STUDY

Requested by the DROI subcommittee

European Parliament

The impact of disinformation on democratic processes and human rights in the world

@Adobe Stock

Authors:
Carme COLOMINA, Héctor SÁNCHEZ MARGALEF, Richard YOUNGS

European Parliament coordinator:
Policy Department for External Relations
Directorate General for External Policies of the Union
PE 653.635 - April 2021

EN

DIRECTORATE-GENERAL FOR EXTERNAL POLICIES
POLICY DEPARTMENT

European Parliament

STUDY

The impact of disinformation on democratic processes and human rights in the world

ABSTRACT

Around the world, disinformation is spreading and becoming a more complex phenomenon based on emerging techniques of deception. Disinformation undermines human rights and many elements of good quality democracy; but counter-disinformation measures can also have a prejudicial impact on human rights and democracy. COVID-19 compounds both these dynamics and has unleashed more intense waves of disinformation, allied to human rights and democracy setbacks. Effective responses to disinformation are needed at multiple levels, including formal laws and regulations, corporate measures and civil society action. While the EU has begun to tackle disinformation in its external actions, it has scope to place greater stress on the human rights dimension of this challenge. In doing so, the EU can draw upon best practice examples from around the world that tackle disinformation through a human rights lens. This study proposes steps the EU can take to build counter-disinformation more seamlessly into its global human rights and democracy policies.

EP/EXPO/DROI/FWC/2019-01/LOT6/R/02
April 2021 - PE 653.635

EN

Policy Department, Directorate-General for External Policies

AUTHORS

- Carme COLOMINA, Research Fellow, Barcelona Centre for International Affairs (CIDOB), Spain
- Héctor SÁNCHEZ MARGALEF, Researcher, Barcelona Centre for International Affairs (CIDOB), Spain
- Richard YOUNGS, Senior Fellow, Carnegie Endowment for International Peace
- Academic reviewer: Kate JONES, Associate Fellow, Chatham House; Faculty of Law, University of Oxford, United Kingdom

PROJECT COORDINATOR (CONTRACTOR)

- Trans European Policy Studies Association (TEPSA)

This study was originally requested by the European Parliament's Subcommittee on Human Rights.

CONTACTS IN THE EUROPEAN PARLIAMENT

Coordination: Marika LERCH, Policy Department for External Policies

Editorial assistant: Daniela ADORNA DIAZ

Feedback is welcome. Please write to marika.lerch@europarl.europa.eu

To obtain copies, please send a request to poldep-expo@europarl.europa.eu

VERSION

English-language manuscript completed on 22 April 2021.

This paper will be published on the European Parliament's online database, 'Think Tank'

ISBN: 978-92-846-8014-6 (pdf)
doi:10.2861/59161 (pdf)
Catalogue number: QA-02-21-559-EN-N (pdf)

ISBN: 978-92-846-8015-3 (paper)
doi:10.2861/677679 (paper)
Catalogue number: QA-02-21-559-EN-C (paper)

Table of contents

Acronyms and Abbreviations

CoE	Council of Europe
CSO	Civil Society Organisations
DEVCO	Directorate-General for International Cooperation and Development
EDAP	European Democracy Action Plan
EDMO	European Digital Media Observatory
EP	European Parliament
EU	European Union
HLEG	High Level Group of Experts on Fake News and Online Disinformation
ICCPR	International Covenant on Civil and Political Rights
ICESCR	International Covenant on Economic, Social and Cultural Rights
OECD	Organisation for Economic Cooperation and Development
OSCE	Organization for Security and Cooperation in Europe
PACE	Parliamentary Assembly of the Council of Europe
UDHR	Universal Declaration on Human Rights
UN	United Nations
UNHRC	United Nations Human Rights Council
WHO	World Health Organization

Executive Summary

The concept of disinformation refers to false, inaccurate or misleading information designed, presented and promoted intentionally to cause public harm or make a profit. Around the world, disinformation is spreading and becoming a more complex trend based on emerging techniques of deception. Hence, it needs to be understood today as being nested within countless techniques from manipulative information strategies. This reflects the acceleration of deep-fake technology, increasingly sophisticated micro-targeted disinformation campaigns and more varied influence operations. In its external relations, the EU needs comprehensive and flexible policy instruments as well as political commitment to deal more effectively with this spiralling phenomenon.

Disinformation also has far-reaching implications for human rights and democratic norms worldwide. It threatens freedom of thought, the right to privacy and the right to democratic participation, as well as endangering a range of economic, social and cultural rights. It also diminishes broader indicators of democratic quality, unsettling citizens' faith in democratic institutions not only by distorting free and fair elections, but also fomenting digital violence and repression. At the same time, as governments and corporations begin to confront this issue more seriously, it is apparent that many of their *counter*-disinformation initiatives also sit uneasily with human rights and democratic standards. Disinformation undermines human rights and many elements of good democratic practice; but counter-disinformation measures can also have a prejudicial impact on human rights and democracy.

The COVID-19 pandemic has intensified these trends and problems. It has unleashed new, more intense and increasingly varied disinformation campaigns around the world. Many non-democratic regimes have made use of the pandemic to crack down on political opposition by restricting freedom of expression and freedom of the media. COVID-19 compounds both disinformation's threat to international human rights, on the one hand, and the dangers of counter-disinformation serving anti-democratic agendas, on the other.

Effective responses to disinformation are needed at different levels, embracing formal legal measures and regulations, corporate commitments and civil society action. In many countries, legislative and executive bodies have moved to regulate the spread of disinformation. They have done so by elaborating codes of practice and guidelines, as well as by setting up verification networks to debunk disinformation. Some corporations have also launched initiatives to contain disinformation, although most have been ambivalent and slow in their efforts. Civil society is increasingly being mobilised around the world to fight against disinformation and often does so through a primary focus on human rights and building democratic capacity at the local level.

The EU needs to support counter-disinformation efforts in its external relations as a means of protecting human rights, making sure it does not support moves that actually worsen human rights. European institutions have begun to develop a series of instruments to fight disinformation, both internally and externally. Having promised a human rights approach in its internal actions, the EU has also formally recognised the need to build stronger human rights and democracy considerations into its external actions against disinformation and deceptive influence strategies. The EU's policy instruments have improved in this regard over recent years, with numerous concrete examples of EU initiatives that adopt a human-rights approach to counter disinformation in third countries.

There is a growing number of practical examples from around the world that offer best practice templates for how the counter-disinformation and human rights agendas can not only be aligned with each other but also be mutually reinforcing. Such examples offer valuable reference points and provide guidance on ways through which the EU should direct its counter-disinformation efforts, in tandem with external support, so as to pursue human rights and democracy. These emergent practices highlight the importance of collaboration between European institutions and civil society as the indispensable basis for building societal resilience to disinformation.

While the EU has begun to tackle disinformation in its external actions, it can and should place greater stress on the human rights dimension within this challenge. Despite the progress made in recent years, EU efforts to tackle disinformation globally need to dovetail seamlessly into the EU's overarching approach towards human rights internationally. The EU has tended to approach disinformation as a geopolitical challenge, to the extent that other powers use deception strategies against the Union itself, but much less as a human-rights problem within third countries. It still seeks to deepen security relations with many regimes guilty of using disinformation to abuse the rights and freedoms of their own citizens.

The EU can take a number of steps to rectify prevailing shortcomings, working at different levels of the disinformation challenge.

- Adopt further measures to exert rights-based external pressure both over corporations and third-country governments.
- Step up efforts to empower third country societies in the fight against disinformation.
- Foster new forms of global dialogue that fuse together disinformation and human rights concerns.

This report details how the EP can play a key role at each level of these recommendations.

- In its ongoing efforts with other parliaments around the world to strengthen global standards, it can push for a UN Convention on Universal Digital (Human) Rights, working together with legislatures from like-minded countries.
- Using these relationships with other parliaments along with its special committee on Foreign Interference in all Democratic Processes in the EU including Disinformation (INGE), the EP should promote best practices with third countries, emphasising the centrality of parliamentary accountability in countering disinformation.
- The Action Plan on Human Rights and Democracy 2020-24 stresses support to parliamentary institutions, thereby offering the EP a reference point and platform from which to exert stronger influence over implementation of the EU's external toolbox, thus ensuring that this gives adequate protection inter alia to human rights in the fight against disinformation.
- The EP can do more to push for increased funding to projects aimed at counteracting disinformation from a human rights perspective.

1 Introduction and methodology

This study considers the impact of online disinformation on democratic processes and human rights in countries outside the European Union. The scope of analysis is limited to legal content shared online; illegal content poses very different political and legal considerations. The study explores both the human rights breaches caused by disinformation as well as those caused by laws and actions aimed at tackling this phenomenon. Our research covered both EU institutional and civil society perspectives in detail.

The study analyses how the EU can better equip itself to tackle disinformation worldwide while protecting human rights. It explores public and private initiatives in third countries, identifying best practices of rights-based initiatives. Special attention is given to recent and current EU proposals, actions and instruments of significant relevance to tackling disinformation in third countries.

Our research methodology included a systematic desk-review of the existing literature on disinformation, human rights and democracy, relying on four types of sources: official documents, communication from stakeholders, scholarly literature and press articles. This study builds from the research published by international organisations, like UNESCO and the Council of Europe, and human rights resolutions from international bodies, including the UN Human Rights Council.

A series of semi-structured interviews were conducted between September 2020 and January 2021. The interviews were conducted under Chatham House rule, meaning that interviewees' comments are taken into account but not attributed in this report. The interviewees included staff members from several divisions of the European External Action Service (EEAS) – including its StratCom unit – and representatives from the Directorate General for International Cooperation (DEVCO, now the DG for International Partnerships, INTPA), as well as MEPs representing different political groups within the European Parliament. We prioritised MEPs who are members of the Subcommittee on Human Rights (DROI) or, in their absence, members of the Special Committee on Foreign Interference in All Democratic Processes within the EU, Including Disinformation (INGE) or Members from the Committee on Foreign Affairs (AFET), respecting gender balance at all times. The authors contacted more than one MEP per group, but the ratio of responses was low. Nonetheless, interviews were conducted with one MEP for each political group except for MEPs from the Identity and Democracy Group (ID), which did not respond to our request. Members from the Non-attached group (NI) were not contacted[1]. The authors would like to thank all the interviewees for their kind contributions.

