapp/)
- Rini, R. (2017), "Fake news and partisan epistemology". *Kennedy Institute of Ethics Journal* 27(2): 43-64
- Roozenbeek, Jon and van der Linden, Sander, The Fake News Game: Actively Inoculating Against the Risk of Misinformation, From: https://www.cam.ac.uk/sites/www.cam.ac.uk/files/fakenews_latest_jrr_aaas.pdf
- Russell, Martin (2016), Briefing "Russia's information war: Propaganda or counter-propaganda?", *European Parliamentary Research Service (EPRS)*, October. (http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/589810/EPRS_BRI(2016)589810_EN.pdf)
- Sajó, A: Militant Democracy and Transition towards Democracy, in: András Sajó (ed.), Utrecht: Eleven International Publishing, 2004, 209-230.
- Scheppele, Kim Lane, 'The Rule of Law and the *FrankenState*: Why Governance Checklists Do Not Work' (2013) 26 Governance 4, 559–562.

- Schulz, W. (2018), “Regulating Intermediaries to Protect Privacy Online – The Case of the German NetzDG” HIIG Discussion Paper Series, 2018.(1). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3216572&download=yes
- Scott Maravilla, Christopher (2008), “Hate Speech as a War Crime: Public and Direct Incitement to Genocide in International Law”, 17 Tul. J. Int'l & Comp. L. 113, 144.
- Stone, Geoffrey R.: Restriction of Speech Because of Its Content: The Peculiar Case of Subject-Matter Restrictions. University of Chicago Law School. Chicago Unbound. 1978. https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1530&context=journal_articles
- Syrovátka, Jonáš, and Jaroslav Hroch (2018), “Czech Presidential Election 2018.” Czech Elections in the Era of Disinformation. Prague: Prague Security Studies Institute, http://www.pssi.cz/download/docs/545_presidential-election-2018-analysis.pdf.
- Szuleka, Małgorzata (2018), "First victims or last guardians? The consequences of rule of law backsliding for NGOs: Case studies of Hungary and Poland”, CEPS Paper in Liberty and Security in Europe, No.6. (https://www.ceps.eu/system/files/MSzuleka_RoLandNGOs.pdf)
- Tandoc E. et al. (2018), Defining “fake news” a typology of scholarly definitions. *Digital Journalism* 6(2): 137-153
- Thorleifsson, Cathrine (2017), “Disposable Strangers: Far-Right Securitisation of Forced Migration in Hungary.” Social Anthropology 25, no. 3, August: 318–34. (https://doi.org/10.1111/1469-8676.12420).
- Toggenburg, G. N. and Grimheden, J. (2016), “The Rule of Law and the Role of Fundamental Rights: Seven Practical Pointers”, in: Closa, C. and Kochenov, D. (eds.), *Reinforcing Rule of Law Oversight in the European Union*, Cambridge: Cambridge University Press.
- Tucker, J. et al. (2018), Social media, political polarization and political disinformation: a review of the scientific literature. Hewlett Foundation, March 2018 (https://hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf)
- Vaidhyanathan, S. (2018), *Antisocial Media How Facebook Disconnects Us and Undermines Democracy,* Virginia Quarterly Review, June.
- Van Ballegooij, Wouter and Tatjana Evas, ‘An EU mechanism on democracy, the rule of law and fundamental rights, interim European added value assessment accompanying the legislative initiative report (Rapporteur Sophie in ‘t Veld)’ European Parliamentary Research Service, October 2016, PE.579.328;
- Vian B. and McStay A. (2018), “Fake news and the economy of emotions: problems, causes, solutions”, *Digital Journalism* 6(2): 154-175.
- Villiger, Mark E.: „Article 17. ECHR and Freedom of Speech in Strasbourg Practice” in Josep Casadevall [et al.] (eds.): *Freedom of Expression: Essays in Honour of Nicolas Bratza* (The Netherlands: Wolf Legal Publishers 2012) 321, 322.
- Vilmer, J-B. et al.(2018), Information Manipulation: A Challenge for Our Democracies, report by the Policy Planning Staff (CAPS) of the Ministry for Europe and Foreign Affairs and the Institute for Strategic Research (IRSEM) of the Ministry for the Armed Forces, Paris, August 2018 (https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf).
- Vosoughi, S. Deb Roy and Sinan Aral (2018), “The spread of true and false news online”, *Science* 359(6380): 1146-1151, 9 March 2018
- Wachter, S. and Mittelstadt, B., A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI, 5 October 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829
- Wardle, Claire and Derakhshan, Hossein (2017), “Information Disorder: Toward an interdisciplinary framework for research and policy making”, Council of Europe, report DGI(2017)09.
- Williams, Alex T. (2017),” Measuring the Journalism Crisis: Developing New Approaches That Help the Public Connect to the Issue”, *International Journal of Communication* 11(2017), Feature 4731–4743.
- Wooley, S. and Guilbeault D. (2017), Computational Propaganda in the United States of America: Manufacturing Consensus Online, Working Paper No 2017.5: Oxford University (http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-USA.pdf)
- Yglesias, Matthew (2018), “The Case for Fox News Studies.”, Political Communication 0, no. 0 (October 23): 1–3. (https://doi.org/10.1080/10584609.2018.1477532).

- Zhdanova, Mariia and Dariya Orlova (2017), "Computational Propaganda in Ukraine: Caught between external threats and internal challenges.", Oxford Internet Institute. Samuel Woolley and Philip N. Howard (eds), Working Paper 2017.9, Oxford, 19 June.
- Zuiderveen Borgesius, F. & Trilling, D. & Möller, J. & Bodó, B. & de Vreese, C. & Helberger, N. (2016). Should we worry about filter bubbles?. Internet Policy Review, 5(1). DOI: 10.14763/2016.1.401.
- Zuiderveen Borgesius, F.J., Möller, J., Kruikemeier, S., Ó Fathaigh, R., Irion, K., Dobber, T., Bodo, B. and de Vreese, C., (2018). Online Political Microtargeting: Promises and Threats for Democracy. Utrecht Law Review, 14(1), pp.82–96. DOI: http://doi.org/10.18352/ulr.420

Academic books:

- Bauman, Zygmunt (2006), *Liquid Fear*, Cambridge: Polity.
- Beck, Ulrich (1986), *Risikogesellschaft: auf dem Weg in eine andere Moderne,* Frankfurt am Main: Suhrkamp.
- Bohman, James (1998), *Deliberative Democracy*, Cambridge, MA.: MIT Press.
- Carey, James W. (1989): Communication as Culture. Essays on Media and Society. New York & London: Routledge.
- Chomsky, N. and Edward S. Herman (1988), *Manufacturing Consent: The Political Economy of the Mass,* Pantheon Books.
- Choucri, Nazli (2012), *Cyberpolitics in international relations,* Cambridge, MA: MIT Press.
- Curran, James, 2011, Media and Democracy. Routledge. London.
- Fisher, W.R. (1995) Narration, knowledge and the possibility of wisdom. In. Rethinking Knowledge: Reflections across the Disciplines. by Goodman, Fisher W.R. (eds.)
- Fukuyama, Francis: Identity: The Demand for Dignity and the Politics of Resentment. Farrar, Straus and Giroux, 2018.
- Giddens, Anthony (1995), "*The consequences of modernity"*, Cambridge: Polity Press.
- Gutmann, Amy and Dennis THOMPSON (1996), "Democracy and disagreement", Cambridge: Harvard
- Habermas, Jürgen (1996), "Between facts and norms: contributions to a discourse theory of law and democracy", Cambridge: MIT Press.
- Hendricks, V.F. – Vestergaard, Mads: Reality Lost. Markets of Attention, Misinformation and Manipulation. Springer, 2018. (Open Access)
- Inglis, F. (1990), *Media Theory*, Blackwell Pub, August.
- Lochocki, T. (2018). The Rise of Populism in Western Europe. [electronic resource] : A Media Analysis on Failed Political Messaging. Cham : Springer International Publishing : Imprint: Springer, 2018.
- Lyotard, Jean-Francois (1979), *La condition postmoderne: rapport sur le savoir*, Paris: Editions de Minuit.
- McLuhan, Marshall: Understanding Media. The extensions of Man. MIT Press, 1994.
- Manning, Martin J. and Herbert Romerstein (2004), *Historical Dictionary of American Propaganda,* Greenwood, 30 November.
- Meiklejohn, Alexander: Free Speech and Its Relation to Self-Government. The Lawbook Exchange, Clark, New Jersey, 2004.
- O'Neil, Cathy: Weapons of Math Destruction (2016) Crown Publishing, New York.
- Oxford English Dictionary, 2nd ed., 1989
- Ritter, Mark and Ulrich Beck (1992), *Risk society: towards a new modernity*, London: Sage Publications.
- Robinson, Linda et al (2018), Modern Political Warfare. Santa Monica, CA: RAND Corporation, 2018
- Schmitt, Michael N. (2017), "Tallinn Manual 2.0 On the International Law Applicable to Cyber Operations', *Cambridge University Press*.
- Sunstein, Cass: #Republic: Divided Democracy in the Age of Social Media. 2007. Princeton University Press.
- Verdross, A. (1923), *Die Einheit des rechtlichen Weltbildes auf Grundlage der Völkerrechtsverfassung*, Tübingen: Mohr.

Legislative instruments, working papers and official reports

- "EU Action Plan on Strategic Communication", Ref. Ares(2015)2608242, 22 June 2015.

- 47 U.S. Code § 230 (c) – Protection for private blocking and screening of offensive material
- Article 29 Data Protection Working Party, Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (n 19)
- Article 29 Data Protection Working Party, Guidelines on the Right to Data Portability (2017) 16/EN WP 242 rev.01 11
- Code Pénal, Livre III. http://www.codes-et-lois.fr/code-penal/article-322-14
- Commission Recommendation (EU) 2017/1584, and its Annex.
- Commission Recommendation (EU) 2017/1584. 13. Sept. 2017. on coordinated response to large-scale cybersecurity incidents and crises. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017H1584&from=EN
- Communication from the Commission on the application of State aid rules to public service broadcasting 2009/C 257/1
- Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Securing free and fair European elections. Brussels, 12.9.2018. COM (2018) 637 final.
- Communications Networks, Content and Technology (2018), "A multi-dimensional approach to disinformation Report of the independent High level Group on fake news and online disinformation", European Union, Brussels, March.
- Council of Europe (2007), Recommendation CM/Rec(2007)16 on measures to promote the public service value of the Internet, Adopted by the Committee of Ministers, 1010th meeting of the Ministers' Deputies, 7 November.
- Council of Europe (2007), Recommendation Rec(2007)2 of the committee of ministers to member states on media pluralism and diversity of media content, adopted by the Committee of Ministers, 985th meeting of the Ministers' Deputies, 31 January 2007.
- Council of Europe (2009), "Convention on Access to Official Documents", Council of Europe Treaty Series – No. 205, Tromsø, 18 June.
- Council of Europe (2010), Recommendation CM/Rec(2010)13 on the protection of individuals with regard to automatic processing of personal data in the context of profiling, adopted by the Committee of Ministers, 23 November.
- Council of Europe (2011), Recommendation CM/Rec(2011)7 on a new notion of media, Adopted by the Committee of Ministers, 1121st meeting of the Ministers' Deputies, 21 September.
- Council of Europe (2012), Recommendation CM/Rec(2012)3 on the protection of human rights with regard to search engines, Adopted by the Committee of Ministers, 1139th meeting of the Ministers' Deputies, 4 April.
- Council of Europe (2012), Recommendation CM/Rec(2012)4 on the protection of human rights with regard to social networking services, adopted by the Committee of Ministers, 1139th meeting of the Ministers' Deputies, 4 April.
- Council of Europe (2013), Recommendation CM/Rec(2013)1 on gender equality and media, Adopted by the Committee of Ministers, 1176th meeting of the Ministers' Deputies, 10 July.
- Council of Europe (2014), Recommendation CM/Rec(2014)6 on a Guide to human rights for Internet users, Adopted by the Committee of Ministers, 1197th meeting of the Ministers' Deputies, 16 April.
- Council of Europe (2015), Recommendation CM/Rec(2015)6 on the free, transboundary flow of information on the Internet, adopted by the Committee of Ministers, 1 April.
- Council of Europe (2016), Recommendation CM/Rec(2016)5 on Internet freedom, Recommendation CM/Rec(2016)1 on protecting and promoting the right to freedom of expression and the right to private life with regard to network neutrality, Adopted by the Committee of Ministers, 1244th meeting of the Ministers' Deputies, 13 January.
- Council of Europe (2017), Consultative Committee of the Convention for the protection of individuals with regard to automatic processing of personal data (T-PD) – Guidelines on the Protection of Individuals with regard to the Processing of Personal Data in a World of Big Data, 23 January, (https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806f06d0)

- Council of Europe (2018), Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, Adopted by the Committee of Ministers, 1309th meeting of the Ministers' Deputies, 7 March.
- Council of the European Union (2008), Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law, L 328/55, 6.12.2008.
- Criminal Code of Hungary.
- Decision n° 2018-774 DC, 20. Dec. 2018. https://www.conseil-constitutionnel.fr/actualites/communique/decision-n-2018-774-dc-du-20-decembre-2018-communique-de-presse
- 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') Articles 12-15.
- Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), Official Journal L 201, 31 July 2002.
- Electoral Commission (2018), "Arron Banks, Better for the Country and Others Referred to the National Crime Agency for Multiple Suspected Offences." The Electoral Commission, November 1, 2018. https://www.electoralcommission.org.uk/i-am-a/journalist/electoral-commission-media-centre/party-and-election-finance-to-keep/arron-banks,-better-for-the-country-and-others-referred-to-the-national-crime-agency-for-multiple-suspected-offences.
- Eurobarometer 461. Designing Europe's Future, April 2017.
- European Commission (2016), Joint Communication to the European Parliament and the Council Joint Framework on countering hybrid threats a European Union response, JOIN(2016) 18 final, Brussels, 6 March.
- European Commission (2017), "Code of Conduct on countering online hate speech – results of evaluation show important progress", Justice and Consumers, 1 June. (http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=71674
- European Commission (2017), "Hungary: Commission takes second step in infringement procedure on Higher Education Law", Brussels, 13 July.
- European Commission (2017), Annex to the Commission Recommendation on Coordinated Response to Large Scale Cybersecurity Incidents and Crises, C(2017) 6100 final, Annex 1, Brussels, 13 September. (http://ec.europa.eu/transparency/regdoc/rep/3/2017/EN/C-2017-6100-F1-EN-ANNEX-1-PART-1.PDF)
- European Commission (2017), Communication to the Parliament and the Council Making the most of NIS – towards the effective implementation of Directive (EU) 2016/1148 concerning measures for a high common level of security of network and information systems across the Union, COM(2017)476 final, Brussels, 13 September.
- European Commission (2017), Follow up to the European Parliament resolution on with recommendations to the Commission on the establishment of an EU mechanism on democracy, the rule of law and fundamental rights, adopted by the Commission on 17 January 2017, SP (2017)16, 17 February.
- European Commission (2017), Joint Communication to the European Parliament and the Council Resilience, Deterrence and Defence: Building strong cybersecurity for the EU, JOIN(2017) 450 final, Brussels, 13 September.
- European Commission (2018), Behavioural Study on Advertising and Marketing Practices in Online Social Media, June 2018 (https://ec.europa.eu/info/sites/info/files/osm-final-report_en.pdf)
- European Commission (2018), Commission Recommendation of 1.3.2018 on measures to effectively tackle illegal content online (C(2018) 1177 final), Brussels, 1 March. (https://ec.europa.eu/digital-single-market/en/news/commission-recommendation-measures-effectively-tackle-illegal-content-online)
- European Commission (2018), Communication from the Commission to the European parliament, the European Council, the Council, The European Economic and Social Committee and the Committee of the Regions Artificial Intelligence for Europe {SWD(2018) 137 final}, COM(2018) 237 final, Brussels, April.
- 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, Brussels, 26 April.

- European Commission (2018), High level Group on fake news and online disinformation, A multi-dimensional approach to disinformation (http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50271)
- European Commission (2018), Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online 2018/0331 (COD), COM(2018) 640 final, Brussels.
- European Commission press release: Code of Practice against disinformation: Commission calls on signatories to intensify their efforts. Brussels, 29 January 2019. http://europa.eu/rapid/press-release_IP-19-746_en.htm
- European Commission for Democracy through Law (Venice Commission), Rule of Law Checklist, 18 March 2016.
- European Commission (2018), Results of Commission's last round of monitoring of the Code of Conduct against online hate speech", 19 January. (http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=612086)
- European Commission, Commission guidance on the application of Union data protection law in the electoral context, 12 September 2018, p. 5, https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-data-protection-law-electoral-guidance-638_en.pdf
- European Data Protection Supervisor (EDPS) (2016), Opinion on coherent enforcement of fundamental rights in the age of big data, 8/2016 (https://edps.europa.eu/sites/edp/files/publication/16-09-23_bigdata_opinion_en.pdf)
- European Data Protection Supervisor (EDPS) (2018), Opinion on online manipulation and personal data, 3/2018 (https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf)
- European Data Protection Supervisor, Opinion 6/2017 on the Proposal for a Regulation on Privacy and Electronic Communications (ePrivacy Regulation), https://edps.europa.eu/sites/edp/files/publication/17-04-24_eprivacy_en.pdf
- European Group on Ethics in Science and New Technologies (2018), "Artificial Intelligence, Robotics and 'Autonomous' Systems", March.
- European Parliament (2015), Understanding propaganda and disinformation, November 2015 (http://www.europarl.europa.eu/RegData/etudes/ATAG/2015/571332/EPRS_ATA(2015)571332_EN.pdf)
- European Parliament (2016), EU strategic communication to counteract anti-EU propaganda by third parties, European Parliament resolution of 23 November 2016 on EU strategic communication to counteract propaganda against it by third parties (2016/2030(INI)), P8_TA(2016)0441, 23 November.
- European Parliament (2016), Resolution of 25 October 2016 with recommendations to the Commission on the establishment of an EU mechanism on democracy, the rule of law and fundamental rights (2015/2254(INL)), P8_TA-PROV(2016)0409, Strasbourg, 25 October.
- European Parliament (2017), Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), P8_TA(2017)0051, Strasbourg, 16 February.
- European Parliament (2017), Resolution on Online platforms and the digital single market (2016/2276(INI)), P8_TA(2017)0272, 15 June.
- European Parliament (2018) Motion for a Resolution to wind up the debate on the statement by the Commission pursuant to Rule 123(2) of the Rules of Procedure on the use of Facebook users' data by Cambridge Analytica and the impact on data protection (2018/2855(RSP)), B8-0480/2018, Committee on Civil Liberties, Justice and Home Affairs, 16 October. (http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+MOTION+B8-2018-0480+0+DOC+PDF+V0//EN)
- European Parliament (2018), Resolution on media pluralism and media freedom in the European Union, (2017/2209(INI)), P8_TA(2018)0204, 3 May.
- European Parliament and the council of the European Union (2000), 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'), L 178/1, 17.7.2000.
- European Parliament and the Council of the European Union (2018), Regulation (EU, Euratom) 2018/673 of the European Parliament and of the Council of 3 May 2018 amending Regulation (EU, Euratom) No 1141/2014 on the statute and funding of European political parties and European political foundations, L114I/1, 4.5.2018. (https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0673&from=EN)