Moreover, input was collected from civil society organisations through the European Endowment for Democracy (EED). The methodology sought to reflect the equal importance of formal institutional approaches and civil society responses to disinformation. The selection of best practices draws on a global civil society research project coordinated by one of the authors, which took place over the four years previous to this study. The project engaged with human rights defenders from all regions of the world who are active on these issues. This research project includes original fieldwork and empirical studies[2].

[1] Interview with an MEP from the European People's Party sitting in the Subcommittee on Human Rights on the 23rd of October, 2020; interview with an MEP from the Socialist and Democrats sitting in the Subcommittee on Human Rights and the Committee on Foreign Affairs on the 5th of October 2020; interview with an MEP from the Renew group sitting in the Committee on Foreign Affairs on the 17th of December 2020; interview with an MEP from the Group of the Greens/European Free Alliance sitting in the Subcommittee on Human Rights and the Committee on Foreign Affairs on the 14th of January 2021; interview with an MEP from the European Conservatives and Reformists Group sitting in the Special Committee on Foreign Interference in all Democratic Processes in the European Union, including Disinformation on the 12th of January 2021; and interview with an MEP from the Left Group sitting in the Committee on Foreign Affairs on the 29th of October 2020. MEPs represented a total of six political groups out of eight, with three women and three men interviewed.

[2] Carnegie's Civic Research Network is a global group of leading experts and activists dedicated to examining the changing patterns of civic activism around the world.

2 Understanding the concept of disinformation

'How can we make sense of democratic values in a world of digital disinformation run amok? What does freedom of speech mean in an age of trolls, bots, and information war? Do we really have to sacrifice freedom to save democracy—or is there another way?'

Peter Pomerantsev, Agora Institute, John Hopkins University and London School of Economics[3]

Key takeaways

- The concept of disinformation refers to false, inaccurate, or misleading information designed, presented and promoted intentionally to cause public harm or make a profit.
- Disinformation has spread rapidly with the rise of social media. Over 40 % of people surveyed in different regions of the world are concerned that it has caused increased polarisation and more foreign interference in politics.
- Disinformation can confuse and manipulate citizens; create distrust in international norms, institutions or democratically agreed strategies; disrupt elections; or feed disbelief in key challenges such as climate change.
- The more this phenomenon expands globally and the more multi-faceted it becomes, the more necessary it is to address not only the content dimension of disinformation but simultaneously the tactics of manipulative influence that accompany it. It is also important to understand the motivations for disinformation – political, financial or reputational (to build influence) – so that these can be addressed.

The internet has provided unprecedented amounts of information to huge numbers of people worldwide. However, at the same time, false, trivial and decontextualised information has also been disseminated. In providing more direct access to content, less mediated by professional journalism, digital platforms have replaced editorial decisions with engagement-optimising algorithms that prioritise clickbait content[4]. Social networks have transformed our personal exposure to information, persuasion and emotive imagery of all kinds.

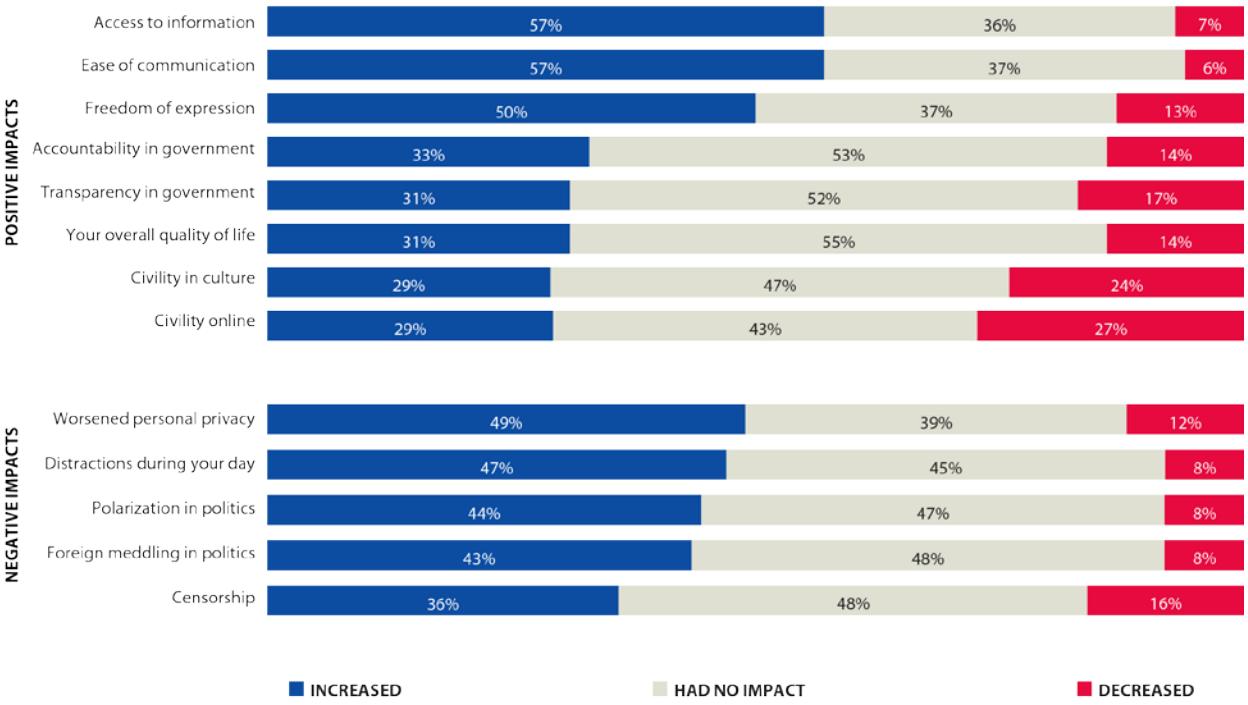
In recent years, the transmission of disinformation has increased dramatically across the world. While providing for a plurality of voices, a democratisation of access to information and a powerful tool for activism, the internet has also created new technological vulnerabilities. An MIT Media Lab research found that lies disseminate 'farther, faster, deeper, and more broadly than the truth' and falsehoods were '70 % more likely to be retweeted than the truth'[5]. False content has potentially damaging impacts on core

[3] Annenberg Public Policy Center, Freedom and Accountability: A Transatlantic Framework for Moderating Speech Online, Final Report of the Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression, June 2020.
[4] Karen Kornbluh and Ellen P. Goodman, Safeguarding Digital Democracy. Digital Innovation and Democracy Initiative Roadmap, *The German Marshall Fund of the United States DIDI Roadmap n 4*, March 2020.
[5] Soroush Vosoughi, Deb Roy and Sinan Aral, The spread of true and false news online. *Science*, vol 359(6380), 2018, pp 1146-1151.

human rights and even the functioning of democracy. Disinformation can serve to confuse or manipulate citizens; create distrust in international norms, institutions or democratically agreed strategies; disrupt elections; or fuel disbelief in key challenges such as climate change[6].

The graph below (Figure 1), on public perceptions of the impact of digital platforms, presents the results of interviews with over 1 000 people in 25 countries in late 2018 and early 2019. It shows that over 50 % of people considered that whilst social media had increased their ability to communicate and access information, they had mixed feelings about the impact on inter-personal relations. Over 40 % of people perceived that social media had contributed to both polarisation and foreign meddling in politics.

Figure 1. The impact of digital platforms: public perceptions[7]



Source: Authors' own elaboration based on CIGI-Ipsos, 2019 CIGI-Ipsos Global Survey on Internet Security and Trust., 2019; CIGI-Ipsos, Internet security and Trust. Part 3, 2019.

2.1 Definition of disinformation

For the European Union, the concept of disinformation refers to 'verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the public and may cause public harm'[8]. A similar definition was also adopted by the Report of the independent High-

[6] Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.

[7] The percentage refers to people interviewed in 25 different countries (Australia, Brazil, Canada, China, Egypt, France, Germany, Great Britain, Hong Kong (China), India, Indonesia, Italy, Japan, Kenya, Mexico, Nigeria, Pakistan, Poland, Russia, South Africa, Republic of Korea, Sweden, Tunisia, Turkey and the United States) accounting for 1.000+ individuals in each country during December 2018 and February 2019: See the original source (CIGI IPSOS Global Survey) for more information about the methodology employed.

[8] European Commission, Joint Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Action Plan against Disinformation, JOIN(2018) 36 final, December 2018.

Level Group on Fake News and Online Disinformation published in March 2018[9]. Under this definition, the risk of harm includes threats to democratic political processes and values. The production and promotion of disinformation can be motivated by economic factors, reputational goals or political and ideological agendas. It can be exacerbated by the ways in which different audiences and communities receive, engage and amplify disinformation[10]. This definition is in line with the one adopted in a note produced by the European Parliamentary Research Service (EPRS) in 2015[11].

The authors of *Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States*[12] use the terms 'disinformation' and 'propaganda' to describe phenomena characterised by four features, namely that: 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'.

Immediately after the 2016 US election, concepts such as 'alternative facts', 'post-truth' and 'fake news' entered into public discourse. Even though the term 'fake news' emerged around the end of the 19th century, it has become too vague and ambiguous to capture the essence of disinformation. Since the term 'fake news' is commonly used as a weapon to discredit the media, experts have called for this term to be abandoned altogether in favour of more precise terminology[13]. The High-Level Group on Fake News and Online Disinformation also took the view that 'fake news' is an inadequate term, not least because politicians use it self-servingly to dismiss prejudicial coverage[14].

Wardle and Derakhshan categorise three types of information disorders to differentiate between messages that are true and those that are false, as well as determining which are created, produced, or distributed by 'agents' who intend to do harm and those that are not[15]. These three types of information disorders – dis-information, mis-information and mal-information – are illustrated in Table 1. The intention to harm or profit is the key distinction between disinformation and other false or erroneous content.

[9] Here, disinformation refers to false, inaccurate or misleading information designed, presented and promoted intentionally to cause public harm or make a profit. See Independent High level Group on fake news and online disinformation, A multi-dimensional approach to disinformation, Report for the European Commission, March 2018.

[10] Independent High level Group on fake news and online disinformation, March 2018.

[11] Authors' note: The European Parliamentary Research Service, in its 2015 *At a glance* paper, used the Oxford English Dictionary's definition of disinformation: 'dissemination of deliberately false information, especially when supplied by a government or its agent to a foreign power or to the media, with the intention of influencing the policies or opinions of those who receive it; false information so supplied'. It also acknowledged that in some official communications, the European Parliament has used the term *propaganda* when referring to Russia's disinformation campaigns or even *misinformation*, although there is consensus that misinformation happens unintentionally.

[12] Judit Bayer, Natalija Bitiukova, Petra Bárd, Judit Szakács, Alberto Alemanno and Erik Uszkiewicz, 'Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States', Directorate General for Internal Policies of the Union (IPOL), European Parliament, 2019.

[13] Margaret Sullivan, It's Time To Retire the Tainted Term Fake News, *Washington Post*, 6 January 2017. [Accessed on 26 March 2021].

[14] Independent High level Group on fake news and online disinformation, A multi-dimensional approach to disinformation, Report for the European Commission, March 2018.

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

Table 1. Types of Information Disorders

	Definition	Example
Misinformation	When **false information** is shared, but **no harm** is meant	A terror attack on the Champs Elysees on 20 April 2017 generated a great amount of misinformation in social networks, spreading rumours and unconfirmed information[16]. People sharing that kind of information didn't mean to cause harm.
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[17].
Malinformation	When **genuine information** is shared to **cause harm**	The intentional leakage of a politician's private emails, as happened during the presidential elections in France 2017[18].

Source and examples: Wardle and Derakhshan, Information Disorder: Toward an interdisciplinary framework for research and policy making, Council of Europe report DGI(2017)09, 2017.

The challenge posed by disinformation comes not only from its content, but also how it is distributed and promoted on social media. The intention to harm or profit that characterises disinformation itself entails that disinformation is commonly accompanied by strategies and techniques to maximise its influence. Hence, the European Democracy Action Plan[19] – the European Commission's agenda to strengthen the resilience of EU democracies – broadens the discussion from tackling disinformation to also tackling 'information influence operations' and 'foreign interference in the information space'. Both these concepts encompass coordinated efforts to influence a targeted audience by deceptive means, the latter involving a foreign state actor or its agents. EU external strategies need to respond not only to the challenge of disinformation, but to these deceptive influence strategies more broadly.

Social media platforms are already tackling these challenges of influence to some extent. For instance, Facebook tackles 'coordinated inauthentic behaviour', understood as 'coordinated efforts to manipulate public debate for a strategic goal where fake accounts are central to the operation'[20]. Twitter's 'platform manipulation' refers to the 'unauthorised use of Twitter to mislead others and/or disrupt their experience by engaging in bulk, aggressive, or deceptive activity'[21]. Similarly, Google reports on 'coordinated influence operation campaigns'[22].

[16] One example of this, mentioned by Wardle and Derakhshan (2017), was the rumour that Muslim population in the UK had celebrated the attack. This was debunked by CrossCheck. For more information on the role of social media that night, also read Soren Seelow 'Attentat des Champs-Elysées : le rôle trouble des réseaux sociaux', in *Le Monde*, 4 May 2017.

[17] EU vs. Disinfo, 'EMMANUEL MACRON'S CAMPAIGN HAS BEEN FUNDED BY SAUDI ARABIA, SINCE…', published on 2 February 2017. [Accessed on 11 February 2021].

[18] Meghan Mohan, Macron Leaks: the anatomy of a hack, *BBC*, published on 9 May 2017. [Accessed on 11 February 2021]

[19] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the European Democracy Action Plan, COM (2020) 790 final, December 2020.

[20] Facebook, December 2020 Coordinated Inauthentic Behavior Report, January 2021. [Accessed on 11 February 2021].

[21] Twitter Transparency Center, Platform Manipulation, January – June 2020, January 2021.

[22] Shane Huntley, Updates about government-backed hacking and disinformation, Google Threat Analysis Group, May 2020. [Accessed on 11 February 2021]

2.2 Instigators and Agents of disinformation

Anyone with a social media account can create and spread disinformation: governments, companies, other interest groups, or individuals. When analysing the different actors responsible for disinformation, the UNESCO Working Group on Freedom of Expression and Addressing Disinformation makes a distinction between those fabricating disinformation and those distributing content: the *instigators* (direct or indirect) are those creating the content and the *agents* ('influencers', individuals, officials, groups, companies, institutions) are those in charge of spreading the falsehoods[23].

The most systemic threats to political processes and human rights arise from organised attempts to run coordinated campaigns across multiple social media platforms. As indicated by Facebook's exposure of coordinated inauthentic behaviour, large disinformation campaigns are often linked with governments, political parties and the military, and/or with consultancy firms working for those bodies[24]. When the instigator or agent of disinformation is – or has the backing of – a foreign state it may be breaching the public international law principle of non-intervention[25]. Experts consider that the principle of non-intervention applies as much to a state's cyber operations as it does to other state activities. Foreign interference should be understood as the 'coercive, deceptive and/or non-transparent effort to disrupt the free formation and expression of individuals' political will by a foreign state actor or its agents'[26]. Since exposure of the role of Russia's Internet Research Agency (IRA) – considered a 'troll farm' spreading pro-Kremlin propaganda online under fake identities – in US politics in 2016, concerns have grown over foreign interference operations involving disinformation. Civil society monitoring indicates that an increasingly assertive China has joined Russia in interfering in democratic processes abroad[27]. This includes interference in elections through concerted disinformation campaigns, fostering democratic regression and promoting 'authoritarian resurgence'[28]. An EEAS special report published in April 2020 noted that 'state-backed sources from various governments, including Russia and – to a lesser extent – China, have continued to widely target conspiracy narratives and disinformation both at public audiences in the EU and the wider neighbourhood' [29].

2.3 Tools and tactics

Disinformation campaigns are becoming increasingly sophisticated and micro-targeted, through marketing strategies that use people's data to segment them into small groups, thus providing apparently tailored content. The fact that content sharing has also moved from open to encrypted platforms (WhatsApp, Facebook Messenger and WeChat are in the top five social media platforms globally) makes it more difficult to track disinformation.

[23] Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.

[24] Nathaniel Gleicher, Removing Coordinated Inauthentic Behavior, *Facebook*, published on 8 July 2021. [Accessed on 11 February 2021]

[25] Carolyn Dubay, A Refresher on the Principle of Non-Intervention, *International Judicial Monitor*, Spring Issue, 2014.

[26] James Pamment, The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework, *Carnegie Endowment for International Peace Future Threats Future Solutions series*, n 2, September 2020 https://carnegieendowment.org/2020/09/24/eu-s-role-in-fighting-disinformation-crafting-disinformation-framework-pub-82720

[27] European Partnership for Democracy (EPD), Louder than words? Connecting the dots of European democracy support. 2019.

[28] Larry Diamond, The Democratic Rollback. The Resurgence of the Predatory State, *Foreign Affairs*, 87(2), 2008, pp 36–48.

[29] EU vs. Disinfo, EEAS SPECIAL REPORT UPDATE: SHORT ASSESSMENT OF NARRATIVES AND DISINFORMATION AROUND THE COVID-19/CORONAVIRUS PANDEMIC (UPDATED 2 – 22 APRIL), published on 24 April 2020. [Accessed on 11 February 2021]

Among the emerging tools available to spread disinformation are:

- Manufactured amplification (artificially boosting the reach of information by manipulating search engine results, promoting hashtags or links on social media) [30]
- Bots (social media accounts operated by computer programmes, designed to generate posts or engage with social platforms' content) [31]
- Astroturf campaigns (masking the real sponsor of a message, giving the false impression that it comes from genuine grass-roots activism)
- Impersonation of authoritative media, people or governments (through false websites and/or social media accounts)
- Micro-targeting (using consumer data, especially on social media, to send different information to different groups[32]. Even if micro-targeting is not necessarily illegal and may be equally used by those spreading legitimate information, the scandal of Cambridge Analytica demonstrates that it poses a serious risk regarding the spread of disinformation) [33]
- 'Deep-fakes' (digitally altered or fabricated videos or audio)[34]

There are different state-sponsored disinformation activities that can be considered harmful practices. In 2019, political parties or leaders in around 45 democratic countries used computational propaganda tools by amassing fake followers to gain voter support; in 26 authoritarian states government entities used computational propaganda as a tool for information control to suppress public opinion and press freedom and discredit opposition voices and political dissent[35].

Manipulation can also be achieved through online selective censorship (removing certain content from a platform by governments' demands or by platforms' curation responses), hacking and sharing or manipulating search engine results[36].