- European Parliament resolution of 25 October 2018 on the use of Facebook users' data by Cambridge Analytica and the impact on data protection (2018/2855(RSP)), http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2018-0433+0+DOC+XML+V0//EN
- European Parliament, Fake news' and the EU's response (April 2017), http://www.europarl.europa.eu/RegData/etudes/ATAG/2017/599384/EPRS_ATA%282017%29599384_EN.pdf
- European Union External Action (2018), "Joint Report on the implementation of the Joint Framework on countering hybrid threats from July 2017 to June 2018", 13 June.
- Facebook Baseline Report on Implementation of the Code of Practice on Disinformation. Published on 29. Jan. 2019. http://ec.europa.eu/information_society/newsroom/image/document/2019-5/facebook_baseline_report_on_implementation_of_the_code_of_practice_on_disinformation_CF161D11-9A54-3E27-65D58168CAC40050_56991.pdf
- Flash Eurobarometer (EBS) 464, Fake news and disinformation online. April 2018.
- German Federal Criminal Code
- Grupo Parlamentario Popular en el Congreso, 2017. dec. 18. : http://www.gppopular.es/wp-content/uploads/2017/12/171219-PNL-Noticias-falsas.pdf?_ga=2.226110172.738521628.1541868032-2076439625.1541868032
- Guiding Principles on Business and Human Rights, Implementing the United Nations "Protect, Respect and Remedy" Framework. New York and Geneva, 2011. https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf
- House of Commons Culture, Media and Sport Select Committee (2018), "'Disinformation and "Fake News": Interim Report.'", July. (https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/36308.htm#_idTextAnchor033.)
- Indictment in the case 18 U.S.C. §§ 2, 371, 1349, 1028A, https://www.justice.gov/file/1035477/download
- Jagland, Thorbjørn (2017), "How Strong are Europe's checks and balances?", Secretary General of the Council of Europe.
- Jeangène Vilmer, J-B, A Escorcia, M Guillaume, and J Herrera (2018), "Information Manipulation: A Challenge for Our Democracies", Report by the Policy Planning Staff (CAPS) of the Ministry for Europe and Foreign Affairs and the Institute for Strategic Research (IRSEM) of the Ministry for the Armed Forces." Paris. (https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf).
- Jourová, Věra (2016), "Code of Conduct on countering illegal hate speech online: First results on implementation", Commissioner for Justice, Consumers and Gender Equality, December. (http://ec.europa.eu/information_society/newsroom/image/document/2016-50/factsheet-code-conduct-8_40573.pdf)
- Le Sénat rejette les propositions de loi sur les "fake news" sans même discuter des textes. 28.07. 2018. https://www.francetvinfo.fr/internet/reseaux-sociaux/facebook/le-senat-rejette-les-propositions-de-loi-sur-les-fake-news-sans-meme-discuter-du-texte_2869295.html.
- Loi du 29 juillet 1881 sur la liberté de la presse. Version consolidée au 13 janvier 2019 https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006070722
- Loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique (1). NOR: ECOX0200175L. Version consolidée au 13 janvier 2019 https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000801164
- Mandates of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression; the Special Rapporteur on the Right to Privacy and the Special Rapporteur on the Promotion and Protection of Human Rights and Fundamental Freedoms while Countering Terrorism, OL OTH 71/2018, Geneva.
- Martens, Bertin, Luis Aguiar, Estrella Gomez-Herrera Frank Mueller-Langer (2018), "The digital transformation of news media and the rise of disinformation and fake news. An economic perspective", JRC Digital Economy Working Paper 2018-02.(https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf)
- Müller, Jan-Werner (2013), "Safeguarding Democracy inside the EU: Brussels and the Future of Liberal Order", *Working Paper* No. 3, Washington DC: Transatlantic Academy.

- Network and Information Security (NIS) Directive (2016/1148); Joint Communication: Resilience, Deterrence and Defence: Building strong cybersecurity for the EU (JOIN(2017) 450 final).
- Nevejans, Nathalie (2016), "European Civil Law Rules in Robotics", the European Parliament's Committee on Legal Affairs (JURI), Brussels, October. (http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf)
- Office of the Director of National Intelligence (2017), "Assessing Russian Activities and Intentions in Recent US Elections.", January 6. (https://www.dni.gov/files/documents/ICA_2017_01.pdf.)
- Parliamentary Assembly of the Council of Europe (2018), Resolution 2212 (2018) on the Protection of editorial integrity.
- Parliamentary Assembly of the Council of Europe (2018), Resolution 2217 (2018) on the Legal challenges related to hybrid war and human rights obligations
- Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online. Brussels, 12.9.2018 COM(2018) 640 final. 2018/0331 (COD) https://ec.europa.eu/commission/sites/beta-political/files/soteu2018-preventing-terrorist-content-online-regulation-640_en.pdf, page 3.
- PROPOSITION DE LOI relative à la lutte contre la manipulation de l'information, 9. Oct. 2018.: http://www.assemblee-nationale.fr/15/ta/ta0180.asp
- PROPOSITION DE LOI relative à la lutte contre les fausses informations, 21.March 2018. http://www.assemblee-nationale.fr/15/propositions/pion0799.asp
- Special Eurobarometer 477. Democracy and elections. Sept-Nov. 2018. http://ec.europa.eu/commfrontoffice/publicopinionmobile/index.cfm/Survey/getSurveyDetail/surveyKy/2198
- The Sounding Board's Unanimous Final Opinion on the So-Called Code of Practice *24 September 2018*. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=54456.
- UK House of Commons Digital (2018), Culture, Media and Sport Committee, Disinformation and 'fake news': Interim Report, 24 July 2018 (https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/363.pdf)
- UK Information Commissioner's Office (2018), Democracy disrupted. Personal information and political influence, 11 July 2018 (https://ico.org.uk/media/2259369/democracy-disrupted-110718.pdf)
- UK's House of Lords, AI in the UK: ready, willing and able?, Select Committee on Artificial Intelligence, Report of Session 2017-19. (https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf)
- UN General Assembly, Resolution adopted Developments in the field of information and telecommunications in the context of international security, [on the report of the First Committee (A/70/455)], A/RES/70/23723, December 2015. (https://unoda-web.s3-accelerate.amazonaws.com/wp-content/uploads/2016/01/A-RES-70-237-Information-Security.pdf)
- UN Guiding Principles on Business and Human Rights Implementing the UN "Protect, Respect and Remedy Framework". 2011. OCHCHR. https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf
- UN, OSCE, OAS, ACHPR (2017), "Joint declaration on freedom of expression and 'Fake News', disinformation and propaganda", A19 and CLD, FOM.GAL/3/17, 3 March. (https://www.osce.org/fom/302796)
- UNESCO (2018), "World Trends in Freedom of Expression and Media Development", Global Report 2017/2018. (http://unesdoc.unesco.org/images/0026/002610/261065e.pdf)
- UNESCO (2018), Journalist, Fake News and Disinformation (https://en.unesco.org/fightfakenews)
- United Nations Human Rights Council (2018), "Report of the Detailed Findings of the Independent International Fact-Finding Mission on Myanmar.", 18 September. (https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_CRP.2.pdf?utm_campaign=The%20Interface&utm_medium=email&utm_source=Revue%20newsletter.)
- United Nations Human Rights Council (2018), "Report of the Independent International Fact-Finding Mission on Myanmar." August 27. (https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/274/54/PDF/G1827454.pdf?OpenElement.)
- United Nations, Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 6 April 2018

- United Nations, Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 6 April 2018, paragraph 30. (https://undocs.org/A/HRC/38/35)
- United Nations' Secretary-General (2015), "Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security", UN GA A/70/174, 22 July.
- United States Department of Justice (2018), "Internet Research Agency Indictment in the United States District Court for the District of Columbia.", February 16. (https://www.justice.gov/file/1035477/download)
- Treaty of the European Union.
- Van Klingeren, M. Orozco, J. van Spanje, C. de Vreese (2015), "Party financing and referendum campaigns in EU Member States", Study for the AFCO Committee, Brussels.
- European Parliament resolution of 27 February 2014 on the situation of fundamental rights in the European Union (2012), European Parliament resolution of 3 July 2013 on the situation of fundamental rights: standards and practices in Hungary (pursuant to the European Parliament resolution of 16 February 2012).
- European Parliament resolution of 8 September 2015 on the situation of fundamental rights in the European Union (2013-2014) (2014/2254(INI)), 8_TA-PROV(2015)0286
- The EU Justice Scoreboard: Towards more effective justice systems in the EU, http://ec.europa.eu/justice/newsroom/effective-justice/news/150309_en.htm.
- Communication from the Commission, Annual Growth Survey 2015, COM (2014) 902 final. For a study of the European semester method refer to 2013 CEPS Study.

Court cases

ECJ

- C-324/09 L'Oréal and others v eBay, judgment of 12 July 2011
- CJEU [GC] decision in the Case C-210/16 Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH, 5 June 2018.
- Opinion of Advocate General (2017), "Case C210/16: Unabhängiges Landeszentrum für Datenschutz
- Joined cases C-402/05 P and C-415/05 P, Opinion of Mr Advocate General Poiares Maduro delivered on 16 January 2008, Yassin Abdullah Kadi and Al Barakaat International Foundation v Council of the European Union and Commission of the European Communities.
- Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH, in the presence of Facebook Ireland Ltd, ertreter des Bundesinteresses beim Bundesverwaltungsgericht", Case Law, CURIA, 24 October.(http://curia.europa.eu/juris/document/document.jsf?text=&docid=195902&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=2546238)

ECtHR

- Animal Defenders International v. the United Kingdom. 48876/08. Judgement of 22.4.2013.
- B.H, M.W, H.P and G.K. v. Austria (application no. 12774/87), decision of the Commission of 12 October 1989;
- Bladet Tromsø and Stensaas v. Norway [GC], 21980/03.
- Communist Party of Germany v. the Federal Republic of Germany, decision of the European Commission on Human Rights of 20 July 1957;
- Delfi v Estonia (App 64569/09) ECtHR 2015.
- ECtHR, Case of Magyar Helsinki Bizottság v. Hungary, Application no. 18030/11, 8 November 2016
- Garaudy v France, Decision of 7. July 2003, no. 65831/01.
- Gaskin v United Kingdom, (1990) 12 E.H.R.R. 36
- Glimmerveen and Hagenbeek v The Netherlands, Judgment of 11 October 1979, nos. 8348/78, 8406/78;
- Guerra v Italy, (1998) 26 E.H.R.R. 357, 14967/89.
- Guseva v. Bulgaria 6987/07.
- Handyside v. the United Kingdom judgment of 7 December 1976, § 49.
- K. v Austria, Judgment of 10 December 1989, no. 12774/87;
- Kommersant Moldovy v. Moldova, Application no. 41827/02.
- LCB v United Kingdom (1998) 27 E.H.R.R. 212,
- Leander v. Sweden, 9248/81

- McGinley and Egan v United Kingdom (1998) 27 E.H.R.R. 1,
- MTE and Index v Hungary (App 22947/13) ECtHR 2016.
- Mustafa Erdoğan and Others v. Turkey, Applications nos. 346/04 and 39779/04, 27 May 2014
- Nachtmann v. Austria, decision of the Commission of 9 September 1998;
- Norwood v UK, Judgment of 16. November 2004, no. 23131/03.
- Oneryildiz v Turkey (2005) 41 E.H.R.R. 325. 48939/99.
- Österreichische Vereinigung zur Erhaltung, Stärkung und Schaffung v. Austria (no. 39534/07, 28 November 2013
- Roche v United Kingdom (2006) 42 E.H.R.R. 599.
- Schimanek v Austria, Judgment of 01 February 2000, nos. 32307/96, 12774/87; H., W., P.
- Schimanek v. Austria, decision of the Court on the admissibility of 1 February 2000.
- Sdruženi Jihočeské Matky v. the Czech Republic (dec.) (2006), 19101/03.
- Seurot v. France no. 57383/00, 18 May 2004 (decision)
- Társaság A Szabadságjogokért v. Hungary, 14 April 2009; 37374/05.
- Times Newspapers Ltd v. United Kingdom (Nos. 1 and 2) (Application nos. 3002/03., 23676/03., March 10. 2009)
- Youth Initiative for Human Rights v. Serbia (2013), 48135/06
- Zana v. Turkey. Application no. 18594/91. 1997. nov. 25.
- Wegrzynowski and Smolczewski v. Poland (Application no. 33846/07., July 16. 2013.)

US court decisions

- Ashcroft v. American Civil Liberties Union, 535 U.S. 564 (2002)
- Federal Communications Commission v. Pacifica Foundation, 438 U.S. 726 (1978)
- First Nat'l Bank v. Bellotti, 435 U.S. 765, 792 (1978).

German court decisions

- Judgment of the Berlin Regional Court dated 16 January 2018, Case no. 16 O 341/15. https://www.vzbv.de/sites/default/files/downloads/2018/02/14/18-02-12_vzbv_pm_facebook-urteil_en.pdf
- BVerfGE 37, 271 – Solange I, 7
- BVerfGE 73, 339 – Solange II,
- BVerfGE 102, 147 – Bananenmarktordnung,
- BVerfGE 89, 155 – Maastricht,
- BVerfGE 123, 267 – Lissabon,
- BVerfG, 21.06.2016 –
- 2 BvR 2728/13;
- 2 BvR 2728/13;
- 2 BvR 2729/13;
- 2 BvR 2730/13.

Journalistic sources and blogs

- #BlueLivesMatter and Beyoncé: Russian Facebook ads hit hot-button US issues, 10 May 2018, (https://www.theguardian.com/us-news/2018/may/10/russia-facebook-ads-us-elections-congress)
- 89up (2018), "89up Releases Report on Russian Influence in the EU Referendum." 89up, February 10. http://89up.org/russia-report.
- Abbruzzese, Jason, and David Ingram (2018), "Iran's Disinformation Campaign Extended to YouTube, Google Says." NBC News, November 11. (https://www.nbcnews.com/tech/tech-news/iran-s-disinformation-campaign-extended-youtube-google-says-n903241)
- Alandete, David (2017), "Russian Network Used Venezuelan Accounts to Deepen Catalan Crisis." El País. November 11, sec. In English. https://elpais.com/elpais/2017/11/11/inenglish/1510395422_468026.html.
- Alandete, David, and Daniel Verdú (2018), "How Russian Networks Worked to Boost the Far Right in Italy.", *El País*. March 1, sec. In English. (https://elpais.com/elpais/2018/03/01/inenglish/1519922107_909331.html.)
- Albright, Jonathan (2016), "The #Election2016 Micro-Propaganda Machine", 18 November 2016, (https://medium.com/@d1gi/the-election2016-micro-propaganda-machine-383449cc1fba)

- Albright, Jonathan (2016), « The #Election2016 Micro-Propaganda Machine », 18 November 2016, (https://medium.com/@d1gi/the-election2016-micro-propaganda-machine-383449cc1fba)
- Albright, Jonathan (2017), "Instagram, Meme Seeding, and the Truth about Facebook Manipulation, Pt. 1." *Medium* (blog), November 8. (https://medium.com/berkman-klein-center/instagram-meme-seeding-and-the-truth-about-facebook-manipulation-pt-1-dae4d0b61db5.)
- Aleem, Z. (2018), « Reddit just shut down nearly 1,000 Russian troll accounts », 11 April 2018, (https://www.vox.com/world/2018/4/11/17224294/reddit-russia-internet-research-agency)
- Atlantic Council's Digital Forensic Research Lab (2018), "#TrollTracker: Facebook Uncovers Iranian Influence Operation", 26 October 2018 (https://medium.com/dfrlab/trolltracker-facebook-uncovers-iranian-influence-operation-d21c73cd71be
- Austria's Post Office under fire over data sharing, France 24, https://www.france24.com/en/20190108-austrias-post-office-under-fire-over-data-sharing
- Bajoran, Donara (2018), "YouTube's Kremlin Disinformation Problem." *DFRLab* (blog), May 3. (https://medium.com/dfrlab/youtubes-kremlin-disinformation-problem-d78472c1b72b)
- Barett, P. et al. (2108), « Combating Russian Disinformation: The Case for Stepping Up the Fight Online ». *NYU Stern Center for Business and Human Rights* (https://disinfoportal.org/wp-content/uploads/ReportPDF/NYU-Stern-CBHR-Combating-Russian-Disinfomration-July-2018-min.pdf)
- Barlow, John Perry (1996), "A Declaration of the Independence of Cyberspace", Davos. (https://www.eff.org/cyberspace-independence)
- Bartlett J et al. (2018), "The Future of Political Campaigning". *DEMOS* (https://www.demos.co.uk/wp-content/uploads/2018/07/The-Future-of-Political-Campaigning.pdf)
- Bashyakarla, V., "Psychometric Profiling: Persuasion by Personality in Elections" (https://ourdataourselves.tacticaltech.org/posts/psychometric-profiling/)
- Bayer, Lili (2016), "Fidesz-Friendly Media Peddling Russian Propaganda." The Budapest Beacon, November 17. (https://budapestbeacon.com/fidesz-friendly-media-peddling-russian-propaganda/)
- BBC (2018), "How WhatsApp Helped Turn a Village into a Mob.", sec. India, July 19.(https://www.bbc.com/news/world-asia-india-44856910.)
- Belise, G. (2018), « Blockchain can eliminate fake news », 27 March 2018 (https://the-blockchain-journal.com/2018/03/27/blockchain-can-eliminate-fake-news/)
- Benedictus, Leo (2016) "Invasion of the Troll Armies: 'Social Media Where the War Goes On.'" *The Guardian*, November 6, 2016 sec. Media. (https://www.theguardian.com/media/2016/nov/06/troll-armies-social-media-trump-russian)
- Bentzen, Naja, (2018), "Online disinformation and the EU's response", *EPRS*, May. (http://www.europarl.europa.eu/RegData/etudes/ATAG/2018/620230/EPRS_ATA(2018)620230_EN.pdf)
- Bing, Christopher (2018), "Exclusive: Twitter Deletes over 10,000 Accounts That Sought To..." *Reuters*, November 2, 2018 (https://www.reuters.com/article/us-usa-election-twitter-exclusive-idUSKCN1N72FA).
- Bort, J. It took only 36 hours for these students to solve Facebook's fake-news problem. Tech Insider 2016, http://www.businessinsider.com/students-solve-facebooks-fake-news-problem-in-36-hours-2016-11
- Brodnig, I. (2017), "7 types of misinformation in the German election", 2 November 2017 (https://firstdraftnews.org/7-types-german-election/)
- BSR (2018), "Human Rights Impact Assessment: Facebook in Myanmar.". (https://fbnewsroomus.files.wordpress.com/2018/11/bsr-facebook-myanmar-hria_final.pdf.)
- Bugorkova, Olga (2015) "Inside the Kremlin's 'Troll Army.'" BBC News, March 19, sec. Europe. (https://www.bbc.com/news/world-europe-31962644).
- Bulckaert, Ninon (2018), "How France Successfully Countered Russian Interference during the Presidential Election." Euractiv.Com (blog), July 17, 2018 (https://www.euractiv.com/section/elections/news/how-france-successfully-countered-russian-interference-during-the-presidential-election/).
- Bump, P.(2018), "Here are the tools that could be used to create the fake news of the future", 12 February 2018 (https://www.washingtonpost.com/news/politics/wp/2018/02/12/here-are-the-tools-that-could-be-used-to-create-the-fake-news-of-the-future/?noredirect=on&utm_term=.19673e753072)
- Carlini, N., "Audio Adversarial Examples" (https://nicholas.carlini.com/code/audio_adversarial_examples)

- Catwalladar, C. (2016), "Google, democracy and the truth about internet search", 4 December 2016 (https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-internet-search-facebook)
- Chafkin, M. (2017), « How Snapchat Has Kept Itself Free of Fake News », 16 October 2017 (https://www.bloomberg.com/news/features/2017-10-26/how-snapchat-has-kept-itself-free-of-fake-news).
- Chi, Zhang (2018), "Study: Chinese-American Immigrants Fall Prey to WeChat's Misinformation Problem." *Columbia Journalism Review*, April 9, 2018 (https://www.cjr.org/tow_center/wechat-misinformation.php).
- Chiuisi F. and Agosti C. (2018), "The Influence Industry Personal Data and Political Influence in Italy", June 2018 (https://ourdataourselves.tacticaltech.org/media/ttc-influence-industry-italy.pdf)
- Collections: VR in Advertising, https://www.adsoftheworld.com/collection/vr_in_advertising
- Collins, Ben (2018), "In Secret Chats, Trolls Struggle to Get Twitter Disinformation Campaigns off the Ground." *NBC News*, November 6, 2018 (https://www.nbcnews.com/tech/tech-news/secret-chats-trolls-struggle-get-twitter-disinformation-campaigns-ground-n931756).
- Committee of Experts on internet intermediaries (MSI-NET)(2017), "Algorithms and Human Rights – Study on the Human Rights Dimensions of Automated Data Processing Techniques and possible regulatory implications", Council of Europe Study DGI (2017)12. (https://rm.coe.int/algorithms-and-human-rights-study-on-the-human-rights-dimension-of-aut/1680796d10)
- Committee of experts on media pluralism and transparency of media ownership (MSI-MED)(2018), "Internet and Electoral Campaigns – Study on the use of internet in electoral campaigns", Council of Europe study, DGI(2017)11, Rapporteur Damian Tambini, April. (https://rm.coe.int/use-of-internet-in-electoral-campaigns-/16807c0e24)
- Constine, J. (2017), "Facebook now has 2 billion monthly users… and responsibility", 27 June 2017 (https://techcrunch.com/2017/06/27/facebook-2-billion-users/)
- Constine, J., Facebook restricts APIs, axes old Instagram platform amidst scandals, 4 April 2018, https://techcrunch.com/2018/04/04/facebook-instagram-api-shut-down/
- Cottee, Simon (2018), "Can Facebook Really Drive Violence?" *The Atlantic*, September. (https://www.theatlantic.com/international/archive/2018/09/facebook-violence-germany/569608/).
- Coyne, Ellen (2018). "Facebook Won't Stop Disputed Pro Life Ad." *The Times*, March 27, sec. Ireland. https://www.thetimes.co.uk/article/facebook-wont-stop-disputed-pro-life-ad-qg0ghgm86.
- Cvetovska, Saska, Aubrey Belford, Craig Silverman, and J. Lester Feder (2018), "The Secret Players Behind Macedonia's Fake News Sites." *OCCRP*, July 18, 2018 (https://www.occrp.org/en/spooksandspin/the-secret-players-behind-macedonias-fake-news-sites)
- Dearden, Lizzie (2017), "Pro-Brexit Twitter Account with 100,000 Followers Could Be Part of Russian 'Disinformation Campaign.'" *The Independent*, August (http://www.independent.co.uk/news/uk/home-news/david-jones-pro-brexit-ukip-twitter-account-russia-fake-bot-troll-trump-disinformation-followers-a7920181.html).
- DeMers, J. (2016), "What Are 'Rich Answers' And How Do They Affect SEO?", 14 August 2016 (https://www.forbes.com/sites/jaysondemers/2016/08/14/what-are-rich-answers-and-how-do-they-affect-seo/#73165da534df)
- DeMers, J. (2017), "7 Predictions For The Shape Of Content Marketing In 2020", 13 April 2017, (https://www.forbes.com/sites/jaysondemers/2017/04/13/7-predictions-for-the-shape-of-content-marketing-in-2020/#2278fa86177d)
- Dewey, C. (2015), « This is what happens when you create an online community without any rules », 13 January 2015 (https://www.washingtonpost.com/news/the-intersect/wp/2015/01/13/this-is-what-happens-when-you-create-an-online-community-without-any-rules/)
- Do You Dare Talk Politics with Family at Holiday Gatherings? by Degges-White, Suzanne, 27. Nov. 2017. https://www.psychologytoday.com/us/blog/lifetime-connections/201711/do-you-dare-talk-politics-family-holiday-gatherings and dozens of other sources.
- Drapeau, Olivier (2017), "I Stand with CEU", video (https://www.youtube.com/watch?v=FoMd_Bxn5rg&feature=youtu.be&fbclid=IwAR2uHR_3m4w5DbB1ndIBRbvYHfHT_eIB9_V9kaaXvfi7_X16dDYclajnDEg).

- Dunai, Márton (2018), "Hungary Approves 'STOP Soros' Law, Defying EU, Rights Groups.", *Reuters*, June 20, 2018. (https://www.reuters.com/article/us-hungary-soros/hungary-approves-stop-soros-law-defying-eu-rights-groups-idUSKBN1JG1VN).
- Dwoskin, Elizabeth, and Tony Romm (2018), "Facebook Purged over 800 U.S. Accounts and Pages for Pushing Political Spam", *Washington Post*, October 11, 2018 (https://www.washingtonpost.com/technology/2018/10/11/facebook-purged-over-accounts-pages-pushing-political-messages-profit/).
- Dwyer, Craig (2018). "How Digital Threats to Democracy Were Tackled During Ireland's Abortion Referendum." Media Policy Project, London School of Economy (blog), July 10. http://blogs.lse.ac.uk/mediapolicyproject/2018/07/10/how-digital-threats-to-democracy-were-tackled-during-irelands-abortion-referendum/.
- Electionland (2016), "Viral 'Rigged' Voting Machine Video Actually User Error", 8 November 2016 (https://projects.propublica.org/electionland/pennsylvania/viral-rigged-voting-machine-actually-user-error/)
- Euractiv (2015), "Hungarian Official Admits Campaign to Generate Hate against Migrants.", September 7. (https://www.euractiv.com/section/justice-home-affairs/news/hungarian-official-admits-campaign-to-generate-hate-against-migrants/.)
- Euractiv (2018), "Whistleblower: Cambridge Analytica Shared Data with Russia.", May 17. (https://www.euractiv.com/section/global-europe/news/whistleblower-cambridge-analytica-shared-data-with-russia/)
- Euractive (2018), "Mark Zuckerberg's full meeting with EU Parliament leaders", Video, *Euractive*, 22 May. (https://www.youtube.com/watch?v=o0zdBUOrhG8)
- European Center for Press and Media Freedom (2018), "Tackling fake news, the Italian way", 22 May 2018 (https://www.rcmediafreedom.eu/Tools/Legal-Resources/Tackling-fake-news-the-Italian-way)
- Facebook (2011),"(2) Case Study: Reaching Voters with Facebook Ads (Vote No on 8).", August 16. (https://www.facebook.com/notes/government-and-politics-on-facebook/case-study-reaching-voters-with-facebook-ads-vote-no-on-8/10150257619200882/)
- Facebook policy chief admits hiring PR firm to attack George Soros. The Guardian. 22. Nov. 2018. https://www.theguardian.com/technology/2018/nov/21/facebook-admits-definers-pr-george-soros-critics-sandberg-zuckerberg
- Facebook Will Not Be Accepting Referendum Related Ads from Advertisers Based Outside of Ireland. Facebook Ireland (blog), May 8, 2018. https://www.facebook.com/notes/facebook-dublin/facebook-will-not-be-accepting-referendum-related-ads-from-advertisers-based-out/10156398786998011/.
- Fazlioglu, M., Analyzing changes in DPA Income and Staff, from 2011 to 2016, 11 December 2017, https://iapp.org/news/a/analyzing-changes-in-dpa-income-and-staff-2011-2016/
- FireEye (2018), "Suspected Iranian Influence Operation Leverages Network of Inauthentic News Sites & Social Media Targeting Audiences in U.S., UK, Latin America, Middle East." Milpitas, CA: FireEye Intelligence, August 1. (https://www.fireeye.com/blog/threat-research/2018/08/suspected-iranian-influence-operation.html).
- Fitzgerald, Cormac (2018). "Concerns over Mystery Facebook Ads Claiming to Offer 'unbiased Facts' on Eighth Referendum." *The Journal*, May 1. http://www.thejournal.ie/eighth-referendum-ads-3986039-May2018/.
- Frier, Sarah, and Camillo Gulia (2018), "WhatsApp Bans More Than 100,000 Accounts in Brazil Election," October 19, 2018 (https://www.bloomberg.com/news/articles/2018-10-19/whatsapp-bans-more-than-100-000-accounts-in-brazil-election).
- Frum, David (2018), "Trump's Crisis of Legitimacy." *The Atlantic*, July 17, 2018 (https://www.theatlantic.com/ideas/archive/2018/07/is-trumps-presidency-legitimate/565451/).
- Fuchs, M.(2017), "Why Social Bots Threaten Our Democracy" in *Das Netz* – English Edition: Digitalization and Society, July 2017 (http://irights-media.de/publikationen/das-netz-english-edition/)
- GDPR misuse in Romania: "independence of DPA" and "transparency" – keywords or buzzwords? 17 December 2018, https://www.gdprtoday.org/gdpr-misuse-in-romania-independence-of-dpa-and-transparency-keywords-or-buzzwords/

- Gensing, Patrick – Lena Kampf: Wie Trolle im Wahlkampf manipulierten. 01. 03. 2018. https://faktenfinder.tagesschau.de/inland/manipulation-wahlkampf-101.htmlGerstein, Josh (2018), "U.S. Brings First Charge for Meddling in 2018 Midterm Elections.", *Politico*, October 19, 2018 (https://politi.co/2Ajbubq).
- Ghattas, K. (2017), "How the Muslim World Lost the Freedom to Choose", *Foreign Policy*, 20 October.(https://foreignpolicy.com/2017/10/20/how-the-muslim-world-lost-the-freedom-to-choose/)
- Ghosh, D., and Scott, B., Digital Deceit. The Technologies Behind Precision Propaganda on the Internet, 23 January 2018, https://www.newamerica.org/public-interest-technology/policy-papers/digitaldeceit/
- Ghosh, S., Facebook approved fake political ads 'paid for' by Cambridge Analytica, 3 October 2018, https://nordic.businessinsider.com/facebook-approved-political-ads-paid-for-by-cambridge-analytica-2018-10?r=US&IR=T
- Gilbert, David (2018), "Iran Is Running an Online Disinformation Campaign on the Scale of Russia's Troll Farm.", *Vice News*, August 22, 2018 (https://news.vice.com/en_ca/article/594ekk/iran-russia-facebook-twitter-disinformation).
- Gleicher, Nathaniel (2018), "Election Update." Facebook Newsroom (blog), November 5, 2018 (https://newsroom.fb.com/news/2018/11/election-update/).
- Gleicher, Nathaniel (2018), "Taking Down Coordinated Inauthentic Behavior from Iran | Facebook Newsroom," October 26, 2018 (https://newsroom.fb.com/news/2018/10/coordinated-inauthentic-behavior-takedown/).
- Gleicher, Nathaniel, and Oscar Rodriguez (2018), "Removing Additional Inauthentic Activity from Facebook." Facebook Newsroom (blog), October 11, 2018 (https://newsroom.fb.com/news/2018/10/removing-inauthentic-activity/).
- Gomes, B. (2017), Our latest quality improvements for Search, 25 April 2017 (https://blog.google/products/search/our-latest-quality-improvements-search/)
- Google Needs To Blacklist 4chan During National Crises, 3 October 2017, https://www.forbes.com/sites/fruzsinaeordogh/2017/10/03/google-needs-to-blacklist-4chan-during-national-crises/#58623f4b3dcd
- Gottfried, Jeffrey, Michael Barthel, and Amy Mitchell (2017), "Trump, Clinton Voters Divided in Their Main Source for Election News | Pew Research Center," January 18, 2017 (http://www.journalism.org/2017/01/18/trump-clinton-voters-divided-in-their-main-source-for-election-news/).
- Grabenwarter, Christoph, *Constitutional Resilience, VerfBlog,* 6 December 2018, https://verfassungsblog.de/constitutional-resilience/.
- Grassegger, H. and Krogerus, M. (2017), "The Data That Turned the World Upside Down", 28 January 2017 (https://publicpolicy.stanford.edu/news/data-turned-world-upside-down)
- Greenup, S. (2018), « Catalogue of all projects working to solve Misinformation and Disinformation », 9 June 2018 (https://misinfocon.com/catalogue-of-all-projects-working-to-solve-misinformation-and-disinformation-f85324c6076c)
- Grimm, Cf. Dieter, *How can a democratic constitution survive an autocratic majority*?, VerfBlog, 13 December 2018, https://verfassungsblog.de/how-can-a-democratic-constitution-survive-an-autocratic-majority/.
- Guess, Andrew, Brendan Nyhan, and Jason Reifler (2018), "Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 U.S. Presidential Campaign." (http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf).
- Guess, Andrew, Jonathan Nagler and Joshua Tucker: Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Sci Adv* **5** (1), eaau4586. DOI: 10.1126/sciadv.aau4586
- Győri, Lóránt, Péter Krekó, Jakub Janda, and Bernhard Weidinger (2017), "Does Russia Interfere in Czech, Austrian and Hungarian Elections? A Study by Political Capital, European Values Think-Tank in Cooperation with DöW". (https://www.kremlinwatch.eu/userfiles/western_experiences_eastern_vulnerabilities_20171012_15273208786863.pdf).

- Harding, Luke (2018), "Revealed: Details of Exclusive Russian Deal Offered to Arron Banks in Brexit Run-Up." *The Guardian*, August 9, sec. UK news. (https://www.theguardian.com/uk-news/2018/aug/09/revealed-detail-of-exclusive-russian-deal-offered-to-arron-banks-in-brexit-run-up).
- Hautala, L. (2018), « Reddit: Russian propaganda spread on our site before 2016 election », 5 March 2018, https://www.cnet.com/news/reddit-russian-propaganda-spread-on-our-site-before-2016-election/)
- Hern, Alex (2017) "Macron Hackers Linked to Russian-Affiliated Group behind US Attack." *The Guardian*, May 8, sec. World news. (https://www.theguardian.com/world/2017/may/08/macron-hackers-linked-to-russian-affiliated-group-behind-us-attack).
- Hern, Alex (2017), "Facebook and Twitter are being used to manipulate public opinion" – report, 19 June 2017 (https://www.theguardian.com/technology/2017/jun/19/social-media-proganda-manipulating-public-opinion-bots-accounts-facebook-twitter)
- Hochmann, T. (2018) , "Shedding light or shooting in the dark – how to define fake news", 5 September 2018 (https://verfassungsblog.de/shedding-light-or-shooting-in-the-dark-how-to-define-fake-news/)
- Holt, Thomas (2018), "Busting Russia's fake news the European Union way", *The Conversation*, 29 March.(https://theconversation.com/busting-russias-fake-news-the-european-union-way-93712).
- Horowitz, Jason (2017) "Italy, Bracing for Electoral Season of Fake News, Demands Facebook's Help." *The New York Times*, November 24 (https://www.nytimes.com/2017/11/24/world/europe/italy-election-fake-news.html).
- Huberman, B.A., Romero, D.M., and Wu, F., Social networks that matter: Twitter under the microscope. 2008, https://arxiv.org/pdf/0812.1045.pdf
- Human Rights Watch (2018), German: flawed social media law, 14 February 2018 (https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law)
- IBM (2018), The science behind the service, 13 May (https://console.bluemix.net/docs/services/personality-insights/science.html#science)
- Illing, S. (2017), "Cambridge Analytica, the shady data firm that might be a key Trump-Russia link, explained", 22 October 2017 (https://www.vox.com/policy-and-politics/2017/10/16/15657512/cambridge-analytica-trump-kushner-flynn-russia)
- Information Commissioner's Office, Investigation into the use of data analytics in political campaigns, a report to Parliament 6 November 2018, https://ico.org.uk/media/2260277/investigation-into-the-use-of-data-analytics-in-political-campaigns-20181107.pdf.
- Ingham, L. (2015), "Holograms, VR and in-game ads: meet the future of the political campaign", 7 September 2015 (https://www.factor-tech.com/feature/holograms-vr-and-in-game-ads-meet-the-future-of-the-political-campaign/)
- Ingram, David (2017), "Facebook Says 126 Million Americans May Have Seen Russia-Linked Political Posts." Reuters, October 30. (https://www.reuters.com/article/us-usa-trump-russia-socialmedia/facebook-says-126-million-americans-may-have-seen-russia-linked-political-posts-idUSKBN1CZ2OI).
- Ingram, David (2018), "Attacks on Jewish People Rising on Instagram and Twitter, Researchers Say." *NBC News*. Accessed November 11. (https://www.nbcnews.com/tech/tech-news/attacks-jewish-people-rising-instagram-twitter-researchers-say-n925086).
- Ingram, Mathew (2017), "In Some Countries, Fake News on Facebook Is a Matter of Life and Death." Columbia Journalism Review, November 21. (https://www.cjr.org/analysis/facebook-rohingya-myanmar-fake-news.php).
- Isaac, Mike, and Kevin Roose (2018), "Disinformation and Fake News Spreads over WhatsApp Ahead of Brazil's Presidential Election." *The New York Times*, October 20. (https://www.independent.co.uk/news/world/americas/brazil-election-2018-whatsap-fake-news-presidential-disinformation-a8593741.html).
- Jäger, Silke: Die betrügerischen Fake-Accounts, die dich wütend machen sollen. 21. Januar 2019. https://krautreporter.de/2762-die-betrugerischen-fake-accounts-die-dich-wutend-machen-sollen.Kalia, Ammar, and Caelainn Barr Angela Giuffrida in Rome (2018). "Revealed: How Italy's Populists Used Facebook to Win Power." *The Guardian*, December 17, sec. World news. https://www.theguardian.com/world/2018/dec/17/revealed-how-italy-populists-used-facebook-win-election-matteo-salvini-luigi-di-maio.