In this acceleration of manipulation, deep-fake technology can be harmful because people cannot tell whether content is genuine or false. Real voices can be manipulated to say things that were never said, making citizens unable to decipher between fact and fiction. That is why deep-fake technology may pose particular harm to democratic discourse and to national and individual security. When we can no longer believe what we see, truth and trust are at risk. Even before entering the political scene, deep-fake technology had already been used to fabricate pornographic content for criminal purposes, thereby

[30] Sam Earle, Trolls, Bots and Fake News: The Mysterious World of Social Media Manipulation, *Newsweek*, published on 14 October 2017. [Accessed on 11 February 2021]
[31] Earle, 2017.
[32] Ghosh Dipayan, What Is Microtargeting and what is it doing in our politics, *Internet Citizen*, Harvard University, 2018.
[33] Carole Cadwalladr, as told to Lee Glendinning, Exposing Cambridge Analytica: 'It's been exhausting, exhilarating, and slightly terrifying, *The Guardian*, published on 29 September 2018. [Accessed on 11 February 2021]; and Dobber, Tom; Ronan Ó Fathaigh and Frederik J. Zuiderveen Borgesius,, The regulation of online political micro-targeting in Europe, *Journal on internet regulation*, Vol. 8(4), 2019.
[34] Beata Martin-Rozumiłowicz and Rasto Kužel, Social Media, Disinformation and Electoral Integrity, *International Foundation for Electoral Systems*, Working Paper, August 2019.
[35] Samantha Bradshaw and Philip N. Howard, The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation, University of Oxford Working Paper 2019(3), 2019.
[36] Judit Bayer, Natalija Bitiukova, Petra Bárd, Judit Szakács, Alberto Alemanno and Erik Uszkiewicz, 'Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States', Directorate General for Internal Policies of the Union (IPOL), European Parliament, 2019.

posing a tangible human rights threat to women in particular[37]. The Brookings Institute notes that harmful deep-fakes may impact democratic discourse in three ways[38]:

- Disinformative video and audio: citizens may believe and remember online disinformation, which can be spread virally through social media.
- Exhaustion of critical thinking: if citizens are unable to know with certainty what news content is true or false, this will exhaust their critical thinking skills leading to an inability to make informed political decisions.
- The Liar's Dividend: politicians will be able to deny responsibility by suggesting that a true audio or video content is false, even if it is true (in the way that 'fake news' has become a way of deflecting media reporting).

Technology is always one step ahead and brings about a near-future where 'we will see the rise of cognitive-emotional conflicts: long-term, tech-driven propaganda aimed at generating political and social disruptions, influencing perceptions, and spreading deception' (Pauwels, 2019: 16). Some studies are already measuring the effects of deep-fakes on people's political attitudes and proving how microtargeting amplifies those effects[39].

2.4 Motivations for disinformation

Research has identified a variety of motivations behind disinformation, including financial or political interests, state actors' agendas, trolling and disruption along with even the desire for fame[40]. In some cases, distorted information does not always seek to convince, but rather to emphasise divisions and erode the principles of shared trust that should unite societies[41]. Lies breed confusion and contribute to the 'decay of truth'[42] and to a clash of narratives. In other cases, disinformation can be a very powerful strategy built on low-cost, low-risk but high-reward tactics in the hands of hostile actors, feeling less constrained by ethical or legal principles and offering very effective influence techniques[43]. The proliferation of social media has democratised the dissemination of such narratives and the high consumption of content undercutting truthfulness creates the perfect environment for some states to innovate in the old playbook of propaganda. Financial reward can also provide substantial motivation as almost USD 0.25 billion is spent in advertising on disinformation sites each year[44].

Hwang[45] identifies political, financial and reputational incentives. Firstly, political motivation to spread disinformation can go from advancing certain political agendas (for instance, linking immigration with criminality) to imposing a narrative that presents a better geopolitical image of certain other nations (illiberal democracies as an opposite to failing western democracies). Secondly, economic motivation is

[37] Madeline Brady, Deepfakes: a new disinformation threat? *Democracy Reporting International*, August 2020.
[38] Alex Engler, Fighting deepfakes when detection fails, *The Brookings Institute*, Published on 24 November 2019. [Accessed on 11 February 2021].
[39] Tom Dobbe, Nadia Metoui, Damian Trilling, Natali Helberger and Claes de Vreese, Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes, *The International Journal of Press/Politics*, Vol 26(1), 2021, pp 69-91.
[40] Rebecca Lewis, and Alice Marwick, Taking the Red Pill: Ideological motivations for Spreading Online Disinformation, in *Understanding and Addressing the Disinformation Ecosystem*, Annenberg School for Communication, December 2017.
[41] Carme Colomina, Techno-multilateralism: The UN in the age of post-truth diplomacy, in Bargués, P., *UN@75: Rethinking multilateralism*, CIDOB Report, vol.6, 2020.
[42] Jennifer Kavanagh and Michael D. Rich, Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life. RAND Corporation Research Reports, 2018.
[43] James Pamment. The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework, *Carnegie Endowment for International Peace Future Threats Future Solutions series*, n 2, September 2020.
[44] Global Disinformation Index, The Quarter Billion Dollar Question: How is Disinformation Gaming Ad Tech?, September 2019.
[45] Tim Hwang, 'Digital disinformation: A primer', *Atlantic Council Eurasia Group*, September 2017.

linked with social platforms' economic models that benefit from the 'click-bait model', hence attracting users through content that is tempting, albeit false. Finally, reputational motivation has to do with the penetration of social networks used in our daily lives. There is an increasing dependence on friends, family or group endorsement[46].

3 The impacts of disinformation and counter-disinformation measures on human rights and democracy

'... technology would not advance democracy and human rights for (and instead of) you'

Zygmunt Bauman in A Chronicle of Crisis: 2011-2016

Key takeaways:

- Online disinformation has an impact on human rights. It affects the right to freedom of thought and the right to hold opinions without interference; the right to privacy; the right to freedom of expression; the right to participate in public affairs and vote in elections.
- More broadly, disinformation diminishes the quality of democracy. It saps trust in democratic institutions, distorts electoral processes and fosters incivility and polarisation online.
- While robust counter-disinformation is needed to protect democracy, it can itself undercut human rights and democratic quality.

The Universal Declaration of Human Rights (UDHR) offers protection to all people in all countries worldwide, but is not legally binding. Conversely, Human Rights are also secured in a series of treaties, such as the International Covenant on Civil and Political Rights (ICCPR) and International Covenant on Economic, Social and Cultural Rights (ICESCR), which *are* legally binding, but do not cover all countries. According to the UN Guiding Principles on Human Rights, states have a duty to protect their population against human rights violations caused by third parties (Implementing the United Nations 'Protect, Respect and Remedy' Framework, 2011, Foundational Principles 1). In addition, businesses are required to respect human rights in their activities (Implementing the United Nations 'Protect, Respect and Remedy' Framework, 2011, Foundational Principles 11).

This chapter explains the different ways in which disinformation infringes human rights and menaces the quality of democratic practice. It unpacks precisely which human rights are endangered by disinformation and which aspects of broader democratic norms it undermines. The chapter then points to an issue that is of crucial importance for EU actions on this issue in third countries: while disinformation threatens human rights, the inverse challenge is that counter-disinformation policies can also restrict freedoms and rights.

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

3.1 Impacts on human rights

Freedom of expression is a core value for democracies (Article 19(2) of ICCPR). This includes press freedom and the right to access information. Under human rights law, even the expression of false content is protected, albeit with some exceptions.

Digitalisation and global access to social networks have created a new set of channels for the violation of human rights, which the UN Human Rights Council confirms must apply as much online as they do offline[47]. Digitalisation has amplified citizens' vulnerability to hate speech and disinformation, enhancing the capacity of state and non-state actors to undermine freedom of expression. Looking more closely at various levels of impact, disinformation threatens a number of human rights and elements of democratic politics[48].

3.1.1 Right to freedom of thought and the right to hold opinions without interference

Article 19 of the UDHR states:

'Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.'

Article 19 of ICCPR states:

'1. Everyone shall have the right to hold opinions without interference.

2. Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice.

3. The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary:

(a) For respect of the rights or reputations of others;

(b) For the protection of national security or of public order (order public), or of public health or morals.'

In its 2011 General Comment on Article 19 ICCPR, which covers both freedom of opinion and freedom of expression, the UN Human Rights Committee proclaims that: 'Freedom of opinion and freedom of expression are indispensable conditions for the full development of the person. They are essential for any society. They constitute the foundation stone for every free and democratic society'. Freedom of thought entails a right not to have one's opinion unknowingly manipulated or involuntarily influenced. It is not yet clear where the dividing line is between legitimate political persuasion and illegitimate manipulation of thoughts, but influence campaigns may well breach this right.

3.1.2 The right to privacy

Article 17 of ICCPR states:

'1. No one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to unlawful attacks on his honour and reputation.

[47] UN Human Rights Council Resolutions (2012-2018), The promotion, protection and enjoyment of human rights on the Internet, UN Doc A/HRC/RES/38/7 (5 July 2018), A/HRC/RES/32/13 (1 July 2016), A/HRC/RES/26/13 (26 June 2014), A/HRC/RES/20/8 (5 July 2012).

[48] Kate Jones (2019) disaggregates the impact of online disinformation on human rights in the political context. She identifies concerns with respect to five key rights: right to freedom of thought and the right to hold opinions without interference; the right to privacy; the right to freedom of expression; the right to participate in public affairs and vote in elections.

2. Everyone has the right to the protection of the law against such interference or attacks.'

Article 12 of UDHR stipulates that: 'No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation.'

The use of disinformation can interfere with privacy rights in two ways: by damaging the individual reputation and privacy of the person it concerns in certain circumstances, and by failing to respect the privacy of individuals in its target audience. The Special Rapporteur on the right to privacy states that 'privacy infringements happen in multiple, interrelated and recurring forms facilitated by digital technologies, in both private and public settings across physical and national boundaries'[49]. Online privacy infringements extend offline privacy infringements. Digital technologies amplify their scope and intensify their impact.

The right to privacy in the digital age is exposed to a new level of vulnerabilities, ranging from personal attacks through social media to the harvesting and use of personal data online for micro-targeting messages. However, as the case *Tamiz v the United Kingdom* showed[50] there is a thin line between freedom of expression and the right to privacy. In that particular case, the European Court of Human Rights (ECtHR) reinforced the protection of freedom of expression by confirming a UK Court decision to reject the libel claim of a British politician against Google Inc. because it hosted a blog which published insulting comments against him.

The UN High Commissioner for Human Rights (OHCHR) has affirmed that there is 'a growing global consensus on minimum standards that should govern the processing of personal data by States, business enterprises and other private actors'[51]. These minimum standards should guarantee that the 'processing of personal data should be fair, lawful and transparent in order to protect citizens from being targeted by disinformation that can for instance cause harm to individual reputations and privacy', even to the point of inciting 'violence, discrimination or hostility against identifiable groups in society'. In this context, the EU General Data Protection Regulation (GDPR) is founded on the right to protection of personal data in EU law. The GDPR imposes controls on the processing of personal data, requiring that data be processed lawfully, fairly and transparently.