- Katz, A.J. (2018), "October 2018 Ratings: Fox News Channel Averaged More Total Viewers Than CNN and MSNBC Combined," October 30 (https://adweek.it/2CP7MrC).
- Keszthelyi, Christian (2016), "Xenophobia Skyrocketing in Hungary, Surveys Reveal." Budapest Business Journal, November 17. (https://bbj.hu/budapest/xenophobia-skyrocketing-in-hungary-surveys-reveal_124920).
- Kingsley, Patrick (2018), "Hungary's Leader Was Shunned by Obama, but Has a Friend in Trump", *NY Times*, 15 August. (https://www.nytimes.com/2018/08/15/world/europe/hungary-us-orban-trump.html)
- Kleinwächter, Wolfgang (2013), 'Internet governance outlook 2013: "Cold internet war" or "peaceful internet coexistence"?', *CircleID*, 3 January. (http://www.circleid.com/posts/20130103_internet_governance_outlook_2013/)
- Kleinwächter, Wolfgang (2018), "Internet Governance Outlook 2018: Preparing for Cyberwar or Promoting Cyber Détente?", *CircleID*, January.
- Klingová, Katarína, Daniel Milo, Veronika Víchová, Lóránt Győri, and Patrik Szicherle n.d. "Information War Monitor for Central Europe: Pro-Kremlin Disinformation Outlets Have a Favourite French Presidential Candidate." GLOBSEC. https://www.globsec.org/publications/information-war-monitor-central-europe-pro-kremlin-disinformation-outlets-favourite-french-presidential-candidate/.
- Kolozsi Ádám (2016): Sosem látott mértékű a magyarországi idegenellenesség, https://index.hu/tudomany/2016/11/17/soha_nem_latott_merteku_az_idegenellenesseg_magyarorszagon/ (letöltés: 2018. XI. 1.).
- Kőműves, Anita (2018), "Target or Ally? Hungary Faces the Elections Battle." *Vsquare.org*, March 4.. (https://vsquare.org/russia-target-or-ally-hungary-faces-the-elections-battle/).
- Koncewicz, Tomasz Tadeusz, The Democratic Backsliding and the European constitutional design in error. When will HOW meet WHY?, VerfBlog, 18 December 2018, https://verfassungsblog.de/the-democratic-backsliding-and-the-european-constitutional-design-in-error-when-how-meets-why/.
- Korzak, Elaine (2017), "UN GGE on Cybersecurity: The End of an Era? The Diplomat", *The Diplomat,* 31 July. (https://thediplomat.com/2017/07/un-gge-on-cybersecurity-have-china-and-russia-just-made-cyberspace-less-safe/).
- Kottasova, Ivana (2016), "Did Fake News Influence Italy's Referendum?" *CNNMoney*, December 5. (https://money.cnn.com/2016/12/05/media/fake-news-italy-referendum/index.html).
- Krejčí, Markéta, Veronika Víchová, and Jakub Janda (2018) "The Role of the Kremlin's Influence and Disinformation in the Czech Presidential Elections." *The Kremlin Watch Report*, European Values, January 29 https://www.kremlinwatch.eu/userfiles/the-role-of-the-kremlin-s-influence-and-disinformation-in-the-czech-presidential-elections_15263778517686.pdf.
- Kumm, Matthias, *How populist authoritarian nationalism threatens constitutionalism or: Why constitutional resilience is a key issue of our time, VerfBlog,* 6 December 2018, https://verfassungsblog.de/how-populist-authoritarian-nationalism-threatens-constitutionalism-or-why-constitutional-resilience-is-a-key-issue-of-our-time/.
- Lewis, P. (2018), "Fiction is outperforming reality": how YouTube's algorithm distorts truth », 2 February 2018 (https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth)
- Lin, H. (2018), "The Danger of Deep Fakes: Responding to Bobby Chesney and Danielle Citron", 27 February 2018 (https://www.lawfareblog.com/danger-deep-fakes-responding-bobby-chesney-and-danielle-citron)
- Lomas, N. (2018), « Tumblr confirms 84 accounts linked to Kremlin trolls », 23 March 2018 (https://techcrunch.com/2018/03/23/tumblr-confirms-84-accounts-linked-to-kremlin-trolls/)
- Lyons, T. et al. (2018), "GDPR and Blockchain". Thematic report prepared by the European Union Blockchain Observatory and Forum (https://www.eublockchainforum.eu/sites/default/files/reports/20181016_report_gdpr.pdf?width=1024&height=800&iframe=true)
- Magenta, Matheus, Juliana Gragnani, and Felipe Souza (2018), "How WhatsApp Is Being Abused in Brazil's Elections," October 24, sec. Technology. (https://www.bbc.com/news/technology-45956557).
- Mamiit, A. (2018), "Alexa, Siri, And Google Assistant Follow Malicious Voice Commands Hidden In Music", 11 May 2018 (https://www.techtimes.com/articles/227465/20180511/alexa-siri-and-google-assistant-follow-malicious-voice-commands-hidden-in-music.htm)

- Maurer, Erik, and Tim Brattberg (2018), "Russian Election Interference: Europe's Counter to Fake News and Cyber Attacks." Carnegie Endowment for International Peace, May 23. (https://carnegieendowment.org/2018/05/23/russian-election-interference-europe-s-counter-to-fake-news-and-cyber-attacks-pub-76435).
- May, B. (2018), "Hey Google, How Do I Optimize For Voice Search?", 20 August 2018 (https://www.forbes.com/sites/forbesagencycouncil/2018/08/20/hey-google-how-do-i-optimize-for-voice-search/#7621acbf3800)
- Mayer, Jane (2018), "How Russia Helped Swing the Election for Trump.", *New Yorker*, September 24. (https://www.newyorker.com/magazine/2018/10/01/how-russia-helped-to-swing-the-election-for-trump).
- McAthy, Rachel (2013), "How Russia Today Reached One Billion Views on YouTube | Media News.", *Journalism.co.uk*, June 4 (https://www.journalism.co.uk/news/how-russia-today-reached-one-billion-views-on-youtube/s2/a553152/).
- McGonagle, Tarlach (2017), ''Fake news'': False fears or real concerns? Netherlands Quarterly of Human Rights", Vol. 35(4), p. 203–209.
- McSorley, Christina (2018). "Google Abortion Poll Ban 'Outrageous.'" *BBC.com*, May 10, sec. Europe. https://www.bbc.com/news/world-europe-44067607.
- Media Bias Fact Check, "Methodology" (https://mediabiasfactcheck.com/methodology/)
- Media Bias/Fact Check, "Questionable source", Voice of Europe, *Media Bias/Fact Check*. (https://mediabiasfactcheck.com/voice-of-europe/)
- Mégane, Fastrez (2018), "A Russian Influence on the French Elections?", EU Disinfo Lab. (https://spark.adobe.com/page/fJxCYVGj8d5Fk/.)
- Meister, Stefan (2016), "The "Lisa case": Germany as a target of Russian disinformation", *NATO Review Magazine*. (https://www.nato.int/docu/review/2016/also-in-2016/lisa-case-germany-target-russian-disinformation/EN/index.htm)
- Melzer, Nils (2011), "Cyberwarfare and International Law". (http://unidir.org/files/publications/pdfs/cyberwarfare-and-international-law-382.pdf).
- Menn, Joseph (2017), "Exclusive: Russia Used Facebook to Try to Spy on Macron Campaign – Sources." *Reuters*, July 27 (https://www.reuters.com/article/us-cyber-france-facebook-spies-exclusive/exclusive-russia-used-facebook-to-try-to-spy-on-macron-campaign-sources-idUSKBN1AC0EI).
- Mérték Médiaelemző Műhely (2016–2018): Szúrópróba, http://mertek.eu/wp-content/uploads/2018/07/Sz%C3%BAr%C3%B3pr%C3%B3ba-25.pdf (letöltés: 2018. XI. 4.).
- Meserole C. and Polyakova A. (2018), "Disinformation Wars", 25 may 2018, https://foreignpolicy.com/2018/05/25/disinformation-wars/ and Greenspan, G., The Blockchain Immutability Myth, 4 May 2017, https://www.multichain.com/blog/2017/05/blockchain-immutability-myth/).
- Mezzofiore, Gianluca. "Wirathu's 'Buddhist Woman Raped' Facebook Post Stokes Anti-Muslim Violence in Mandalay.", *International Business Times UK*, July 2 (https://www.ibtimes.co.uk/wirathus-buddhist-woman-raped-facebook-post-stokes-anti-muslim-violence-mandalay-1455069).
- Ministry of the Foreign Affairs of Denmark (2018), Strengthened safeguards against foreign influence on Danish elections and democracy, 7 September 2018 (http://um.dk/en/news/NewsDisplayPage/?newsID=1DF5ADBB-D1DF-402B-B9AC-57FD4485FFA4)
- Mohan, M. (2017), « Macron Leaks: the anatomy of a hack », 9 May 2017 (https://www.bbc.com/news/blogs-trending-39845105).
- Moon, Mariella (2018), "Facebook's Free Basics Quietly Pulled from Myanmar, Other Markets." *Engadget*, May 2 (https://www.engadget.com/2018/05/02/facebook-free-basics-quietly-pulled-myanmar/).
- Mozur, Paul (2018), "A Genocide Incited on Facebook, With Posts From Myanmar's Military." *The New York Times*, October 18, sec. Technology. (https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html).
- Mr Juncker, be Bob the Builder. 25. Nov. 2016. https://euobserver.com/opinion/136030
- Nardelli, Alberto, and Craig Silverman (2016), "Italy's Most Popular Political Party Is Leading Europe In Fake News And Kremlin Propaganda." *BuzzFeed,* 29 November. https://www.buzzfeed.com/albertonardelli/italys-most-popular-political-party-is-leading-europe-in-fak.

- Nardelli, Alberto, and Craig Silverman (2017), "One Of The Biggest Alternative Media Networks In Italy Is Spreading Anti-Immigrant News And Misinformation On Facebook." *BuzzFeed*, 21 November, https://www.buzzfeed.com/albertonardelli/one-of-the-biggest-alternative-media-networks-in-italy-is.
- National Endowment for Democracy, "The Big Question: how will 'deepfakes' and emerging technology transform disinformation" (https://www.ned.org/the-big-question-how-will-deepfakes-and-emerging-technology-transform-disinformation/)
- Nemer, David (2018), "The Three Types of WhatsApp Users Getting Brazil's Jair Bolsonaro Elected." *The Guardian*, October 25, sec. World news. (https://www.theguardian.com/world/2018/oct/25/brazil-president-jair-bolsonaro-whatsapp-fake-news).
- Newton, C.(2018), « How white supremacists are thriving on YouTube », 19 September 2018 (https://www.theverge.com/2018/9/19/17876892/youtube-extremism-report-rebecca-lewis-data-society)
- Nielsen, Nikolaj (2017), "Bannon's The Movement to launch with January summit", *Euobserver*, 22 October, (https://euobserver.com/political/143125)
- NiemanLab (2018), "Disinformation gets worse" (http://www.niemanlab.org/2017/12/disinformation-gets-worse/)
- Nikolaj Nielsen, Bannon's The Movement to launch with January summit, 22 October 2017, https://euobserver.com/political/143125.
- Nikolaj Nielsen, Bannon's The Movement to launch with January summit, 22 October 2017, https://euobserver.com/political/143125.
- Nimmo, B. (2017), « Spread it on Reddit: How a fake story about Angela Merkel led to a far-right cluster on Reddit », 12 February 2017 (https://www.stopfake.org/en/spread-it-on-reddit-how-a-fake-story-about-angela-merkel-led-to-a-far-right-cluster-on-reddit/).
- Nimmo, Ben (2017), "#ElectionWatch: Scottish Vote, Pro-Kremlin Trolls." DFRLab (blog), December 13. (https://medium.com/dfrlab/electionwatch-scottish-vote-pro-kremlin-trolls-f3cca45045bb).
- Nimmo, Ben (2017), "The Kremlin's Audience in France." DFRLab (blog), April 14. (https://medium.com/dfrlab/the-kremlins-audience-in-france-884a80515f8b).
- Nimmo, Ben, and Anna Pellegatta (2018), "#ElectionWatch: Italy's Self-Made Bots." DFRLab (blog), January 25. (https://medium.com/dfrlab/electionwatch-italys-self-made-bots-200e2e268d0e).
- Nimmo, Ben, and Graham Brookie (2018), "#TrollTracker: Facebook Uncovers Iranian Influence Operation." Medium (blog), October 26. (https://medium.com/dfrlab/trolltracker-facebook-uncovers-iranian-influence-operation-d21c73cd71be).
- OpenSecrets.org, Center for Responsive Politics, Cost of Election. www.opensecrets.org/overview/cost.php
- Ordway, Denise-Marie (2018), "Facebook and the newsroom: 6 questions for Siva Vaidhyanathan", *Journalist's Resource*, 12 September. (https://journalistsresource.org/studies/society/social-media/facebook-siva-vaidhyanathan-news?utm_source=JR-email&utm_medium=email&utm_campaign=JR-email&quot;%20target=&quot;_self&quot)
- Ordway, Denise-Marie (2018), "Top 10 research studies on digital news, social media in 2017", *Journalist's Resource*, 5 January. (https://journalistsresource.org/studies/society/news-media/top-10-research-studies-digital-news-social-media-2017)
- Orlowski, R. (2018), « Facebook says deleted many fake accounts in German campaign », 27 September 2018 (https://www.reuters.com/article/us-germany-elections-facebook/facebook-says-deleted-many-fake-accounts-in-german-campaign-idUSKCN1C22Q5)
- Owen, L. (2018), « News in a disintegrating reality: Tow's Jonathan Albright on what to do as things crash around us », 28 February 2018 (http://www.niemanlab.org/2018/02/news-in-a-disintegrating-reality-tows-jonathan-albright-on-what-to-do-as-things-crash-around-us/)
- Owen, L.(2017), « Instagram is also a huge source of Russian propaganda on social media (Pinterest's not safe either) », 9 November 2017 (http://www.niemanlab.org/2017/11/instagram-is-also-a-huge-source-of-russian-propaganda-on-social-media-pinterests-not-safe-either/)
- Palma, Bethania (2017), "Germany Election So Far Unaffected by 'Fake News.'", *Snopes.com*, September 23. (https://www.snopes.com/news/2017/09/23/german-election-so-far-unaffected-by-fake-news/).

- Panetta, K. (2017), "Gartner Top Gartner Top Strategic Predictions for 2018 and Beyond. Gartner", 3 October 2017 (https://www.gartner.com/smarterwithgartner/gartner-top-strategic-predictions-for-2018-and-beyond/).
- Panyi, Szabolcs, and András Dezső (2017), "We Are Not Paid Agents of Russia, We Do It out of Conviction.", *Index*, January 30 (http://index.hu/belfold/2017/01/30/we_are_not_paid_agents_of_russia_we_do_it_out_of_conviction/)
- Patrick Kingsley, Hungary's Leader Was Shunned by Obama, but Has a Friend in Trump, 15 August 2018, https://www.nytimes.com/2018/08/15/world/europe/hungary-us-orban-trump.html.
- Patrick Kingsley, Hungary's Leader Was Shunned by Obama, but Has a Friend in Trump, 15 August 2018, https://www.nytimes.com/2018/08/15/world/europe/hungary-us-orban-trump.html.
- Peinado, F. (2018), « The business of digital manipulation in Spain », 24 May 2018 (https://elpais.com/elpais/2018/05/24/inenglish/1527147309_000141.html)
- Penzenstadler, Nick, Brad Heath, and Jessica Guynn (2018), "What We Found in Facebook Ads by Russians Accused of Election Meddling." *USA Today*, May 13. (http://www.usatoday.com/story/news/2018/05/11/what-we-found-facebook-ads-russians-accused-election-meddling/602319002/).
- Perez-Rosas, V. et al., Automatic Detection of Fake News, Proceedings of the 27th International Conference on Computational Linguistics, pages 3391–3401, Santa Fe, New Mexico, USA, August 20-26, 2018, http://aclweb.org/anthology/C18-1287
- Plucinska, Joanna, and Mark Scott (2018), "How Italy Does Putin's Work.", *Politico*, March 3. (https://www.politico.eu/article/italy-election-fake-news-sunday-bufale-misinformation-vladimir-putin-russia/).
- POLYGRAPH.info (2018), "Disinfo News: Media Consolidation and 'Fake News' Plague Hungary in Orban Era." *POLYGRAPH.info*, April 26, 2018. (https://www.polygraph.info/a/fake-news-in-hungary/29194591.html).
- Popken, Ben (2018), "How WhatsApp Became Linked to Mob Violence and Fake News — and Why It's Hard to Stop.", *NBC News*, November 2. (https://www.nbcnews.com/tech/tech-news/how-whatsapp-became-linked-mob-violence-fake-news-why-it-n929981).
- Poynter, "A guide to anti-misinformation actions around the world" (https://www.poynter.org/news/guide-anti-misinformation-actions-around-world)
- Provost, Claire, and Lara Whyte (2018). "Foreign and 'alt-Right' Activists Target Irish Voters on Facebook Ahead of Abortion Referendum." openDemocracy, April 25. https://www.opendemocracy.net/5050/claire-provost-lara-whyte/north-american-anti-abortion-facebook-ireland-referendum.
- Reichel, C. (2018), "3 quick tips for debunking hoaxes in a hurricane", *Journalists' Resources*, 14 September. (https://journalistsresource.org/tip-sheets/reporting/hurricane-florence-hoax-tips?utm_source=JR-email&utm_medium=email&utm_campaign=JR-email&quot;%20target=&quot;_self&quot)
- Removal of online hate speech in numbers. 2018. http://blogs.lse.ac.uk/mediapolicyproject/2018/08/16/removals-of-online-hate-speech-in-numbers/
- Rogers, K. and Bromwich, J. (2016), « The Hoaxes, Fake News and Misinformation We Saw on Election Day », 9 November 2016 (https://www.nytimes.com/2016/11/09/us/politics/debunk-fake-news-election-day.html)
- Satariano, Adam (2018). "Ireland's Abortion Referendum Becomes a Test for Facebook and Google." *The New York Times*, May 25, sec. Technology. https://www.nytimes.com/2018/05/25/technology/ireland-abortion-vote-facebook-google.html.
- Schmitt, Michael – Liis Vihul: International Cyber Law Politicized: The UN GGE's Failure to Advance Cyber Norms. Just Security Blog. 2017. https://www.justsecurity.org/42768/international-cyber-law-politicized-gges-failure-advance-cyber-norms/
- Seddon, Max (2014), "Documents Show How Russia's Troll Army Hit America.", *BuzzFeed News*, June 2. (https://www.buzzfeednews.com/article/maxseddon/documents-show-how-russias-troll-army-hit-america).
- Serhan, Yasmeen (2018), "Italy Scrambles to Fight Misinformation Ahead of Its Elections.", *The Atlantic*, February 24.(https://www.theatlantic.com/international/archive/2018/02/europe-fake-news/551972/).
- Shane, Scott, and Mark Mazzetti (2018), "The Plot to Subvert an Election: Unraveling the Russia Story So Far.", *The New York Times*, September 20, sec. U.S.

(https://www.nytimes.com/interactive/2018/09/20/us/politics/russia-interference-election-trump-clinton.html).

- Shu, K., et al., Fake News Detection on Social Media: A Data Mining Perspective, https://arxiv.org/pdf/1708.01967.pdf
- Silverman, Alberto Nardelli, Craig (2016), "Italy's Most Popular Political Party Is Leading Europe In Fake News And Kremlin Propaganda.", *BuzzFeed*, November 29 (https://www.buzzfeed.com/albertonardelli/italys-most-popular-political-party-is-leading-europe-in-fak).
- Silverman, C. and Alexander, L. How Teens In The Balkans Are Duping Trump Supporters With Fake News. 2016, https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trumpmisinfo?utm_term=.jxX7xRvNr0#.pnj8qZRxW
- Singer, N. (2018), "Microsoft Urges Congress to Regulate Use of Facial Recognition", 13 July 2018 (https://www.nytimes.com/2018/07/13/technology/microsoft-facial-recognition.html)
- Smik, D. (2018), Three Major Developments That Will Shape SEO In 2018, 14 February 2018 (https://www.forbes.com/sites/forbesagencycouncil/2018/02/14/three-major-developments-that-will-shape-seo-in-2018/#1abc104f1b8e)
- Smith, C. (2018), Interesting 4Chan statistics and facts, June 2018 (https://expandedramblings.com/index.php/4chan-statistics-facts/)
- Solon, Olivia (2018), "Facebook Struggling to End Hate Speech in Myanmar, Investigation Finds." *The Guardian*, August 16, sec. Technology. https://www.theguardian.com/technology/2018/aug/15/facebook-myanmar-rohingya-hate-speech-investigation
- Solon, Olivia and Levin, S. (2016), "How Google's search algorithm spreads false information with a rightwing bias", 16 December 2016 (https://www.theguardian.com/technology/2016/dec/16/google-autocomplete-rightwing-bias-algorithm-political-propaganda).
- Stamos, Alex (2017), "An Update On Information Operations On Facebook | Facebook Newsroom," September 6 (https://newsroom.fb.com/news/2017/09/information-operations-update/).
- Statista (2018), Number of active virtual reality users worldwide from 2014 to 2018 (in millions), https://www.statista.com/statistics/426469/active-virtual-reality-users-worldwide/
- Statista, Most popular social networks worldwide as of October 2018, ranked by number of active users (in millions), https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/
- Statista, Tumblr – Statistics & Facts, https://www.statista.com/topics/2463/tumblr/
- Statista, Virtual reality software and hardware market size worldwide from 2016 to 2020, by platform (in billion U.S. dollars), https://www.statista.com/statistics/528793/virtual-reality-market-size-worldwide-by-platform/
- Stecklow, Steve (2018). "Why Facebook Is Losing the War on Hate Speech in Myanmar." *Reuters*. Accessed October 27.(https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/).
- Stegemann, Patrick – Sören Musyal, Diana Kulozik, Lisa Wandt, Markus Pohl, Patrick Gensing. AfD-Funktionär an Troll-Attacken beteiligt. 01. 03. 2018. https://faktenfinder.tagesschau.de/inland/manipulation-wahlkampf-103.html
- Stop Fake.org (2017), "How Russian news networks are using Catalonia to destabilize Europe", By David Alandete, *El País*, October. (https://www.stopfake.org/en/how-russian-news-networks-are-using-catalonia-to-destabilize-europe/)
- Storyful Team (2017), "Debunking the Hurricane Irma Fakes", 11 September (https://storyful.com/seeing-is-not-believing-debunking-the-hurricane-irma-fakes/)
- Sumramanian, S. (2017), « Inside the Macedonian Fake-News Complex », 15 February 2017 (https://www.wired.com/2017/02/veles-macedonia-fake-news/)
- Sweden Democrats (2018) – anti-immigration, anti-EU party set to win more votes than ever, 6 September 2018 (https://theconversation.com/sweden-democrats-anti-immigration-anti-eu-party-set-to-win-more-votes-than-ever-102675)
- Teyssou, D. et al., The InVID Plug-in: Web Video Verification on the Browser, 2017, https://www.researchgate.net/publication/320570485_The_InVID_Plug-in_Web_Video_Verification_on_the_Browser

- The Economist (2017), "Russian Twitter Trolls Meddled in the Brexit Vote. Did They Swing It?" November 23. (https://www.economist.com/britain/2017/11/23/russian-twitter-trolls-meddled-in-the-brexit-vote-did-they-swing-it.)
- The Economist (2017), How the World Was Trolled (November 4-10), Vol. 425, No 9065
- The Local (2018), "Russian 'troll Factory' Tweets Tried to Influence Italian Voters," August 2. (https://www.thelocal.it/20180802/russian-troll-factory-tweets-attempted-influence-italian-elections.)
- The Local (2018), "Was 'the Year of Fake News in Italy', Regulator Warns." *The Local*, February 20. https://www.thelocal.it/20180220/fake-news-spread-italy-agcom-election.
- Thompson, I., How Irish anti-abortion activists are drawing on Brexit and Trump campaigns to influence referendum, 2 May 2018, https://www.opendemocracy.net/5050/isobel-thompson/irish-anti-abortion-campaigners-brexit-trump-data-companies
- Tiku, N. (2016), « Why Snapchat And Apple Don't Have A Fake News Problem », 1 December 2016 (https://www.buzzfeednews.com/article/nitashatiku/snapchat-fake-news).
- Timberg, Craig, and Shane Harris (2018), "Russian Operatives Blasted 18,000 Tweets Ahead of a Huge News Day during the 2016 Presidential Campaign. Did They Know What Was Coming?" *Washington Post*, July 20. (https://www.washingtonpost.com/technology/2018/07/20/russian-operatives-blasted-tweets-ahead-huge-news-day-during-presidential-campaign-did-they-know-what-was-coming/).
- Tough new German law puts tech firms and free speech in spotlight. 5. Jan. 2018. https://www.theguardian.com/world/2018/jan/05/tough-new-german-law-puts-tech-firms-and-free-speech-in-spotlight
- Transparent Referendum Initiative, Questions we're working on part 2: When is an ad political? 27 February 2018, https://medium.com/@TransparentRef/questions-were-working-on-part-2-what-counts-as-a-political-ad-d410209c5df6
- Tufekci, Zeynep (2018), "Opinion | YouTube, the Great Radicalizer.", *The New York Times*, June 8, 2018, sec. Opinion. (https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html).
- Two eras of the internet: pull and push. cdixon blog. 21.12.2014. http://cdixon.org/2014/12/21/two-eras-of-the-internet-pull-and-push/
- Tynan, Dan (2016), "How Facebook Powers Money Machines for Obscure Political 'news' Sites.", *The Guardian*, August 24, 2016, sec. Technology. (https://www.theguardian.com/technology/2016/aug/24/facebook-clickbait-political-news-sites-us-election-trump).
- Ugolik, K. (2017), "Can Virtual Reality Change Minds on Social Issues?" 11 January 2017 (https://narratively.com/can-virtual-reality-change-minds-social-issues/)
- Viner, Katharine (2016), "How Technology Disrupted the Truth." *The Guardian*, July 12, sec. Media. (https://www.theguardian.com/media/2016/jul/12/how-technology-disrupted-the-truth).
- Vogels, R. (2016), "Trump, Micro Targeting And The Mechanisms Of Data Capitalism", 17 December 2016 (https://www.huffingtonpost.com/entry/trump-micro-targeting-and-the-mechanisms-of-data_us_585433c0e4b0d5f48e164efc)
- Wakabayashi, Daisuke, and Nicholas Confessore (2017), "Russia's Favored Outlet Is an Online News Giant. YouTube Helped.", *The New York Times*, December 27, sec. Technology. https://www.nytimes.com/2017/10/23/technology/youtube-russia-rt.html.
- Willan, Philip: Twitter a trap for Italy's communications gurus. 22. June, 2015. PCWorld. https://www.pcworld.com/article/2938832/twitter-a-trap-for-italys-communications-gurus.html
- Winter, Jana (2018), "Hackers Targeting Election Networks across Country in Lead up to Midterms – The Boston Globe.", *The Boston Globe*, November 5. (https://www.bostonglobe.com/metro/2018/11/04/hackers-targeting-election-networks-across-country-lead-midterms/d0EzG4Cmh2jeMqllhXo4WP/story.html).
- Witness and First Draft (2018), Mal-uses of AI-generated Synthetic Media and Deepfakes: Pragmatic Solutions Discovery Convening, 11 June 2018, http://witness.mediafire.com/file/q5juw7dc3a2w8p7/Deepfakes_Final.pdf/file
- Wong, J. (2018), "It might work too well': the dark art of political advertising online", 19 March 2018 (https://www.theguardian.com/technology/2018/mar/19/facebook-political-ads-social-media-history-online-democracy)

- Zimdars, M. False, Misleading, Clickbait-y, and/or Satirical "News" Sources, https://docs.google.com/document/d/10eA5-mCZLSS4MQY5QGb5ewC3VAL6pLkT53V_81ZyitM/preview
- Zuckerberg, Mark (2017), "I Just Went Live a Minute Ago. Here's What I...," September 21, 2017. (https://www.facebook.com/zuck/posts/10104052907253171).

Other sources:

- Adams, D. et al. (2018), Ethics Emerging: the Story of Privacy and Security Perceptions in Virtual Reality. Conference paper for Fourteenth Symposium on Usable Privacy and Security ({SOUPS}**,** August 2018
- Buttarelli, Giovanni (2018), "Speech to LIBE on the Facebook/Cambridge Analytica Case", 25 June. (https://edps.europa.eu/sites/edp/files/publication/25-06-18_speech_giovanni_hearing_libe_en.pdf)
- European Centre for Press and Media Freedom (ECPMF) (2017), "Conference e-book Promoting dialogue between the European Court of Human Rights and the media freedom community. Freedom of expression and the role and case law of the European Court of Human Rights: developments and challenges", Strasbourg, 24 March. (https://ecpmf.eu/files/ecpmf-ecthr_conference_e-book.pdf)
- Tumblr, https://staff.tumblr.com/post/172170432865/were-taking-steps-to-protect-against-future
- Tumblr, https://www.tumblr.com/business
- Twitter, https://business.twitter.com/en/solutions/twitter-promote-mode.html

ANNEXES

Annex 1. Overview of selected digital platforms

No	Online platform	Main features	Examples of involvement in the informational manipulation actions
1.	**Facebook** **2 196 million users worldwide**[634]	• A social networking site that allows users to create profiles and interact with their 'friends' and promotional pages. Users can also create pages themselves and advertise their products and services. Facebook offers a variety of integrated services, including messenger services, marketplace, games and others. • Notable advertising features include *Custom Audiences* (allows advertisers to upload data from their database to identify and target users' Facebook profiles), *Lookalike Audiences* (enables to reach people who have similar profiles as those in a known group) and *Brand Lift* survey (measuring the success of the ads). • Facebook's unprecedented user base, its sophisticated targeting tools combined with a low level of transparency in terms of page ownership and ad purchases, led it to become one of the main (if not the main) vehicles for online manipulation.	• Facebook estimated that as many as 60 million bots might be infesting its platform. According to the researchers, they were responsible for a substantial portion of political content posted during the 2016 U.S. Presidential campaign, and some of the same bots were later used to attempt to influence the 2017 French elections.[635] • In the 2016 US Presidential Elections, Facebook was extensively used by domestic, foreign and unidentified actors to promote divisive content, particularly in the swing states. Some of the promoted ads advocated for specific candidates, others focussed more generally on issues such as guns, immigration and race relations. [636] The Trump campaign was known to use Facebook's 'dark posts' (sponsored Facebook posts which can only be seen by users with very specific profiles) to micro-target groups of voters with "40-50,000 variants of ads every day".[637] • Facebook was used as a vehicle by local, foreign and anonymous groups during the 2018 Ireland's abortion referendum,[638] the 2016

[634] As of July 2018. Information (except for Tumblr and 4Chan) from: Most popular social networks worldwide as of October 2018, ranked by number of active users (in millions), https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/; For the share of the EU traffic, see: European Commission, Behavioural Study on Advertising and Marketing Practices in Online Social Media, June 2018, p. 16 https://ec.europa.eu/info/sites/info/files/osm-final-report_en.pdf

[635] Lazer, D. et al, The science of fake news. Science 359(6380) : 1094-1096, http://science.sciencemag.org/content/359/6380/1094, p. 1094.

[636] Closing the Digital Loopholes that Pave the Way for Foreign Interference in U.S. Elections. Report based on study by Kim, Y. (University of Wisconsin), 16 April 2018, pp. 3-4, https://campaignlegal.org/sites/default/files/04-16-18%20CLC-IO%20Issue%20Brief%20Young%20Mie%20Report%20FINAL.pdf

[637] Illing, S., Cambridge Analytica, the shady data firm that might be a key Trump-Russia link, explained, 22 October 2017, https://www.vox.com/policy-and-politics/2017/10/16/15657512/cambridge-analytica-trump-kushner-flynn-russia

[638] Meserole C. and Polyakova A., Disinformation Wars, 25 may 2018, https://foreignpolicy.com/2018/05/25/disinformation-wars/

			UK referendum,[639] the 2017 German elections[640] and other important political events in the EU.
2.	**YouTube** **1 900 million users worldwide**	• A video-sharing platform that allows users to upload, view, rate, report, share, like and comment on video content, save content as a favourite, add it to playlists, subscribe to (i.e. follow) other users and channels. • Some recent new features include 360-degree videos, mobile live streams and virtual reality.[641] • YouTube advertising is based on TrueView ads, which work via a customised pricing model based on user engagement: traders only pay for viewers who watch the advertisement for at least 30 seconds. [642] • YouTube's proprietary recommendation algorithm has been criticised by both regulators and former employees for prioritising sensationalist and violent content, driven by the objective to extend the amount of time people spend online and to increase advertising revenue.[643]	• A report by Data Society found that "YouTube gives a platform to conspiracy theorists and fringe groups who can make persuasive, engaging videos on outrageous topics" and is used by "the far-right to spread extreme messaging to large numbers of people and to seed topics for journalists".[644] • According to Albright from Tow Center for Digital Journalism who investigated "fake news propaganda networks" during the 2016 US Presidential Elections, "many sites, domains, tweets, and Facebook pages were linking into YouTube — not just YouTube channels or single videos, but previews in tweets or Facebook pages".[645] His research led him to almost 80 000 fake videos uploaded to YouTube: "they were all keyword-stuffed. Very few of them had even a small number of views, so what these really were was about impact — these were a gaming system".[646] • According to Oxford University research, "cyber troops have been known to create and upload YouTube videos that 'contain

[639] Vote Leave's targeted Brexit ads released by Facebook, 26 July 2018, https://www.bbc.com/news/uk-politics-44966969

[640] Orlowski, R., Facebook says deleted many fake accounts in German campaign, 27 September 2018, https://www.reuters.com/article/us-germany-elections-facebook/facebook-says-deleted-many-fake-accounts-in-german-campaign-idUSKCN1C22Q5

[641] European Commission, Behavioural Study on Advertising and Marketing Practices in Online Social Media, June 2018, p. 22 https://ec.europa.eu/info/sites/info/files/osm-final-report_en.pdf

[642] Ibid.

[643] Lewis, P., "Fiction is outperforming reality": how YouTube's algorithm distorts truth, 2 February 2018, https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth

[644] Marwick, A. and Lewis R., Media, Manipulation and Disinformation Online. Data and Society, 2017, p. 26, https://centerformediajustice.org/wp-content/uploads/2017/07/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf; also see Newton, C., How white supremacists are thriving on YouTube, 19 September 2018, https://www.theverge.com/2018/9/19/17876892/youtube-extremism-report-rebecca-lewis-data-society

[645] Albright, J., The #Election2016 Micro-Propaganda Machine, 18 November 2016, https://medium.com/@d1gi/the-election2016-micro-propaganda-machine-383449cc1fba; Owen, L., News in a disintegrating reality: Tow's Jonathan Albright on what to do as things crash around us, 28 February 2018, http://www.niemanlab.org/2018/02/news-in-a-disintegrating-reality-tows-jonathan-albright-on-what-to-do-as-things-crash-around-us/

[646] Owen, L., News in a disintegrating reality: Tow's Jonathan Albright on what to do as things crash around us, 28 February 2018, http://www.niemanlab.org/2018/02/news-in-a-disintegrating-reality-tows-jonathan-albright-on-what-to-do-as-things-crash-around-us/

			persuasive messages' under online aliases" in the UK.[647] However, in its recent statement in early 2018, YouTube said that it found no "no evidence of Russian interference in the Brexit referendum".[648]
3.	**WhatsApp** **1 500 million users worldwide**	- An instant messaging platform, where users can exchange messages and calls individually or in groups of up to 256 members. Security of messaging services is guaranteed by end-to-end encryption. All WhatsApp accounts are tied to mobile phone numbers. - Information delivered via WhatsApp has more penetration and appears to come from a reliable – or known – source. WhatsApp messages come as personal, individual messages from specific contacts.[649] - WhatsApp does not display any advertisements, but in August 2016 WhatsApp announced that they would start sharing user data (i.e. phone numbers and aggregated analytical data) with Facebook.[650]	- WhatsApp is becoming an increasingly influential tool in political campaigning and disinformation actions in the Global South (e.g. Brazil, Malaysia, Kenya, Colombia). This is mostly attributed to a zero-rating practice found in the region, whereby telecoms offer free data to mobile phone users if they exclusively use Facebook or WhatsApp. According to Tactical Tech's investigation, "[w]hile zero-rating reduces the cost of accessing a service such as WhatsApp, it also discourages users from going on other platforms or accessing the web. As a result, it also limits their chances to fact-check information that comes from those platforms via other sources".[651] - For example, in Brazil, anti-vaccination groups spread disinformation on WhatsApp about yellow fever vaccinations, contributing to a measured uptick of the disease.[652] In Kenya, 'keyboard warriors' hired by political parties engaged in the dissemination of negative messages about opponents or general disinformation.[653]
4.	**Instagram** **1 000 million users worldwide**	- A photo and video-sharing social networking service. - Advertisements, in the form of sponsored posts, only started appearing on Instagram after November 2013.	- Instagram was reportedly being used by the Kremlin to engage in informational manipulation during the 2016 US Presidential elections. For example, a fake Instagram account linked to the

[647] Bradshaw, S., and Howard, P., Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation. University of Oxford: Working Paper, 2017(12), p. 12.