3.1.3 The right to freedom of expression

The UDHR and the International Covenant on Civil and Political Rights (ICCPR), in Article 19, protect the right to freedom of expression (see full quotation above, under Section 3.1.1).

The right to disseminate and access information is not limited to true information. In March 2017, a joint declaration by the UN Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples all stressed that 'the human right to impart information and ideas is not limited to 'correct' statements, that the right also protects information and ideas that may shock, offend and disturb' [52]. They declared themselves alarmed both 'at instances in which public authorities denigrate, intimidate and threaten the media', as well as others stating that the media is 'the opposition' or is 'lying' and has a hidden political agenda'. They further warned that 'general prohibitions on the dissemination of

[49] OHCHR, Report of the Special Rapporteur on the right to privacy for the 43th session of the Human Rights Council, February 2020.

[50] Tamiz v the United Kingdom (Application no. 3877/14) [2017] ECHR (12 October 2017), seen in European Court of Human Rights upholds the right to freedom of expression on the Internet, Human Rights Law Centre, 2017 [Accessed on 28/02/21]

[51] OHCHR, The right to privacy in the digital age, Report of the United Nations High Commissioner for Human Rights for the 39th session of the Human Rights Council, August 2018.

[52] Organization for Security and Co-operation in Europe (OSCE). Joint declaration on freedom of expression and 'fake news', disinformation and propaganda, March 2017.

information based on vague and ambiguous ideas, including 'false news' or 'non-objective information', are incompatible with international standards for restrictions on freedom of expression [...] and should be abolished'.

In April 2020, the same signatories were parties to a new 'Joint Declaration on Freedom of Expression and Elections in the Digital Age', in which they expressed 'grave concern' about the threats and violent attacks that journalists may face during elections, adding that targeted smear campaigns against journalists, especially female journalists, undermine their work as well as public trust and confidence in journalism. The agreed text not only calls for protecting freedom of expression, but also points at political authorities passing laws limiting rights, restricting Internet freedom or 'abusing their positions to bias media coverage, whether on the part of publicly-owned or private media, or to disseminate propaganda[53] that may influence election outcomes'. The signatories put forward a reminder that the states' obligation to respect and protect freedom of expression is especially pronounced in relation to female journalists and individuals belonging to marginalised groups[54].

3.1.4 Economic, social and cultural rights

Disinformation impacts not only the political sphere, but also economic, social and cultural aspects of life, from personal mindsets about vaccinations to disavowing cultures or different opinions. Disinformation feeds polarisation and erodes trust both within institutions and amongst communities. Such manipulation tactics can damage personal rights to health and education, participation in cultural life and membership of a community.

There are several economic, social and cultural rights that can be disrupted by disinformation, such as those included in Article 25(1) UDHR: 'Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control'.

Article 12 of the ICESCR affirms:

'1. The States Parties to the present Covenant recognise the right of everyone to the enjoyment of the highest attainable standard of physical and mental health.'

The most prominent recent example relates to disinformation around COVID-19, which has distorted freedom of choice even within a health context (see also Chapter 4 on COVID-19). Blackbird[55], an organisation whose purpose it is to enhance decision making and empower the pursuit of information integrity'[56], released studies analysing the volume of disinformation generated on Twitter as a result of the COVID-19 outbreak. In one of its reports[57], Blackbird identified one of the disinformation campaigns unfolding in Twitter as 'Dem Panic', which had the goal of de-legitimising the Democratic Party in the US for their early warnings about the coronavirus and the need to introduce preventative measures. Downplaying effects of the virus can clearly have negative impacts on public health. The Lancet warned in October 2020 that the anti-vaccine movement together with digital platforms are becoming wealthier by hosting and spreading disinformation campaigns in social media[58]. It is claimed that 'anti-vaxxers have

[53] In Article 20(1) of the 1966 International Covenant on Civil and Political Rights (ICCPR), the term propaganda 'refers to the conscious effort to mould the minds of men [sic] so as to produce a given effect' (Whitton, 1949; via Jones, 2019).
[54] OSCE, Joint Declaration on Freedom of Expression and Elections in the Digital Age, April 2020.
[55] Blackbird AI, Disinformation reports, 2020.
[56] The founders of this organisation 'believe that disinformation is one of the greatest global threats of our time impacting national security, enterprise businesses and the general public'. Accordingly, their 'mission is to expose those that seek to manipulate and divide'.
[57] Blackbird AI, COVID-19 Disinformation Report – Vol. 2, March 2020.
[58] Center for Countering Digital Hate, The Anti-vaxx industry. How Big-Tech powers and profits from vaccine misinformation' in T. Burki, The online anti-vaccine movement in the age of COVID-19, *The Lancet*, Vol. 2, October 2020.

increased their following by at least 7·8 million people since 2019' and hence 'the anti-vaccine movement could realise USD 1 billion in annual revenues for social media firms'.

3.2 Impact on democratic processes

3.2.1 Weakening of trust in democratic institutions and society

Disinformation has an impact on the basic health and credibility of democratic processes. This has become the core of recent positions taken by international organisations, such as Resolution 2326 (2020) of the Parliamentary Assembly of the Council of Europe (PACE) expressing concern 'about the scale of information pollution in a digitally connected and increasingly polarised world, the spread of disinformation campaigns aimed at shaping public opinion, trends of foreign electoral interference and manipulation'[59]. Information and shared narratives are a precondition for good quality democratic public discourse.

In this context, the European Parliament views disinformation as an 'increasing systematic pressure' on European societies and their electoral stability[60]. The European Commission's strategy *Shaping Europe's Digital Future*[61] considers that 'disinformation erodes trust in institutions along with digital and traditional media and harms our democracies by hampering the ability of citizens to take informed decisions'. It also warns that disinformation is set to polarise democratic societies by creating or deepening tensions and undermining democratic pillars such as electoral systems.

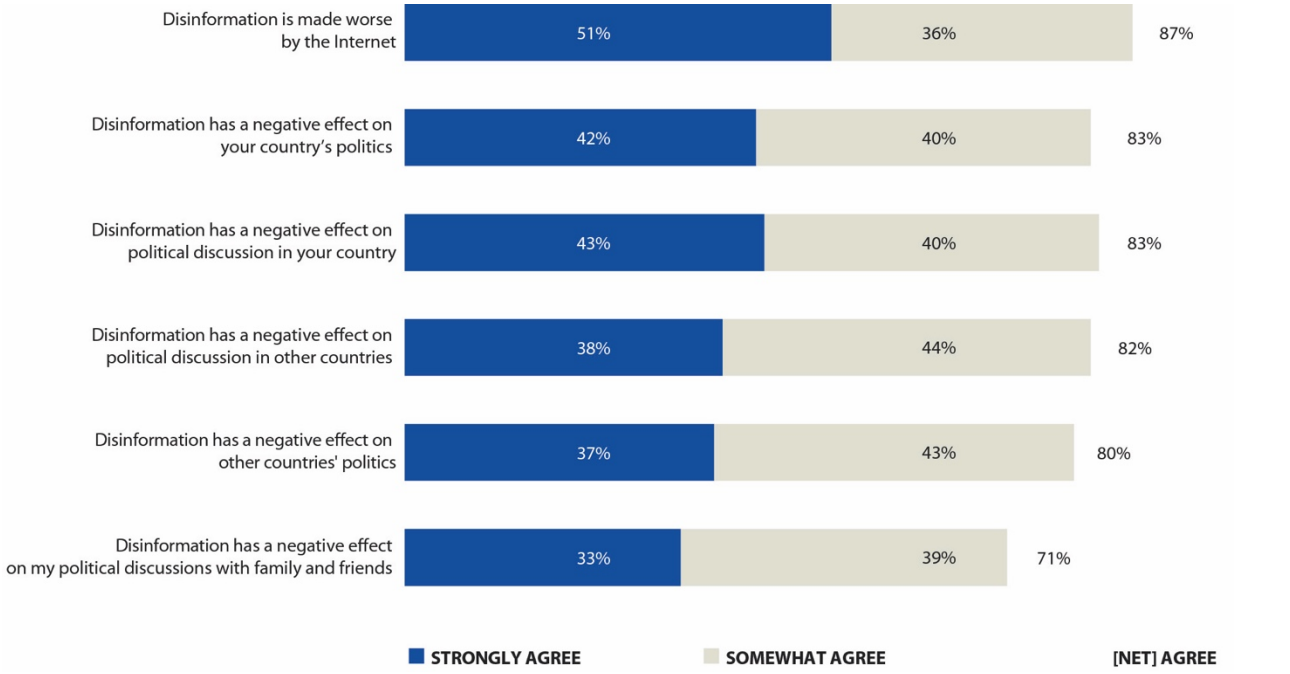
There are a number of ways in which disinformation weakens democratic institutions. These include the use of social media to channel disinformation in coordinated ways so as to undermine institutions' credibility. As trust in mainstream media has plummeted[62], alternative news ecosystems have flourished. Online platforms' business model pushes content that generates clicks and this has increased polarisation. This favours the creation of more homogeneous audiences, undercuts tolerance for alternative views[63].

Figure 2 below suggests that around 80 % of people believe that disinformation has negative impacts in their own countries' politics, in other countries' politics and in political discussions among families and friends, which increases polarisation.

Surveys also show that disinformation can sow distrust in different pillars of democratic institutions, including public institutions such as governments, parliaments and courts or their processes, public figures, as well as journalists and free media[64]. For example, a survey undertaken by Ipsos Public Affairs and Centre for International Governance Innovation (CIGI) reports that, due to the spread of disinformation, many citizens have less trust in media (40 %) and government (22 %)[65].

[59] PACE, Democracy Hacked? How to Respond?, Resolution 2326 of the Parliamentary Assembly of the Council of Europe on 31 January 2020 (9th Sitting), January 2020.
[60] European Parliament, European Parliament resolution of 23 November 2016 on EU strategic communication to counteract propaganda against it by third parties, P8_TA(2016)0441, November 2016.
[61] European Commission, Tackling online disinformation, webpage. [Accessed on 15 October 2020]
[62] Nic Newman, Richard Fletcher, Anne Schulz, Simge Andi and Rasmus Kleis Nielsen, Reuters Institute Digital News Report 2020. *Reuters Institute for the Study of Journalism*, 2020.
[63] Massimo Flore, Understanding Citizen's Vulnerabilities: form Disinformation to Hostile Narratives, *JRC Technical Report*, European Commission, 2020.
[64] IPSOS Public Affairs and Centre for International Governance Innovation (CIGI), Internet security and trust, CIGI IPSOS Global Survey 2019, Vol 3, 2019.
[65] IPSOS and CIGI, 2019, p 138.

Figure 2. The Political Impacts of Disinformation[66]



Source: Authors' own elaboration based on CIGI-Ipsos, 2019 CIGI-Ipsos Global Survey on Internet Security and Trust., 2019; CIGI-IPSOS, Internet security and Trust. Part 3, 2019.

3.2.2 The right to participate in public affairs and election interference

Article 21 of the UDHR states:

'1. Everyone has the right to take part in the government of his country, directly or through freely chosen representatives';

3. The will of the people shall be the basis of the authority of government; this will shall be expressed in periodic and genuine elections which shall be by universal and equal suffrage and shall be held by secret vote or by equivalent free voting procedures.'

Article 25 of the ICCPR defends that:

'Every citizen shall have the right and opportunity, without any of the distinctions mentioned in Article 2 and without unreasonable restrictions:

- To take part in the conduct of public affairs, directly or through freely chosen representatives;
- To vote and to be elected at genuine periodic elections which shall be by universal and equal suffrage and shall be held by secret ballot, guaranteeing the free expression of the will of the electors;
- To have access, on general terms of equality, to public service in his country.'

According to the UN Human Rights Committee, states are obliged to ensure that 'Voters should be able to form opinions independently, free of violence or threat of violence, compulsion, inducement or

[66] The percentage refers to people interviewed in 25 different countries (Australia, Brazil, Canada, China, Egypt, France, Germany, Great Britain, Hong Kong (China), India, Indonesia, Italy, Japan, Kenya, Mexico, Nigeria, Pakistan, Poland, Russia, South Africa, Republic of Korea, Sweden, Tunisia, Turkey and the United States) accounting for 1.000+ individuals in each country during December 2018 and February 2019. See the original source (IPSOS-CIGI, 2019) for more information about the methodology.

manipulative interference of any kind'[67]. Election interference can be defined as unjustified and illegitimate ways of influencing people's minds and voters'choices, thereby reducing citizens' abilities to exercise their political rights[68]. This means that the right to vote has to be exercised without interference with freedoms of thought and opinion, with the right to privacy and without hate speech. Many governments' use of disinformation contradicts this injunction. Even where they are not directly using disinformation in electoral campaigns, other states may be falling short in protecting this right on behalf of their citizens. Foreign states and non-state actors are also able to influence and undermine elections through digital disinformation[69]. Russia's Internet Research Agency (IRA) purchased around 3 400 advertisements on Facebook and Instagram during the US 2016 election campaign and, according to a 2019 Report 'Russian-linked accounts reached 126 million people on Facebook, at least 20 million users on Instagram, 1.4 million users on Twitter, and uploaded over 1.000 videos to YouTube'[70].

Whether or not successful, manipulating elections by affecting voters' opinions and choices through disinformation damages democracy and creates a trail of doubt as to whether democratic institutions actually work well in reflecting citizens' choices.

3.3 Digital violence and repression

Article 20(2) of the ICCPR states that:

'2. Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.'

Disinformation is associated with a rise in more dramatic and grave digital violence. Digital violence has been defined as using mobile phones, computers, video cameras and similar electronic devices with intention to frighten, insult, humiliate or hurt a person in some other way[71]. The term 'cyber-violence', includes a range of controlling and coercive behaviours, such as cyber-stalking, harassment on social media sites, or the dissemination of intimate images without consent. Here, the instigators of digital violence can be state and non-state actors as well as private groups or individuals.

In this context, digital repression means the coercive use of information and communication technologies by a state to exert control over not only potential, but also existing challenges and challengers. Digital repression includes an assortment of tactics through which states can use digital technologies to monitor and restrict the actions of its citizens, including digital surveillance, advanced biometric monitoring, disinformation campaigns and state-based hacking[72]. These actions self-evidently infringe core democratic rights like the right to privacy and to freedom of expression.

In multiple political systems, cyber militias and 'troll farms' are used to drown out dissenting voices, accusing them of spreading 'fake news' or being 'enemies of the people', a sort of censorship through noise[73]. Experts warn against the practice of 'state-sponsored trolling', which consists of governments

[67] UN Committee on Human Rights, General Comment 25, 'The Right to Participate in Public Affairs, Voting Rights and the Right to Equal Access to Public Service', 1510th meeting (fifty-seventh session), 12 July 1996.
[68] UNHR, 'Monitoring Human Rights in the Context of Elections', in *Manual on Human Rights Monitoring*, chapter 23, 2011.
[69] Suzanne Spaulding, Devi Nair, and Arthur Nelson, Beyond the Ballot: How the Kremlin Works to Undermine the U.S. Justice System, *Center for Strategic and International Studies*, May 2019.
[70] Robert Mueller, Report on the Investigation into Russian Interference in the 2016 Presidential Election, U.S. Department of Justice, March 2019.
[71] Dragan Popadic Dragan and Dobrinka Kuzmanovic, Utilisation of Digital Technologies, Risks, and Incidence of Digital Violence among students in Serbia. Unicef Report, 2013.
[72] Steven Feldstein, How Artificial Intelligence Is Reshaping Repression, *Journal of Democracy*, Vol. 30, 2019, pp 40-52.
[73] Peter Pomerantsev, 'Human rights in the age of disinformation', *Unherd*, published on 8 July 2020. [Accessed on 15 August 2020]

creating online content to attack opposition and discredit critical voices of dissent, thereby cutting across basic standards of democratic debate which should be open and pluralistic.

While this study's remit is limited to legal content, current disinformation challenges cannot be divorced from the concept of 'hate speech', commonly referring to any communication that disparages a person or a group on the basis of some characteristic such as race, colour, ethnicity, gender, sexual orientation, nationality, religion, or other characteristic. Disinformation can be an important part of this. Incitement to hatred and discrimination (without advocacy of violence) is illegal in many EU Member States, unlike in the United States. To address this challenge, in May 2016 the European Commission agreed with Facebook, Microsoft, Twitter and YouTube, on a Code of Conduct on Countering Illegal Hate Speech Online[74].

Coordinated online hate speech against racial and ethnic minorities has led to violence in different places and disinformation has been used to attack minorities and human rights defenders around the world[75]. For instance: in Sri Lanka and Malaysia targeted disinformation resulted in an outburst of violence against Muslims; in Myanmar, the military used Facebook to incite violence against the Rohingya; and rumours about Muslims in India circulating on WhatsApp have resulted in lynchings[76].

3.4 Counter-disinformation risks

Disrupting human rights and democracy by disinformation is clearly of serious concern. However, there is another side to the democratic equation, namely that action *against* disinformation also carries risks. Tackling disinformation through a human rights prism involves difficult trade-offs and delicate policy balances. Disinformation itself threatens to produce core breaches in human rights. Hence, countering disinformation is an important contribution in efforts to safeguard global human rights. Yet, countering disinformation can also itself constrict human rights. Even if disinformation can easily damage human rights, both within the EU and at international level, the legal and political abuse of what has been labelled the 'fight against fake news' has in some countries also resulted in reduced freedom of expression and political dissent. This is becoming a more serious problem both inside and outside the European Union[77].

Consequently, in defending measures to tackle disinformation, the European Union must be careful to tackle both human rights impacts resulting from disinformation and any rights abuses inadvertently caused by attempts to counter disinformation in such a way that does not encourage third-country governments' human rights breaches. Erosion of rights can be caused, for instance, by: government interferences with internet services; state censorship or restrictions to online speech; and obstacles to the proper functioning of media outlets. Any actions – be they legal, administrative, extra-legal or political – which have the potential to breach freedoms of expression, assembly and association, will ultimately result in an erosion of the democratic space[78].

[74] European Commission, Code of Conduct on Countering Illegal Hate Speech Online, June 2016.
[75] Interview with and MEP from the GUE-NGL group, 29 October 2020.
[76] Timothy McLaughlin, 'How WhatsApp Fuels Fake News and Violence in India,' *Wired,* Published on 12 December 2018; Vindu Goel, Suhasini Raj and Priyadarshini Ravichandran, 'How WhatsApp Leads Mobs to Murder in India,' *New York Times*, published on 18 July 2018. [Accessed on 15 August 2020]
[77] ARTICLE 19, Responding to 'Hate Speech': Comparative Overview of Six EU Countries, 2018.
[78] European Partnership for Democracy in collaboration with the Netherlands Institute for Multiparty Democracy, Thinking democratically: recommendations for responding to the phenomenon of 'shrinking space'., EPD Report, 2020.

There is also a risk that the activities of the digital platforms in combating disinformation may restrict freedom of expression. In two reports published in 2018[79] and 2019[80], the former UN Special Rapporteur on Freedom of Expression David Kaye warned against regulation that entrusts platforms with even more powers to decide on content removals without public oversight. As moves towards platform regulation are developing within the EU, similar regulatory updates are now under discussion in the United States.

Awareness about dangers associated with using counter-disinformation techniques which could inadvertently compromise rights has also grown in the UN. In December 2019, a Russian-led and Chinese-backed resolution on cybercrime entitled 'Countering the use of information and communications technologies for criminal purposes', was adopted by 79 votes to 60 with 33 abstentions, despite opposition from several major Western powers[81]. Opponents of the text feared that the resolution would serve to erode freedom of expression online. It sought to make advances in the 'fighting of cybercrime' through information control and the suppression of political dissidents. Votes in favour were cast by countries such as Cambodia, North Korea, Burma, Venezuela, Algeria, Syria, Belarus and Kazakhstan. All EU member states, Canada, Australia and the United States voted against. As geopolitical debates about cybersecurity have deepened, it is increasingly apparent that human rights risk becoming a casualty of *both* disinformation and counter-disinformation agendas.