[648] YouTube finds no evidence of Russian interference in Brexit referendum, 8 February 2018, https://www.abc.net.au/news/2018-02-09/youtube-finds-no-evidence-of-russian-interference-in-brexit/9412036

[649] Renno, R., WhatsApp: The Widespread Use of WhatsApp in Political Campaigning in the Global South, 3 July 2018, https://ourdataourselves.tacticaltech.org/posts/whatsapp/

[650] European Commission, Behavioural Study on Advertising and Marketing Practices in Online Social Media, June 2018, p. 21

[651] Renno, R., WhatsApp: The Widespread Use of WhatsApp in Political Campaigning in the Global South, 3 July 2018, https://ourdataourselves.tacticaltech.org/posts/whatsapp/

[652] How WhatsApp Could Worsen Brazil's Yellow Fever Outbreak , 9 March 2018, https://www.wired.com/story/when-whatsapps-fake-news-problem-threatens-public-health/

[653] Renno, R., WhatsApp: The Widespread Use of WhatsApp in Political Campaigning in the Global South, 3 July 2018, https://ourdataourselves.tacticaltech.org/posts/whatsapp/

		• Advertisements on Instagram must be ordered via Facebook's in-house advertising platform,[654] and advertisers can make use of Facebook's targeting features. According to Albright from Tow Center for Digital Journalism, these advanced functionalities makes Instagram "more pervasive than Twitter for political meme-spreading as well as viral outrage video-based behavioural re-targeting".	Kremlin's Internet Research Agency posted voter suppression messages aimed at African Americans.[655] • According to Albright from Tow Centre for Digital Journalism, during the 2016 US Presidential Elections "[t]wo unofficial third-party 're-sharing' apps on Instagram have circulated and pushed IRA content far beyond the realm of Instagram and Facebook and embedded it all over the internet. This includes cross-posting of memes and posts from accounts removed from Instagram back into Facebook, Instagram, and also into Twitter. These apps also helped the memes get over to Pinterest".[656]
5.	**Tumblr** **345 million**[657] **users worldwide**	• A microblogging and social networking website, allowing users to post multimedia and other content to a short-form blog. • With respect to advertising, Sponsored Posts on Tumblr operate like regular Tumblr posts, except for they are more visible.[658]	• In March 2018, Tumblr made a statement explaining that it had uncovered 84 accounts linked to the Kremlin through its Internet Research Agency. According to the statement, these accounts were being used as part of a disinformation campaign leading up to the 2016 US Presidential elections.[659] • Trolls were reportedly using Tumblr to push anti-Clinton messages, including by actively promoting Democrat rival Bernie Sanders. Other themes included racial injustice and police violence.[660]
6.	**Twitter** **336 million** **users worldwide**	• A social networking service where users post and interact with messages known as 'tweets'. Tweets are restricted to 280 characters. The distinguishing characteristics of Twitter	• The indictment from the US Special Counsel detailed how Twitter was used in the 2016 US Presidential Elections. According to the available data, the Kremlin's Internet Research Agency controlled 3 814 human accounts and 50 258 bots on Twitter, with which 1.4

[654] European Commission, Behavioural Study on Advertising and Marketing Practices in Online Social Media, June 2018, p. 22 (internal references omitted)

[655] Barett, P. et al., Combating Russian Disinformation: The Case for Stepping Up the Fight Online. NYU Stern Center for Business and Human Rights : 2018, p. 6, https://disinfoportal.org/wp-content/uploads/ReportPDF/NYU-Stern-CBHR-Combating-Russian-Disinfomration-July-2018-min.pdf6

[656] Owen, L., Instagram is also a huge source of Russian propaganda on social media (Pinterest's not safe either), 9 November 2017, http://www.niemanlab.org/2017/11/instagram-is-also-a-huge-source-of-russian-propaganda-on-social-media-pinterests-not-safe-either/; Albright, J., Instagram, Meme Seeding, and the Truth about Facebook Manipulation, Pt. 1, 8 November 2017, https://medium.com/berkman-klein-center/instagram-meme-seeding-and-the-truth-about-facebook-manipulation-pt-1-dae4d0b61db5

[657] See: Statista, Tumblr - Statistics & Facts, https://www.statista.com/topics/2463/tumblr/

[658] Tumblr, https://www.tumblr.com/business

[659] Tumblr, https://staff.tumblr.com/post/172170432865/were-taking-steps-to-protect-against-future

[660] Lomas, N., Tumblr confirms 84 accounts linked to Kremlin trolls, 23 March 2018, https://techcrunch.com/2018/03/23/tumblr-confirms-84-accounts-linked-to-kremlin-trolls/

		include followers, @replies, #hashtags, direct private messaging, trending topics, verified accounts and polls.[661] • Twitter offers a possibility to promote existing tweets (*Promote Mode*) or to launch a separate Twitter ad campaign to a pre-selected target audience (*Twitter Ads*).[662] • Twitter is considered particularly vulnerable to disinformation campaigns. The reasons for this are twofold: first, Twitter accounts are not verified. Second, Twitter's application programming interface (API) still allows for false content to be easily created and spread.[663] • According to the Knight Foundation, most of the accounts spreading disinformation included in their study show evidence of automated posting, i.e. bot activity, and these accounts appear to be densely connected.[664]	million Americans interacted.[665] Some of the controlled Twitter accounts (e.g. ‘Tennessee GOP’ with 100 000 online followers) were falsely claimed to be operated by the US state political party and posted allegations of the voter fraud.[666] Others (e.g., @March_for_Trump) were used to organise political rallies in the US.[667] However, according to the Knight Foundation study, “[p]lenty of other accounts, though, do tweet in lockstep with the Kremlin’s message, including hundreds of accounts with more followers than top IRA trolls”. • Oxford University’s research did not find significant IRA activity on Twitter during the 2016 UK referendum. According to the research, 105 accounts tweeted almost 16 000 times in two separate weeks ahead of the 2016 vote, but it is not known how many people saw these tweets.[668] The prevalence of Twitter bots during the French and German 2017 elections was also not substantial, although, in Germany, they were particularly active in the context of the refugee debate.[669]

[661] European Commission, Behavioural Study on Advertising and Marketing Practices in Online Social Media, June 2018, p. 23

[662] Twitter, https://business.twitter.com/en/solutions/twitter-promote-mode.html

[663] Meserole C. and Polyakova A., Disinformation Wars, 25 may 2018, https://foreignpolicy.com/2018/05/25/disinformation-wars/

[664] Seven ways misinformation spread during the 2016 election. Knight Foundation, 4 October 2018, https://medium.com/trust-media-and-democracy/seven-ways-misinformation-spread-during-the-2016-election-a45e8c393e14

[665] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, p. 85, https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf

[666] Indictment in the case 18 U.S.C. §§ 2, 371, 1349, 1028A, https://www.justice.gov/file/1035477/download

[667] Ibid.; also see Howard, P., Social Media, News and Political Information during the US Election: Was Polarizing Content Concentrated in Swing States? 28 September 2017, https://www.recode.net/2017/9/28/16378186/twitter-fake-news-misinformation-russia-oxford-swing-states

[668] Russian tweets on Brexit were minimal, study shows, 18 December 2017, https://www.ft.com/content/fbf8ab4c-e41d-11e7-97e2-916d4fbac0da

[669] Howard, P., Junk News and Bots during the French Presidential Election: What Are French Voters Sharing Over Twitter? COMPROP data memo: Oxford University, 22 April 2017; Neudert L-M. et al., Junk News and Bots during the German Parliamentary Election: What are German Voters Sharing over Twitter? COMPROP data memo: Oxford University,

19 September 2017. A recent study by the Swedish Defense Research Institute found that 2,618 automated Twitter accounts were sending 6% of the content bearing the hashtags #svpol and #val2018 gathered since March 2018, see: Sweden Democrats – anti-immigration, anti-EU party set to win more votes than ever, 6 September 2018, https://theconversation.com/sweden-democrats-anti-immigration-anti-eu-party-set-to-win-more-votes-than-ever-102675

			• Specific disinformation cases in the EU that obtained traction due to Twitter bots include the 'Skripal affair' (the poisoning of a former Russian spy in the United Kingdom in April 2018)[670] and #MacronGate or #MacronLeaks, discussed below.[671]
7.	**Reddit** **330 million users worldwide**	• A platform for web content rating and discussion. It consists of a collection of opt-in communities called 'subreddits' (public or private), revolving around specific topics. Registered users can submit content on these subreddits and vote other users' posts up or down. • Some subreddits are a major source of conspiracy theories. In these cases, Reddit users act as amateur detectives, combing through documents and images to string together a theory. Once these theories gain traction on Reddit, they can be covered by fringe or even mainstream news outlets, further spreading the claims.[672] • Reddit is also said to be used in 'black hat' SEO tactics. According to Ghosh and Scott, "[o]ne example of how this is done—reputed to be a popular current tactic—is coordinated posting of a particular URL (or URLs) on Reddit. Hundreds or thousands of posts across relevant Reddit sub-threads are crawled and indexed by Google's search algorithm and may play a role in driving up search rank	• In the US, Reddit was instrumental in spreading the 'Pizzagate' scandal and disinformation about the Boston Marathon Bombing in April 2013, to name just two examples.[674] • During the 2017 elections in Germany, a story accusing Angela Merkel of deliberately allowing Islamic State terrorists to operate in Europe was posted on different conspiracy and politics subreddits. However, it failed to obtain traction or coverage.[675] • In 2018, Reddit reported that it had removed hundreds of accounts it suspects are of Russian origin or that linked directly to known sources of propaganda.[676] According to Reddit, Kremlin influence appeared to come mainly in the form of content posted by the troll accounts, rather than paid advertising.[677]

[670] Allegedly, different unverified conspiracy stories were posted by about 2,800 Twitter accounts which are likely to be bot-managed and have reached 7.5 million users. French study, p. 76

[671] Vilmer, J-B. et al., Information Manipulation: A Challenge for Our Democracies, pp. 107-108, https://www.diplomatie.gouv.fr/IMG/pdf/information_manipulation_rvb_cle838736.pdf

[672] Finkel, J. et al., Fake news and disinformation: The roles of the nation's digital newsstands, Facebook, Google, Twitter and Reddit. Stanford Law School : 2017, p. 157, https://www-cdn.law.stanford.edu/wp-content/uploads/2017/10/Fake-News-Misinformation-FINAL-PDF.pdf

[674] Finkel, J. et al., Fake news and disinformation: The roles of the nation's digital newsstands, Facebook, Google, Twitter and Reddit. Stanford Law School : 2017, p. 158, https://www-cdn.law.stanford.edu/wp-content/uploads/2017/10/Fake-News-Misinformation-FINAL-PDF.pdf

[675] Nimmo, B., Spread it on Reddit: How a fake story about Angela Merkel led to a far-right cluster on Reddit, 12 February 2017, https://www.stopfake.org/en/spread-it-on-reddit-how-a-fake-story-about-angela-merkel-led-to-a-far-right-cluster-on-reddit/

[676] Hautala, L., Reddit: Russian propaganda spread on our site before 2016 election, 5 March 2018, https://www.cnet.com/news/reddit-russian-propaganda-spread-on-our-site-before-2016-election/

[677] Aleem, Z., Reddit just shut down nearly 1,000 Russian troll accounts, 11 April 2018, https://www.vox.com/world/2018/4/11/17224294/reddit-russia-internet-research-agency

		before Reddit moderators intervene or Google spots an anomaly".[673]	
8.	**SnapChat** **255 million** **users worldwide**	• An image messaging social platform to exchange pictures with friends. The basic premise consists of privately sharing images that are only temporarily available and disappear after a short period or view.[678] • Differently from Facebook, posts from people are displayed chronologically, not by popularity or a personalised algorithm.[679] • Snapchat also allows ads to be purchased, including video ads, sponsored lenses and geo-filters.[680]	• Although Snapchat's images were criticised as racist and often sexist, the platform has found no evidence of disinformation dissemination or political ad buys from third countries for election interference purposes. As explained by Snapchat, this is mostly due to the involvement of human editors and content vetting.[681] For example, the app news section *Discover* is limited to professionally edited content. Before they can post in *Discover,* news publishers are vetted as a potential partner, an agreement that comes with strict terms.[682]
9.	**4Chan** **11 million**[683] **users worldwide**	• An image-based bulletin board where users can post comments and share images anonymously.[684] • Posts disappear very quickly, often after only a few hours. Each sub-board has a designated topic and specific norms which are enforced by other users.[685]	• 4chan is considered to breed conspiracy theorists and trolls. In 4Chan, similarly in Reddit, "users can dissect event footage in real time and instantaneously form theories that align with their worldviews. These groups often undergo polarisation effects: as sceptical users opt out of these communities, they become echo chambers of like-minded believers without exposure to any differing views".[688]

[673] Ghosh D. and Scott B., #Digitaldeceit. The technologies behind precision propaganda and the Internet. Harvard Kennedy School, January 2018, p. 20.

[678] European Commission, Behavioural Study on Advertising and Marketing Practices in Online Social Media, June 2018, p. 24

[679] Bossetta, M., The Digital Architectures of Social Media: Comparing Political Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election.

Journalism & Mass Communication Quarterly I-26, 2018, p. 17.

[680] European Commission, Behavioural Study on Advertising and Marketing Practices in Online Social Media, June 2018, p. 25

[681] Chafkin, M., How Snapchat Has Kept Itself Free of Fake News, 16 October 2017, https://www.bloomberg.com/news/features/2017-10-26/how-snapchat-has-kept-itself-free-of-fake-news

[682] Tiku, N., Why Snapchat And Apple Don't Have A Fake News Problem, 1 December 2016, https://www.buzzfeednews.com/article/nitashatiku/snapchat-fake-news

[683] Smith, C., Interesting 4Chan statistics and facts, June 2018, https://expandedramblings.com/index.php/4chan-statistics-facts/

[684] 4Chan, http://www.4chan.org/

[685] Marwick, A. and Lewis R., Media, Manipulation and Disinformation Online. Data and Society, 2017, p. 5 https://centerformediajustice.org/wp-content/uploads/2017/07/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

[688] Marwick, A. and Lewis R., Media, Manipulation and Disinformation Online. Data and Society, 2017, pp. 17-18 (internal references omitted), https://centerformediajustice.org/wp-content/uploads/2017/07/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

		- 4chan is often referred to as a birthplace of 'image- and humour- based' online subculture, which the modern 'alt-right' relies on.[686] - A 'spin-off' of 4chan is 8chan platform, which is described as "the more-lawless, more-libertarian, more 'free' follow-up to 4chan".[687]	- In the US, one of this kind of theory that was supported and widely shared in the 4Chan community was that Hillary Clinton was involved in a child sex ring and satanic rituals, otherwise known as the 'Pizzagate' scandal.[689] 4Chan is also known for actively spreading disinformation during crises and achieving number one search spots on Google and Facebook.[690] - In the EU, 4Chan was instrumental in circulating 'Macron leaks' during the 2017 French Presidential Elections. A link to the documents posted by an anonymous user on 4Chan was widely circulated and shared on Twitter and in media.[691]

Source: authors.[692]

[686] Ibid., p. 2.

[687] Dewey, C., This is what happens when you create an online community without any rules, 13 January 2015, https://www.washingtonpost.com/news/the-intersect/wp/2015/01/13/this-is-what-happens-when-you-create-an-online-community-without-any-rules/

[689] Marwick, A. and Lewis R., Media, Manipulation and Disinformation Online. Data and Society, 2017, p. 55, https://centerformediajustice.org/wp-content/uploads/2017/07/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

[690] Google Needs To Blacklist 4chan During National Crises, 3 October 2017, https://www.forbes.com/sites/fruzsinaeordogh/2017/10/03/google-needs-to-blacklist-4chan-during-national-crises/#58623f4b3dcd

[691] Mohan, M., Macron Leaks: the anatomy of a hack, 9 May 2017, https://www.bbc.com/news/blogs-trending-39845105

[692] The authors do not attempt to provide an in-depth analysis of platforms' digital architecture or an exhaustive account of their involvement in online manipulation campaigns, but rather to highlight certain features of the platforms that make them particularly (not) attractive for certain manipulation actions and give illustrative examples of such actions. During this exercise, the scarcity of quality empirical research into how exactly the listed digital platforms are used by different actors behind the informational manipulation in the EU, became apparent. Majority of the inquires to-date focussed on the Facebook, YouTube and Twitter, however, anecdotal evidence suggest that platforms like Reddit and 4Chan were behind some of the prominent disinformation actions in the EU. On the other hand, platforms, like Snapchat, seem to be largely immune to disinformation actions presumably because of the strict editorial control of the content. It is therefore important to include these players into the future research, also appreciating a diversity of national practices in using the platforms across the Union.

Annex 2. A description of the main disinformation actions since 2014

This section will provide a detailed overview of the cases assessed in chapter 1.3. It will also provide information on additional cases that, while important, could not have been included in chapter 1.3 due to space limitations.

Disinformation actions targeting foreign populations

2016 US Presidential Elections

The **Internet Research Agency** (IRA), a 'troll farm' set up in in 2013 in St. Petersburg, Russia, has been reportedly tasked with spreading pro-Kremlin propaganda online. Its employees, estimated to number in the thousands, are **paid to comment** on websites extensively, as well as **to run inauthentic social media accounts**.[693] IRA had reportedly been active in Russia,[694] as well as in Ukraine.[695] It started its US operations in 2014. In the run-up to the 2016 elections, the US operations reportedly employed **80 people** and had a **monthly budget of USD 1.25 million**.[696] The fake accounts, pretending to belong to US citizens, were primarily active on Facebook and Twitter.

According to Facebook, between June 2015 and August 2017, **120 IRA-linked accounts published 80 000 posts**, which may have been **seen by 126 million Americans**.[697] The posts aimed to create tension.[698] A fake Facebook group named Blacktivist posting "militant slogans and stomach-churning videos of police violence against African-Americans" had higher engagement rates than the real Black Lives Matter page.[699]

In September 2017, Facebook also revealed that about **470 fake accounts and pages spent USD 100 000 on** about **3 000 ads**, focusing on divisive issues "from LGBT matters to race issues to immigration to gun rights".[700] A quarter of the ads were geographically targeted.[701] The posts made and shared by the inauthentic accounts also discouraged minorities from voting and spread rumours about alleged voter fraud.[702]

[693] Benedictus, Leo. "Invasion of the Troll Armies: 'Social Media Where the War Goes On.'" The Guardian, November 6, 2016, sec. Media. https://www.theguardian.com/media/2016/nov/06/troll-armies-social-media-trump-russian.

[694] Seddon, Max. "Documents Show How Russia's Troll Army Hit America." BuzzFeed News, June 2, 2014. https://www.buzzfeednews.com/article/maxseddon/documents-show-how-russias-troll-army-hit-america.