[79] UN Special rapporteur on the promotion and protection of the right to freedom of opinion and expression, 'Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment', to the General Assembly, 73rd session, 29 August 2018.

[80] UN Special rapporteur on the promotion and protection of the right to freedom of opinion and expression, Report on the promotion and protection of the right to freedom of opinion and expression on 'Surveillance and human rights, to the, 41st session Human Rights Council, 24 June-12 July 2019.

[81] Carme Colomina, Techno-multilateralism: The UN in the age of post-truth diplomacy, in Bargués, P., *UN@75: Rethinking multilateralism*, CIDOB Report, vol.6, 2020.

4 The impact of disinformation during the COVID-19 crisis

'As COVID-19 spreads, a tsunami of misinformation, hate, scapegoating and scare-mongering has been unleashed'

Antonio Guterres, United Nations Secretary-General

Key takeaways

- The COVID-19 infodemic has led to a spike in 'fake news' about hoaxes, pseudoscience and conspiracy theories, thereby breeding distrust in public institutions and putting lives at risk.
- Most of the disinformation spread during the COVID-19 crisis has been reconfigured content, false or misleading, coming from prominent public figures.
- The pandemic has exacerbated pre-existing trends of foreign interference but has also unleashed new dynamics. In this acceleration of disinformation, tech companies have become more powerful in shaping public opinion but also political actors with an essential responsibility in the global responses to disinformation.
- Some regulations targeting disinformation during the pandemic have involved new limits on press freedom and censorship tools that are likely to persist beyond the COVID-19 crisis.

The COVID-19 pandemic has forced the EU to adapt its external policies so as to counter evolving threats and online manipulation[82]. During this time of uncertainty, there has been an increase in information consumption. This has led to data overexposure, a spike in 'fake news', hoaxes, pseudoscience and conspiracy theories, breeding distrust in public institutions. The World Health Organisation (WHO) noted that global citizens were victims of both the pandemic itself and an 'infodemic' that arose around it. It launched a pilot programme called EPI-WIN[83], the Information Network for Epidemics, to combat misinformation.

Between 20 January and 10 February 2020, two million messages posted on Twitter – 7 % of the total – spread conspiracy theories about the coronavirus. Amidst the need for information and reliable sources, an increased awareness of the vulnerability caused by disinformation arose. According to Reporters Without Borders, as many as 74 % of Internet users expressed their concern about 'fake news' on social media[84]. This overload of unreliable information spread rapidly among the population, making it hard for people to find trustworthy sources and reliable guidance just when they needed it most. Disinformation put lives at risk as it led people to ignore official health advice and engage in risky behaviour. In Madagascar, for instance, president Andry Rajoelina showed his support for an unproven herbal tea to cure COVID-19[85]. Just one month earlier, US president Donald Trump suggested that disinfectant could kill the virus[86]. As the Reuters Institute pointed out, politicians, celebrities and other prominent public figures were

[82] European Commission, Joint Communication to the European Parliament, the Europan Council, the Council, The European Economic and Social Commitee and the Committee of the Regions. Tackling COVID-19 disinformation - Getting the facts right, JOIN(2020) 8 final, June 2020.
[83] See WHO EPI-WIN webpage.
[84] See Reporters without Borders Tracker-19 webpage.
[85] Arin Baker, 'Could It Work as a Cure? Maybe.' A Herbal Remedy for Coronavirus Is a Hit in Africa, But Experts Have Their Doubt, *TIME*, published on 22 May 2020.
[86] Clark Dartunorro, Trump suggests 'injection' of disinfectant to beat coronavirus and 'clean' the lungs, *NBC news*, published on 24 April 2020.

conspicuous in spreading misinformation online during the pandemic. One research survey revealed that falsehoods coming from presidents, opinion-makers and global public figures represented 20 % of total misinformation, accounting for 69 % of such actions on social media[87].

Fact-checking organisations in Africa report to have debunked more than 1 000 of such misleading reports since the pandemic's onset[88]. COVID-19 disinformation has recycled pre-existing conspiracy theories and harmful medical practices. According to a Reuters Institute research survey[89], 87 % of online disinformation on the pandemic spread between January and March 2020 used reconfigured content, impersonated genuine sources or updated longstanding conspiracy theories.

Table 2. COVID-19 Information Disorder
(Identified by international fact-checkers between January and March 2020[90])

Types	*Contents*	*Scope*
Reconfiguration	• 'Misleading content' containing some true information, but the details were reformulated, selected and re-contextualised in ways that made them false or misleading • Imposter Content: when genuine sources are impersonated • Images or videos labelled or described as being something other than what they are • Conspiracy theories	59 % of disinformative content identified[91]
Completely fabricated	New content 100 % false, designed to deceive and do harm	38 % of disinformative content identified[92]
Satire/Parody	No intention to cause harm, but with potential to fool	3 % of the misinformative content identified[93]
Monetised disinformation	Content that potentially capitalises on promoting or advocating for harmful health or medical claims or practices.	IBM's X-Force Threat Intelligence identified in March a 6 000 % increase in COVID-19 related spam, compared to the year before. From January to August 2020,Google blocked or removed over 82.5 million COVID-19 related ads, suspended more than 1 300 accounts from EU-based advertisers.

Source: Author's own elaboration based on Brennen et al., Types, Sources and Claims of COVID-19 Misinformation, Reuters Institute and Google, EU & COVID-19 Disinformation Google Report, September 2020'.

[87] Scott J. Brennen, Felix Simon, Philip N. Howard, Rasmus Kleis Nielsen. Types, Sources, and Claims of COVID-19 Misinformation, Reuters Institute for the Study of Journalism, April 2020.
[88] WHO, Landmark alliance launches in Africa to fight COVID-19 misinformation, *WHO Africa*, published on 3 December 2020.
[89] This Reuters Institute research analysed 225 pieces of misinformation sampled from a corpus of English-language fact-checks gathered by First Draft, focusing on content rated false or misleading. See S. Brennen et al., 2020.
[90] S. Brennen et al., 2020.
[91] S. Brennen et al., 2020.
[92] S. Brennen et al., 2020.
[93] S. Brennen et al., 2020.

During the pandemic, social media acted as a double-edged sword[94]. On the one hand, digital platforms were useful in promoting debate within the scientific community and disseminating valuable information as well as investigative results. On the other hand, they also helped disseminate flawed studies and misinformation. Even if the pandemic has strengthened the pressure of accountability on big tech companies, their capacity to filter information has been much weaker than their impact on the circulation of content. However, in the framework of the Code of Practice agreed between the European Commission and big tech companies, there have been monthly reports on the actions undertaken by these platforms to tackle COVID-19-related disinformation[95]. Social media platforms were also asked to collaborate with the UN system to prevent the spread of disinformation. A team of WHO 'myth-busters' worked with search and media companies such as Facebook, Google, Pinterest, Tencent, Twitter, TikTok, YouTube and others to counter the spread of lies about the pandemic[96]. In December 2020, the WHO launched its *Africa Infodemic Response Alliance* to coordinate the actions of international organisations, governments and fact-checkers to tackle disinformation on COVID-19[97].

Social networks' policies of content control have been strengthened to some degree. For instance, Twitter reported that it had removed 20 000 tweets, challenged 8.5 million accounts and has the capacity to detect 92 % of content which violates Twitter's rules and policies.[98] Microsoft centred its efforts on making sure that searches about coronavirus would end at authoritative webpages with trustworthy information. Facebook promised to use 'strong warning labels' tagging disinformation that they removed, for instance, a post by Brazilian president Jair Bolsonaro claiming that hydroxychloroquine was a remedy for COVID-19. Instagram filtered and removed COVID-19 related content that was not posted by official health organisations.

However, such advances remain insufficient. The Oxford Internet Institute shows that Facebook's content moderating policies have detected less than 1 % of videos with 'disinformative' content about the virus[99]. An Avaaz report published in August 2020 pointed out that o**nly 16 % of all health misinformation analysed had a warning label from Facebook.** Despite their content being fact-checked by Avaaz, the other 84 % of articles and posts sampled remained online without warnings.[100] One of the reasons for this failure lies in the possibilities for cross-sharing between different social networks[101] that make traceability more difficult.

As Europol states[102], coordinated disinformation campaigns can feed distrust in the ability of democratic institutions to deliver effective responses to the current health crisis[103]. Subsequently, criminal organisations as well as state and state-backed actors have sought to exploit the public health crisis to advance their interests. According to Lea Gabrielle, Special Envoy and Coordinator of the Global Engagement Centre at the US State Department, countries such as China, Russia and Iran have sharply

[94] Carlos Chaccour and Rafael Vilasanjuan, Infodemic: has the epidemic of misinformation affected the response to COVID-19?v, ISGlobal, Barcelona Institute for Global Health, September 2020.

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

[96] See WHO Mythbusters webpage.

[97] Scott J. Brennen, Felix Simon, Philip N. Howard, Rasmus Kleis Nielsen. Types, Sources, and Claims of COVID-19 Misinformation, Reuters Institute for the Study of Journalism, April 2020.

[98] S. Brennen et al., 2020.

[99] Aleksi Knuutila, Aliaksandr Herasimenka, Hubert Au, Jonathan Bright and Philip N. Howard, The Spread of Misinformation Videos on Social Media and the Effectiveness of Platform Policies, Oxford Internet Institute, September 2020.

[100] Avaaz, 'Facebook's Algorithm: A Major Threat to Public Health', Avaaz Report, August 2020.

[101] EU Disinfo Lab, Hydroxychloroquine and Facebook: The Challenge of Moderating Scientifically Debatable Claims. Published on 27 October 2020 [Accessed 07 January 2021]

[102] EUROPOL, Catching the virus cybercrime, disinformation and the COVID-19 pandemic, April 2020.