[695] Bugorkova, Olga. "Inside the Kremlin's 'Troll Army.'" BBC News, March 19, 2015, sec. Europe. https://www.bbc.com/news/world-europe-31962644.

[696] "Internet Research Agency Indictment in the United States District Court for the District of Columbia." 2018. United States Department of Justice. p4. https://www.justice.gov/file/1035477/download.

[697] Ingram, David. "Facebook Says 126 Million Americans May Have Seen Russia-Linked Political Posts." Reuters, October 30, 2017. https://www.reuters.com/article/us-usa-trump-russia-socialmedia/facebook-says-126-million-americans-may-have-seen-russia-linked-political-posts-idUSKBN1CZ2OI.

[698] For example, in May 2016, an IRA-linked account masquerading as a local organisation called for a protest against an Islamic centre in Houston. Another IRA-linked account organised a pro-Islamic demonstration for the same time and the same place (Shane, Scott, and Mark Mazzetti. "The Plot to Subvert an Election: Unraveling the Russia Story So Far." The New York Times, September 20, 2018, sec. U.S. https://www.nytimes.com/interactive/2018/09/20/us/politics/russia-interference-election-trump-clinton.html).

[699] Mayer, Jane. 2018. "How Russia Helped Swing the Election for Trump." New Yorker, September 24, 2018. https://www.newyorker.com/magazine/2018/10/01/how-russia-helped-to-swing-the-election-for-trump, para. 44.

[700] Stamos, Alex. "An Update On Information Operations On Facebook | Facebook Newsroom," September 6, 2017. https://newsroom.fb.com/news/2017/09/information-operations-update/.

[701] In May 2018, the ads, totalling 3,500, were released, revealing that half of them referenced race, while a quarter focused on crime and policing (Penzenstadler, Nick, Brad Heath, and Jessica Guynn. "What We Found in Facebook Ads by Russians Accused of Election Meddling." USA Today, May 13, 2018. http://www.usatoday.com/story/news/2018/05/11/what-we-found-facebook-ads-russians-accused-election-meddling/602319002/.)

[702] "Internet Research Agency Indictment in the United States District Court for the District of Columbia." United States Department of Justice, February 16, 2018. https://www.justice.gov/file/1035477/download.

The IRA reportedly had **170 Instagram accounts** that reached **20 million people with 120 000 posts**.[703] Some research found evidence that Instagram's reach and influence went far beyond this figure.[704]

In January 2018, **Twitter** announced that it had removed over **3 800 IRA-linked accounts** and more than **50 000 Russian bots** in connection with its investigation of the US Presidential Elections.[705]

In March 2018, **Tumblr** identified **84 IRA-linked accounts** that spread disinformation before the US Presidential Elections.[706]

In April 2018, **Reddit** banned **944 accounts** it suspected belonged to IRA trolls.[707]

State-funded Russian media outlets television channel **RT** and website/news agency/radio station **Sputnik** also played a role in the disinformation campaign. They consistently presented Democratic candidate Hillary Clinton unfavourably.[708] Social media, often bots, were used to amplify their messages.

RT also has a carefully curated, successful **YouTube** channel,[709] which was the first news organisation on YouTube to reach, in 2013, **1 billion views**.[710]

Additionally, Russian **hackers** stole thousands of emails from the staff of the Democratic National Committee as well as Clinton campaign chair John Podesta. The stolen data were released via the websites **DCLeaks** and **Wikileaks**.

In July 2016, Russian hackers also managed to hack a state election office, stealing personal information of 500 000 voters, US federal investigators claim.[711] In the months leading to the 2018 midterm elections, hackers used similar methods to **try and steal electoral data**. The attacks have not been officially attributed to anyone.[712]

As discussed in chapter 1.3, disinformation produced for financial gain lies outside the scope of this study. Yet it is worth mentioning that apparently many of the websites running made-up news stories in relation to the elections were set up and operated by young men in the town of Veles, Former Yugoslav Republic of

[703] Shane, Scott, and Mark Mazzetti. "The Plot to Subvert an Election: Unraveling the Russia Story So Far." The New York Times, September 20, 2018, sec. U.S. https://www.nytimes.com/interactive/2018/09/20/us/politics/russia-interference-election-trump-clinton.html.

[704] Albright, Jonathan. "Instagram, Meme Seeding, and the Truth about Facebook Manipulation, Pt. 1." Medium (blog), November 8, 2017. https://medium.com/berkman-klein-center/instagram-meme-seeding-and-the-truth-about-facebook-manipulation-pt-1-dae4d0b61db5.

[705] Swaine, Jon. "Twitter Admits Far More Russian Bots Posted on Election than It Had Disclosed." The Guardian, January 20, 2018, sec. Technology. https://www.theguardian.com/technology/2018/jan/19/twitter-admits-far-more-russian-bots-posted-on-election-than-it-had-disclosed.

[706] Sommerlad, Joe. "Russian Hackers Targeted Tumblr during the US Election." The Independent, March 26, 2018. https://www.independent.co.uk/life-style/gadgets-and-tech/news/tumblr-russian-hacking-us-presidential-election-fake-news-internet-research-agency-propaganda-bots-a8274321.html.

[707] Huffman, Steve. "R/Announcements - Reddit's 2017 Transparency Report and Suspect Account Findings." reddit, April 17, 2018. https://www.reddit.com/r/announcements/comments/8bb85p/reddits_2017_transparency_report_and_suspect/.

[708] "Assessing Russian Activities and Intentions in Recent US Elections." Office of the Director of National Intelligence, January 6, 2017. https://www.dni.gov/files/documents/ICA_2017_01.pdf, p. 26.

[709] McAthy, Rachel. "How Russia Today Reached One Billion Views on YouTube | Media News." Journalism.co.uk, June 4, 2013. https://www.journalism.co.uk/news/how-russia-today-reached-one-billion-views-on-youtube/s2/a553152/.

[710] Wakabayashi, Daisuke, and Nicholas Confessore. "Russia's Favored Outlet Is an Online News Giant. YouTube Helped." The New York Times, December 27, 2017, sec. Technology. https://www.nytimes.com/2017/10/23/technology/youtube-russia-rt.html.

[711] "Internet Research Agency Indictment in the United States District Court for the District of Columbia." United States Department of Justice, February 16, 2018. https://www.justice.gov/file/1035477/download.

[712] Winter, Jana. "Hackers Targeting Election Networks across Country in Lead up to Midterms - The Boston Globe." The Boston Globe, November 5, 2018. https://www.bostonglobe.com/metro/2018/11/04/hackers-targeting-election-networks-across-country-lead-midterms/d0EzG4Cmh2jeMqllhXo4WP/story.html.

Macedonia.[713][714] "The **Macedonian teenagers**", as they came to be known, were apparently motivated by purely financial goals, although recent investigation suggests there might have been more to the story than that.[715]

The 2016 US Presidential Elections were not only subject to disinformation action from Russia but also from the US **alt-right movement**. These groups developed **"attention hacking"** techniques to manipulate public opinion. These often evidently untrue news stories and **conspiracy theories** were widely circulated and were amplified not only by the alt-right gaming the system but also, inadvertently, by mainstream media. More than **a quarter of US voting-age adults visited a disinformation website** in the weeks before the elections,[716] and the 20 top false news stories about the elections generated more engagements on Facebook than the top 20 mainstream news stories in the last three months before the elections.[717]

Brexit referendum (2016)

Evidence suggests Russian meddling in UK affairs well before the vote to leave the EU. After **the 2014 Scottish independence referendum**, pro-Russian accounts on Twitter **amplified claims** of election **fraud**, as well as demands for a revote.[718]

As for the Brexit referendum, links to Russia are less well-established than in the US case. There are allegations of **illicit campaign financing** from Russia. Leave.EU campaign donor Arron Banks is to be investigated by the UK's National Crime Agency because the UK's Electoral Commission suspects that he used money "from impermissible sources" for the Leave campaign and concealed their origin.[719] Banks was also reportedly offered lucrative **business deals** in Russian gold companies ahead of the Brexit referendum.[720]

Russia utilised its propaganda channels **RT** and **Sputnik** to campaign for Brexit. In the first six months of 2016, they published **261 pieces** on the referendum that showed an **anti-EU slant**.[721] These were then **amplified on social media**, with up to 134 million potential impressions, researchers calculate. This is a much wider potential reach than that of content shared from the Vote Leave and Leave.EU websites.[722]

[713] Tynan, Dan. "How Facebook Powers Money Machines for Obscure Political 'news' Sites." The Guardian, August 24, 2016, sec. Technology. https://www.theguardian.com/technology/2016/aug/24/facebook-clickbait-political-news-sites-us-election-trump.

[714] Silverman, Craig, and Lawrence Alexander. "How Teens In The Balkans Are Duping Trump Supporters With Fake News." BuzzFeed News, November 3, 2016. https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo.

[715] Cvetovska, Saska, Aubrey Belford, Craig Silverman, and J. Lester Feder. "The Secret Players Behind Macedonia's Fake News Sites." OCCRP, July 18, 2018. https://www.occrp.org/en/spooksandspin/the-secret-players-behind-macedonias-fake-news-sites.

[716] Guess, Andrew, Brendan Nyhan, and Jason Reifler. "Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 U.S. Presidential Campaign." 2018. http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf.

[717] Silverman, Craig. "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook." BuzzFeed News, November 16, 2016. https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook.

[718] Nimmo, Ben. "#ElectionWatch: Scottish Vote, Pro-Kremlin Trolls." DFRLab (blog), December 13, 2017. https://medium.com/dfrlab/electionwatch-scottish-vote-pro-kremlin-trolls-f3cca45045bb. It is impossible to tell definitively whether these accounts belonged to the IRA or just share Kremlin's ideology.

[719] "Arron Banks, Better for the Country and Others Referred to the National Crime Agency for Multiple Suspected Offences." The Electoral Commission, November 1, 2018. https://www.electoralcommission.org.uk/i-am-a/journalist/electoral-commission-media-centre/party-and-election-finance-to-keep/arron-banks,-better-for-the-country-and-others-referred-to-the-national-crime-agency-for-multiple-suspected-offences.

[720] Harding, Luke. "Revealed: Details of Exclusive Russian Deal Offered to Arron Banks in Brexit Run-Up." The Guardian, August 9, 2018, sec. UK news. https://www.theguardian.com/uk-news/2018/aug/09/revealed-detail-of-exclusive-russian-deal-offered-to-arron-banks-in-brexit-run-up.

[721] "'Disinformation and "Fake News": Interim Report.'" House of Commons Culture, Media and Sport Select Committee, July 2018. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/363/36308.htm#_idTextAnchor033.

[722] "89up Releases Report on Russian Influence in the EU Referendum." 89up, February 10, 2018. http://89up.org/russia-report.

2017 French Presidential Elections

In July 2017 it was reported that **Facebook** had taken action against **70 000 accounts** before the elections in May, because they were spreading disinformation,[723] although these were not linked to Russia. At the same time, Facebook revealed a **spying campaign** against Emmanuel Macron. Russian agents tried to gather personal information about Emmanuel Macron by posing as friends of friends of his associates.[724]

Some security experts linked **'Macron Leaks'** the hacking of the email accounts of Emmanuel Macron's campaign team, to the same Russian group that hacked the US Democratic National Committee's emails.[725]
Correlation (though no causation) was established between those who shared pro-Russia content and those who shared disinformation on social media.[726]

As in other cases, Russian state media outlets **RT** and **Sputnik** were found to be biased and spreading **unsubstantiated rumours** about Emmanuel Macron, which were amplified on Twitter and Facebook by a network of bots.[727]

Iranian efforts 2018

In August 2018, an "influence operation" launched by **Iran targeting** people in **the US, the UK, Latin-America and the Middle East** was uncovered.[728] **Facebook** removed **655** "inauthentic" pages, groups and **accounts** linked to Iran. **Twitter** did the same with nearly **300 accounts. Google** removed 58 accounts, including **39 YouTube channels** it suspects belongs to Iranian entities.[729] In October 2018, **Facebook** announced another round of action against "coordinated inauthentic behaviour from Iran," this time removing **82 Facebook and Instagram accounts**.[730]

Not as high-profile, but the operation was comparable in scale to Russian efforts, with the false personas and groups publishing a lot of content on divisive issues.[731] **Bots** were also used to amplify the messages.[732]

[723] Menn, Joseph. "Exclusive: Russia Used Facebook to Try to Spy on Macron Campaign - Sources." Reuters, July 27, 2017. https://www.reuters.com/article/us-cyber-france-facebook-spies-exclusive/exclusive-russia-used-facebook-to-try-to-spy-on-macron-campaign-sources-idUSKBN1AC0EI.

[724] Ibid.

[725] Hern, Alex. "Macron Hackers Linked to Russian-Affiliated Group behind US Attack." The Guardian, May 8, 2017, sec. World news. https://www.theguardian.com/world/2017/may/08/macron-hackers-linked-to-russian-affiliated-group-behind-us-attack.

[726] "A Russian Influence on the French Elections?" EU Disinfo Lab, 2018. https://spark.adobe.com/page/fJxCYVGj8d5Fk/.

[727] Nimmo, Ben. 2017. "The Kremlin's Audience in France." DFRLab (blog). April 14, 2017. https://medium.com/dfrlab/the-kremlins-audience-in-france-884a80515f8b.

[728] "Suspected Iranian Influence Operation Leverages Network of Inauthentic News Sites & Social Media Targeting Audiences in U.S., UK, Latin America, Middle East." Milpitas, CA: FireEye Intelligence, August 1, 2018. https://www.fireeye.com/blog/threat-research/2018/08/suspected-iranian-influence-operation.html.

[729] Abbruzzese, Jason, and Ingram, David. "Iran's Disinformation Campaign Extended to YouTube, Google Says." NBC News. Accessed November 11, 2018. https://www.nbcnews.com/tech/tech-news/iran-s-disinformation-campaign-extended-youtube-google-says-n903241.

[730] Gleicher, Nathaniel. "Taking Down Coordinated Inauthentic Behavior from Iran | Facebook Newsroom," October 26, 2018. https://newsroom.fb.com/news/2018/10/coordinated-inauthentic-behavior-takedown/.

[731] Gilbert, David. "Iran Is Running an Online Disinformation Campaign on the Scale of Russia's Troll Farm." Vice News, August 22, 2018. https://news.vice.com/en_ca/article/594ekk/iran-russia-facebook-twitter-disinformation.

[732] Nimmo, Ben, and Graham Brookie. "#TrollTracker: Facebook Uncovers Iranian Influence Operation." Medium (blog), October 26, 2018. https://medium.com/dfrlab/trolltracker-facebook-uncovers-iranian-influence-operation-d21c73cd71be.

Domestic campaigns

Hungarian government campaigns against migrants and against George Soros

The current wave of Hungarian governmental information campaigns **dates back to 2015**.

The first anti-migrant campaign was followed by several others, including one against George Soros and the EU.

The campaigns featured **newspaper ads**, **billboards**, **radio and television spots,** online **banners** and direct marketing **letters** to voting-aged Hungarians (in the framework of "national consultations").

These were presented as information in the public interest, commissioned by the government of Hungary and funded by **tax-payers**.

The campaigns peddled well-known **far-right tropes** such as: migrants taking away the jobs of Hungarians and increasing the terror threat, or George Soros having a "plan" to "flood" Europe with migrants[733].

The campaigns received support from **pro-government media**, both public media and outlets owned by businessmen with ties to the government. In the weeks in the run-up to the 2018 general elections, for example, the landing page of news site Origo.hu featured nothing but stories on "Muslim criminals".[734] Many of these stories were clearly fabricated. The Hungarian mainstream media have become so adept at producing fabricated content, that "genuine" disinformation websites such as *Napi migráns* (Daily Migrant) reportedly started featuring stories from Hungarian public media as well as pro-government mainstream media.[735]

Receiving its share of Russian propaganda, from 2013-16, Hungary had about **90** Hungarian-language **websites and blogs** promoting disinformation along **Kremlin lines**, mostly producing false news about the migration crisis.[736]

2018 Italian general elections

In 2016, BuzzFeed News linked Movimento Cinque Stelle (**M5S**) to a **network of websites and social media accounts** that spread disinformation.[737] Some of these sites that pose as independent news sites were allegedly found to be **under the direct control of M5S party leadership**; the site *TzeTze,* for instance, is reportedly owned by Casaleggio Associati, a firm founded by late M5S co-founder Gianroberto Casaleggio. M5S co-founder Beppe Grillo's personal blog, M5S party websites and some of these news sites were also exposed as sharing IP addresses, Google Analytics and AdSense IDs.[738] **Lega Nord** was also reported to share Google codes with websites that were not officially associated with the party and that spread pro-Putin propaganda and conspiracy theories.[739] Another BuzzFeed report exposed a large network of news sites and social media accounts that spread *hyperpartisan*, anti-immigration and Islamophobic content; some of the most popular non-legacy news sites were found to belong to this network.[740]

[733] See footnote 180.

[734] "Disinfo News: Media Consolidation and 'Fake News' Plague Hungary in Orban Era." POLYGRAPH.info, April 26, 2018. https://www.polygraph.info/a/fake-news-in-hungary/29194591.html.

[735] Kőműves, Anita. "Target or Ally? Hungary Faces the Elections Battle." Vsquare.org, March 4, 2018. https://vsquare.org/russia-target-or-ally-hungary-faces-the-elections-battle/.

[736] Panyi, Szabolcs, and András Dezső. "We Are Not Paid Agents of Russia, We Do It out of Conviction." Index, January 30, 2017. http://index.hu/belfold/2017/01/30/we_are_not_paid_agents_of_russia_we_do_it_out_of_conviction/.

[737] Silverman, Alberto Nardelli, Craig. "Italy's Most Popular Political Party Is Leading Europe In Fake News and Kremlin Propaganda." BuzzFeed, November 29, 2016. https://www.buzzfeed.com/albertonardelli/italys-most-popular-political-party-is-leading-europe-in-fak.

[738] Ibid.

[739] Horowitz, Jason. "Italy, Bracing for Electoral Season of Fake News, Demands Facebook's Help." The New York Times, November 24, 2017. https://www.nytimes.com/2017/11/24/world/europe/italy-election-fake-news.html.

[740] Silverman, Alberto Nardelli, Craig. "One Of The Biggest Alternative Media Networks In Italy Is Spreading Anti-Immigrant News And Misinformation On Facebook." BuzzFeed. Accessed November 5, 2018. https://www.buzzfeed.com/albertonardelli/one-of-the-biggest-alternative-media-networks-in-italy-is.

False information in Italy was allegedly rampant in the 2018 elections campaign.[741] Disinformation actions mostly concerned **Facebook**, which has 30 million users in the country. Facebook introduced **fact-checking** measures to counter the sharing of fake news.[742] When fabricated content was shared by politicians, Facebook opted not to remove these but to alert local politicians instead if false information was shared widely.[743] Most of the disinformation shared concerned migrants.[744]

In the last few elections, Lega Nord allegedly also made use of 'selfbots' on Twitter. Selfbots are human users who volunteered to become Lega Nord 'spokespeople' by authorising an app to automatically like and retweet content from Lega leader Matteo Salvini's account.[745]

Some signs of Russian election meddling in the campaign have also been reported. Some of the trolls identified in the US investigation as belonging to the IRA tweeted about the Italian elections.[746] Additionally, an analysis by big data firm Alto Data Analytics revealed that Russian bots largely amplified Sputnik's anti-immigration messages on Twitter in 2017.[747]

2018 Brazilian Presidential Elections

With unlimited access on many Brazilian mobile phone networks, **WhatsApp** has become the main way to discuss news for many Brazilians.[748] It also became the main source of disinformation in the recent Brazilian Presidential Elections. WhatsApp has 120 million users in Brazil, out of a population of 209.3 million.[749] A study found that **over half of 100 000 images** widely shared on WhatsApp in Brazil before the elections contained **false or misleading** information.[750] People were reportedly often **spammed** with messages. During the campaign, **phone numbers**, obtained from legal databases or illicitly, **were added to WhatsApp groups without their owners' consent**.[751]

On 18 October, Brazilian newspaper *Folha de Sao Paulo* reported that presidential frontrunner, populist Jair Bolsonaro encouraged his supporters among **the business elite to bankroll WhatsApp campaigns** attacking his rival, Fernando Haddad. Brazil's top electoral court opened a formal investigation into the issue, and **WhatsApp banned more than 100 000 accounts**.[752]

WhatsApp and similar group messaging apps seem uniquely suited to spread information as well as to wage disinformation campaigns. **WhatsApp messages are spread in closed groups, lending them an aura of**

[741] "Digital News Report 2018." Reuters Institute, 2018. http://media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf?x89475, p.87.

[742] Serhan, Yasmeen. "Italy Scrambles to Fight Misinformation Ahead of Its Elections." The Atlantic, February 24, 2018. https://www.theatlantic.com/international/archive/2018/02/europe-fake-news/551972/.

[743] Plucinska, Joanna, and Mark Scott. "How Italy Does Putin's Work." Politico, March 3, 2018. https://www.politico.eu/article/italy-election-fake-news-sunday-bufale-misinformation-vladimir-putin-russia/.

[744] Alaphilippe, A, C Ceccarelli, L Charlet, and M Mycielski. "Disinformation Detection System: 2018 Italian Elections." Brussels: EU Disinfo Lab, June 1, 2018. https://disinfo.eu/wp-content/uploads/2018/06/2018-Italian-elections-Case-report.pdf.

[745] Nimmo, Ben, and Anna Pellegatta. "#ElectionWatch: Italy's Self-Made Bots." DFRLab (blog), January 25, 2018. https://medium.com/dfrlab/electionwatch-italys-self-made-bots-200e2e268d0e.

[746] "Russian 'troll Factory' Tweets Tried to Influence Italian Voters," August 2, 2018. https://www.thelocal.it/20180802/russian-troll-factory-tweets-attempted-influence-italian-elections.

[747] Alandete, David, and Daniel Verdú. "How Russian Networks Worked to Boost the Far Right in Italy." El País. March 1, 2018, sec. In English. https://elpais.com/elpais/2018/03/01/inenglish/1519922107_909331.html.

[748] Nemer, David. "The Three Types of WhatsApp Users Getting Brazil's Jair Bolsonaro Elected." The Guardian, October 25, 2018, sec. World news. https://www.theguardian.com/world/2018/oct/25/brazil-president-jair-bolsonaro-whatsapp-fake-news.

[749] Magenta, Matheus, Juliana Gragnani, and Felipe Souza. "How WhatsApp Is Being Abused in Brazil's Elections," October 24, 2018, sec. Technology. https://www.bbc.com/news/technology-45956557.

[750] Isaac, Mike, and Kevin Roose. "Disinformation and Fake News Spreads over WhatsApp Ahead of Brazil's Presidential Election." The New York Times, October 20, 2018. https://www.independent.co.uk/news/world/americas/brazil-election-2018-whatsap-fake-news-presidential-disinformation-a8593741.html.

[751] Magenta, Matheus, Juliana Gragnani, and Felipe Souza. "How WhatsApp Is Being Abused in Brazil's Elections," October 24, 2018, sec. Technology. https://www.bbc.com/news/technology-45956557.

[752] Frier, Sarah, and Camillo Gulia. "WhatsApp Bans More Than 100,000 Accounts in Brazil Election," October 19, 2018. https://www.bloomberg.com/news/articles/2018-10-19/whatsapp-bans-more-than-100-000-accounts-in-brazil-election.

authenticity and trust. They are easy to forward to other groups, which allows for quick dissemination. From a regulatory/fact-checking perspective, they are hard to track down, due to the fact that the groups are closed, and the messages are encrypted.[753]

Disinformation campaign against the Rohingya minority in Myanmar

Facebook became accessible and affordable in Myanmar in 2013; it is so popular today, that for many of the country's 18 million Internet users, it became synonymous with **'the Internet'.**[754] Facebook's **Free Basics** programme, available in Myanmar in 2016-2017,[755] offering Facebook and some other apps access without data charges also contributed to its popularity.

Hate speech against the Muslim Rohingya minority appeared on Facebook **very early** on, with groups like the Buddhist nationalist Ma Ba Tha inciting hatred by spreading unsubstantiated rumours about the Muslim minority. Facebook was alerted to the problem in **November 2013** but took no countermeasures.[756] Facebook even failed to respond adequately after **a post** of a Muslim man raping a Buddhist woman **led to deadly riots** in the city of Mandalay in 2014.[757]

Facebook only started to address the problem in April 2018, but hate speech on Myanmar Facebook pages continues to flourish.[758]

In October 2018, the New York Times exposed an even darker, **systematic disinformation campaign** against the Rohingya, run **by Myanmar's military** on Facebook for half a decade.[759] As many as 700 Myanmar military personnel were tasked to create **fake pages** on Facebook dedicated to celebrities, news or beauty. These pages were then used **to spread hate speech and false rumours** against the Rohingya.

In September 2017, military accounts posing as fan pages sent warnings to both Buddhist and Muslim groups saying that "an attack from the other side was imminent".[760] That month, then-UN High Commissioner for Human Rights Zeid Raad Al Hussein called the ongoing violence "a textbook example of ethnic cleansing".[761]

On 5 November 2018, Facebook released a third-party assessment it had commissioned on its impact on human rights in Myanmar.[762] The report contains no earth-shattering insight, and blames most of the problems on low levels of digital literacy in the country. At the same time, the Human Rights Council also published two reports

[753] Popken, Ben. "How WhatsApp Became Linked to Mob Violence and Fake News — and Why It's Hard to Stop." NBC News, November 2, 2018. https://www.nbcnews.com/tech/tech-news/how-whatsapp-became-linked-mob-violence-fake-news-why-it-n929981.

[754] Solon, Olivia. "Facebook Struggling to End Hate Speech in Myanmar, Investigation Finds." The Guardian, August 16, 2018, sec. Technology. https://www.theguardian.com/technology/2018/aug/15/facebook-myanmar-rohingya-hate-speech-investigation.

[755]Moon, Mariella. "Facebook's Free Basics Quietly Pulled from Myanmar, Other Markets." Engadget, May 2, 2018. https://www.engadget.com/2018/05/02/facebook-free-basics-quietly-pulled-myanmar/.

[756]Ibid.

[757] Mezzofiore, Gianluca. "Wirathu's 'Buddhist Woman Raped' Facebook Post Stokes Anti-Muslim Violence in Mandalay." International Business Times UK, July 2, 2014. https://www.ibtimes.co.uk/wirathus-buddhist-woman-raped-facebook-post-stokes-anti-muslim-violence-mandalay-1455069.

[758] Stecklow, Steve. "Why Facebook Is Losing the War on Hate Speech in Myanmar." Reuters. Accessed October 27, 2018. https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/.

[759] Mozur, Paul. "A Genocide Incited on Facebook, With Posts From Myanmar's Military." The New York Times, October 18, 2018, sec. Technology. https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html.

[760] Ibid.

[761] Ingram, Mathew. "In Some Countries, Fake News on Facebook Is a Matter of Life and Death." Columbia Journalism Review, November 21, 2017. https://www.cjr.org/analysis/facebook-rohingya-myanmar-fake-news.php.

[762] "Human Rights Impact Assessment: Facebook in Myanmar." BSR, 2018. https://fbnewsroomus.files.wordpress.com/2018/11/bsr-facebook-myanmar-hria_final.pdf.

on Myanmar in August and September, which assign more responsibility to Facebook.[763] The second report concludes that Facebook should assess the human rights impact of its products in new markets before entering the market.[764]

[763] "Report of the Independent International Fact-Finding Mission on Myanmar." United Nations Human Rights Council, August 27, 2018. https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/274/54/PDF/G1827454.pdf?OpenElement.

[764] "Report of the Detailed Findings of the Independent International Fact-Finding Mission on Myanmar." United Nations Human Rights Council, September 18, 2018. https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_CRP.2.pdf?utm_campaign=The%20Interface&utm_medium=email&utm_source=Revue%20newsletter.

Annex 3. Legislative provisions of national electoral acts on campaign financing and countering disinformation and propaganda (in certain Member States)

Table 14: Election regulations of 15 selected Member States, which could be used to counter disinformation and propaganda

EU Member States	Year of last parliamentary elections	Provisions countering fake news, propaganda and disinformation and regulations concerning the use of campaign finances in electoral codes/election acts
Austria[765]	2017	N/A[766]
Bulgaria[767]	2017	**Articles 165, 166 and 167** define and restrict the amount and sources of money that can be spent on financing election campaigns. **Article 168** defines that a party, a coalition or a nomination committee shall not receive donations from certain sources, like anonymous donators, legal persons and religious institutions. **Articles 171 and 172** request that a Single Public Register of the parties, coalitions and nomination committees registered for participation in the respective type of elections shall be created at the Bulgarian National Audit Office before the elections, and within 30 days after the election day a report shall be made about the activities of providers of media services, the sociological and advertising agencies, as well as the public relations agencies. **Article 185** prohibits the display of election canvassing materials outside of the election campaign period. **Articles 476 and 477** claim that person(s) who breach the provisions of the above articles shall be liable to a fine.
Finland[768]	2015	N/A
Germany[769]	2017	N/A
Hungary[770]	2018	**Chapter VIII (Election campaigns)** gives a clear explanation and use of political advertising, its role, time and place. The exact definition can be found in Section 146. This chapter stipulates equal opportunity for all parties to demonstrate their view and the number of minutes to be dedicated to a political party in a certain kind of medium.
Ireland[771]	2016	N/A

[765] Federal Law on National Council Elections (1992) Source: http://www.ris.bka.gv.at/Dokumente/Erv/ERV_1992_471/ERV_1992_471.pdf

[766] N/A means that the legal documents (shown in the footnotes) did not contain any reference to the mentioned kind of provisions. The lack of the provisions can also be informative, by showing that in certain Member States legislation has not yet followed social patterns.

[767] Election Code of Bulgaria adopted on 5 March 2014, Source: https://www.venice.coe.int/webforms/documents/default.aspx?pdffile=CDL-REF(2014)025-e

[768] Election Act (714/1998) Source: https://www.finlex.fi/en/laki/kaannokset/1998/en19980714.pdf .

[769] Federal Electoral Law (Bundeswahlgesetz, BWG), 1956, Source: https://germanlawarchive.iuscomp.org/?p=228

[770] Act XXXVI of 2013 on Electoral Procedure, Source: http://www.valasztas.hu/documents/538536/548702/Act+XXXVI+of+2013+on+Electoral+Procedure.pdf/2e82a257-b592-4819-923f-eac4a18cfec6

[771] Electoral Act 1992, Source: http://www.irishstatutebook.ie/eli/1992/act/23/enacted/en/print#sec1

Italy[772]	2018	N/A
Latvia[773]	October 2018	**Section 19** states that there should be no obstacle to the exercise of voting rights, no public disturbance and no campaigning inside the polling station or within 50 meters from the entrance to the building in which the polling station is located.
Netherlands[774]	2017	N/A
Poland[775]	2015	**Article 107** prohibits campaigning on the day of the vote, and 24 hours beforehand, including convening meetings, organising marches and demonstrations, giving speeches and distributing materials. **Article 111 paragraph 1** states that election material disseminated in the press (posters, leaflets and slogans, as well as speech or other forms of election propaganda), which contain information that is untrue, is prohibited and the advertiser shall pay an amount of money up to of 100 000 zlotys to an organisation of public benefit. **Paragraph 4** requests that disinformation shall be corrected, replied to or apologised for at the latest within 48 hours, at the expense of person ordered to do so. **Chapter 13** contains the detailed rules and regulations for campaigning in radio and television programmes. It also stipulates that airtime allocated to one election committee cannot be transferred to another one. **Chapter 15** defines strict rules for financing of election campaigns on a general basis. **Articles 252-254** (Sejm) **and 284-285** (Senate) stipulate that an election committee has the right to free broadcasting of electoral programmes through public radio and television broadcasters, and even set the number of hours of airtime.
Portugal[776]	2015	**Articles 75 and 77** stipulates that political parties must make detailed accounts of all the revenues received and expenses incurred as well as stating that no party or coalition may spend more on the election campaign than fifteen times the monthly national minimum wage per candidate. **Article 143** reflects the above two by defining certain fines and sanctions on those parties and members thereof who breach the provisions of Article 75 and 77. There are many articles that provide detailed regulations on propaganda and campaign activities. **Article 62** strictly defines those media (radio and television) and broadcasting times where and when political advertising is allowed. **Article 139** contains sanctions for causing physical damages in election propaganda materials (a prison term of up to six months or a fine).

[772] Changes to the electoral system of the Chamber of Deputies and the Senate of the Republic. Delegation to the government of the establishment of uninominal and plurinominal electoral colleges. 3 November 2017, Source: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=2ahUKEwixglHqqIPeAhWvlYsKHVI9Ar4QFjABegQICBAC&url=https%3A%2F%2Fwww.legislationline.org%2Fdocuments%2Fid%2F21996&usg=AOvVaw3Pba9rxf6Deumxtqmg6BSs

[773] The Saeima Election Law 18 January 2018, Source: https://www.cvk.lv/pub/upload_file/Saeima_Election_Law_2018_ENG.pdf Access: 13.10.2018

[774] Act of 28 September 1989 containing new provisions governing the franchise and elections (Elections Act) Last amended by Act of 29 October 2009, Bulletin of Acts and Decrees 2009, no. 452, Source: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=2ahUKEwiG1JqY5oLeAhWpmIsKHT4aDqMQFjABegQICBAC&url=https%3A%2F%2Fwww.government.nl%2Fbinaries%2Fgovernment%2Fdocuments%2Fleaflets%2F2010%2F06%2F25%2Felections-act%2Fpdf-voor-engelse-site-elections-act-2010.pdf&usg=AOvVaw1hihzlGz03alCuVs3Q7tzl,

[775] Act of 5 January, 2011 Election Code (Journal of Laws 31 January, 2011), Unofficial translation for OSCE/ODIHR, Source: http://aceproject.org/ero-en/poland-2011-election-code/view,

[776] Assembly of the Republic Electoral Law, Law no. 14/79 of 16 May 1979, Source: http://aceproject.org/ero-en/regions/europe/PT/portugal-electoral-law-english-2011/view

		Article 141 stipulates that on the day of the election or on the day before it, it is forbidden to engage in electoral propaganda by any means (a prison term of up to six months or a fine). **Article 53** defines the campaign period as beginning on the fourteenth day, and ending at midnight on the second day, before the election day. **Article 58** is important as it states that "during election campaigns no limitation may be imposed on the expression of political, economic and social principles, without prejudice to any civil or criminal liability". And also, "during election campaign periods no sanctions whatsoever may be applied to enterprises that operate the media, or to their agents, for acts that form part of the campaign, without prejudice to any liability they incur, which may only be actioned after election day".
Romania[777]	2016	**Article 38** grants political parties free access to public radio and television services. Private radio and television stations shall practice the same tariff per programme and per time unit for all the electoral competitors participating in the elections. Also, electoral advertisements are only allowed to be transmitted within electoral programmes.
Spain[778]	2016	**Section 53** stipulates the period when electoral campaigning is prohibited, – no electoral propaganda may be disseminated nor can any electoral campaign event be held once the campaign is legally finished. Also, "from the call of the election to the legal start of the campaign, no commercial publicity or propaganda shall be allowed". **Section 55** defines two types of placing electoral propaganda. Campaign materials can only be put on display on surfaces provided free of charge by local councils or on authorised commercial surfaces. **Section 58** allows candidates to engage in advertising in the periodical press and on private broadcasting stations. **Section 65** defines the competent authority responsible for allotment of free propaganda space as the Central Electoral Commission under which a Radio and Television Committee shall work **Section 93** states that on polling day no electoral propaganda should be carried out. **Section 175** provides subsidy of expenses of campaign activities based on election results. Parties can get a certain amount of subsidy related to seats obtained in the Congress and Senate and votes obtained.
Sweden[779]	2018	**Chapter 8, Section 3** stipulates that "Propaganda or other activities aimed at influencing or impeding voters in making their choice may not occur at a voting station or in a space adjacent to it".
United Kingdom[780]	2017	N/A

[777] Regulations on the elections to the Chamber of Deputies and the Senate (2008), Source: https://publicofficialsfinancialdisclosure.worldbank.org/sites/fdl/files/assets/law-library-files/Romania_%20Law%20No%2035%20on%20Elections%20to%20Chamber%20of%20Deputies%20and%20Senate_2008_EN.pdf

[778] Representation of the people Institutional Act, 1st April 2015, Source: http://www.juntaelectoralcentral.es/cs/jec/documentos/LOREG_ENG

[779] The Elections Act (2005:837) Source: https://www.government.se/49150c/contentassets/4e2fdee5a8e342e88289496d34701aec/the-elections-act-2005837

[780] Political Parties and Elections Act 2009, Source: https://www.legislation.gov.uk/ukpga/2000/41/pdfs/ukpga_20000041_en.pdf

This study, commissioned by the European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs and requested by the European Parliament's Committee on Civil Liberties, Justice and Home Affairs, assesses the impact of disinformation and strategic political propaganda disseminated through online social media sites. It examines effects on the functioning of the rule of law, democracy and fundamental rights in the EU and its Member States.

The study formulates recommendations on how to tackle this threat to human rights, democracy and the rule of law. It specifically addresses the role of social media platform providers in this regard.

PE 608.864

Print ISBN 978-92-846-4638-8 | doi: 10.2861/209374 | QA-02-19-140-EN-C
PDF ISBN 978-92-846-4639-5 | doi: 10.2861/603773 | QA-02-19-140-EN-N