[103] EUROPOL, 2020.

increased their dissemination of disinformation about the coronavirus since January 2020, repeating and amplifying one another's propaganda and falsehoods[104]. Conspiracy theories spun about false remedies and dark origins in the manufacture of the virus, incited attacks on 5G telecommunications masts in Europe and fuelled anti-Asian racism[105].

4.1 Acceleration of existing trends

Disinformation about the pandemic has become a global phenomenon. Acceleration of the digitalisation process, together with hyper-connectivity and Internet penetration capacity have amplified effects of what the EU High Representative for foreign affairs and security policy, Josep Borrell, called the 'global battle of narratives'[106]. Examples gathered by the Oxford Internet Institute show how Russian and Iranian outlets generated polarising content in Spanish to target Latin America and Spanish-speaking social media users in the United States[107]. Furthermore, numerous Chinese state-backed outlets targeted non-English audiences across the globe to support China's soft power ambitions. For instance, China launched a CGTN *Español* channel, targeting Spanish-speaking countries – especially in Latin America and the Caribbean where China has economic, political and cultural interests. Meanwhile, Iran had also launched a Latin American-focused channel with *HispanTV* in 2012. This channel represents the Islamic Republic of Iran Broadcaster's (IRIB) attempt to establish stronger connections with their allies in Latin America and counterbalance Western media. These channels were increasingly used during the pandemic, causing a surge in the level of disinformation challenging Latin American democracies and affecting efforts to fight COVID-19[108]. Finally, Turkey has also shown recent interest in Arabic, Spanish and German-speaking audiences, within the context of President Recep Tayyip Erdogan's use of soft power tools to bolster his country's global reputation.

4.2 The impact of the COVID-19 *infodemia* on Human Rights

The COVID-19 'infodemic' has demonstrated how disinformation may prevent individuals from realising their right to health, given legal force through Article 12 of the International Covenant on Economic, Social and Cultural Rights (ICESCR). The impact of disinformation, particularly at a time of uncertainty around the pandemic and a hunger for information, is that individuals may fail to follow the guidance of science and authority in accessing the health care and medicines they need, in participating in measures necessary to combat the pandemic and in having vaccines that may protect them and their societies from further disease. 'Without the appropriate trust and correct information, diagnostic tests go unused, immunisation campaigns (or campaigns to promote effective vaccines) will not meet their targets and the virus will continue to thrive'[109].

[104] Special envoy and coordinator of the Global Engagement Center at the US State Department, Lea Gabrielle, hearing before the subcommittee on State Department and USAID management, international operations, and bilateral international development of the committee on Foreign Relations. US Senate, 5 March 2020. Available here.
[105] Thomas, Seal, 5G Virus Conspiracy Theory Drives Phone Mast Attacks in UK, *Bloomberg*, published on 6 April 2020. [Accessed 10 October 2020]
[106] EEAS, The Coronavirus pandemic and the new world it is creating, published on 23 March 2020. [Accessed 10 October 2020]
[107] Katarina Rebello, Christian Schwieter, Marcel Schliebs, Kate Joynes-Burgess, Mona Elswah, Jonathan Bright & Philip N. Howard. COVID-19 News and Information from State-Backed Outlets Targeting French, German and Spanish-Speaking Social Media Users, University of Oxford - Oxford Internet Institute, June 2020.
[108] Tom Philips, David Agren, Dan Collyns and Uki Goñi, 'Tsunami of fake news hurts Latin America's effort to fight coronavirus', *The Guardian*, 26 July 2020.
[109] Joint statement by WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, and IFRC, Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation, September 2020.

As noted previously, in responding to the COVID-19 outbreak, at least 17 governments have enforced regulations targeting disinformation during the pandemic which involve new censorship[110]. The International Press Institute Information's COVID-19 tracker has been monitoring press freedom restrictions, warning that they may outlast the pandemic[111].

According to the International Observatory of Human Rights, many governments 'have cracked down hard on media outlets and journalists' trying to provide accurate information and facts. This London-based NGO gathered a long list of rights violations, allegedly triggered by the emergency resulting from the fight against this pandemic. In Iran, the government has imposed sweeping restrictions on journalistic coverage as part of a systematic effort to downplay the scope of this public health crisis. Similarly, Egypt has pressured journalists to understate the number of infections and forced a Guardian journalist to leave the country after she reported on a scientific study which stated that Egypt was likely to have many more coronavirus cases than have been officially confirmed[112]. Turkey has launched legal proceedings against 316 social media users, charging them with inciting hatred and enmity by spreading concern about COVID-19[113]. In Russia, the state media regulator, *Roskomnadzor*, ordered two media outlets to remove COVID-19 reports from their websites and social media[114]. The regulator also published a warning that it would take punitive measures against the 'dissemination of false information' and attempts to 'sow panic among the public and provoke public disturbance'. All this exemplifies a violation of the right to information and a free press.

Reporters Without Borders (RSF) has denounced China's coronavirus disinformation campaign which was designed to drown out critics who blame Beijing for the virus' spread. Chinese officials have gone so far as to suggest that it might have been the US army that 'brought the epidemic to Wuhan' or that the coronavirus might have been 'circulating in parts of Italy before doctors were aware of the outbreak in China'[115]. As a further example, the government in Honduras responded to the outbreak by suspending a clause in its constitution that prohibits censorship and protects the right to free expression.

[110] International Press Institute, Rush to pass 'fake news' laws during COVID-19 intensifying global media freedom challenges, October 2020.

[111] See IPI COVID-19 Media Freedom Monitor.

[112] Michael Safi, Egypt forces Guardian journalist to leave after coronavirus story, *The Guardian*, Published on 26 March 2020. [Accessed 10 October 2020]

[113] EU vs. Disinfo, EEAS SPECIAL REPORT UPDATE: SHORT ASSESSMENT OF NARRATIVES AND DISINFORMATION AROUND THE COVID-19 PANDEMIC , April 2020.

[114] Alexander Avilov, Russian News Outlets Ordered to Take Down 'Fake' Coronavirus News, *The Moscow Times*, Published on 20 March 2020. [Accessed 10 October 2020]

[115] Reporters Without Borders, Beware of China's coronavirus disinformation, RSF says, *RWB News*, published on 18 April 2020.

5 Mapping responses: Legislative and regulatory bodies, corporate activities and civil society

'What's disappeared from the internet isn't truth so much as trust'

Philip Seargent in *The Art of Political Storytelling. Why stories win votes in post-truth politics*

Key takeaways

- Three groups of actors are involved in responses to disinformation: legislative and regulatory bodies, private sector (digital platforms) and civil society. Certain responses target the actors deemed responsible for disinformation; some target the disruptive techniques used; others aim at improving citizens resilience to disinformation.
- Most of the actions taken by the largest social platforms are related to content curation. However, in line with a broader concept of disinformation, digital platforms' responsibility should not be narrowed to the factuality of content spread, but to the complex structure that determines how content is transmitted, including the business model that can enforce systemic biases.
- One of the main problems in holding social platforms accountable for the spread of disinformation is how to determine who is responsible for the content posted.
- There is a risk with some of these responses of turning governments or private digital platforms into arbiters of acceptable speech.

In taking measures to address the entire spectrum of disinformation responses, it is important to look at the different actors involved, together with the typology of decisions and actions taken. Certain responses target the actors deemed responsible for disinformation, some target the disruptive techniques used, whilst others aim at improving citizens resilience to disinformation.

The UNESCO Working Group on Freedom of Expression and Addressing Disinformation distinguishes four top-level categories or responses to disinformation[116]:

- identification responses whose goal is to raise awareness on the deceptive content and the importance of truth;
- responses aimed at producers and distributors, including legislative and regulatory decisions;
- responses aimed at production and distribution mechanisms, which involve the role of technological platforms;
- and responses aimed at the targeted audiences of disinformation campaigns, in particular involving measures to build resilience.

[116] Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.

Table 3. Disinformation responses

Type of responses	Measures and actions	Actors involved
Identification responses	• Monitoring and fact-checking responses • Investigative responses	News organisations, internet communication companies, academia, civil society organisations and independent fact-checking organisations
Responses aimed at instigators, agents and intermediaries	• Legislative, pre-legislative, and policy responses (to deter disinformation or affirm the right of freedom of expression) • National and international counter-disinformation campaigns • Electoral responses (to protect integrity and credibility of elections, through measures that detect, track, and counter disinformation)	Legislative and regulatory bodies, international organisations This category of responses needs a multi-dimensional approach involving a combination of monitoring and fact-checking, legal, curatorial and technical actors.
Responses aimed at the distribution mechanisms	• Curatorial responses (editorial and content policies and 'community standards') • Technical and algorithmic responses • De-monetisation responses	Regulatory bodies and international organisations (for instance, ensuring privacy-preserving environments or facilitating independent arbitration), social platforms
Responses aimed at the target audiences of disinformation campaigns	• Ethical and normative responses • Educational responses • Empowerment and credibility labelling efforts • Support for a free and diverse media	Regulatory bodies, international organisations, academia and education systems, news organisations

Source: Author's own elaboration based on Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.

From the point of view of actors involved, these responses could be divided into three groups: legislative and regulatory bodies, the private sector (digital platforms) and civil society.

5.1 Legislative and regulatory bodies

Different political and regulatory traditions play a role in shaping responses to online disinformation. As seen in previous sections, while some governments have been considering how to respond to disinformation without damaging pluralism and human rights, others are using legislation against disruptive content to limit freedoms. By March 2020, at least 28 countries had passed laws related to disinformation, either updating existing regulations or passing new rulings. The scope of established legislation varies from media and electoral laws to cybersecurity and penal codes[117]. The Poynter Institute published a guide of anti-disinformation actions around the world, mapping the different responses[118]. These range from bills to punish politicians nationwide for 'the dissemination, promotion or financing of false news' – e.g. Chile in 2019 –, to China's criminalisation of starting or spreading rumours that 'undermine economic and social order' in 2016[119], media regulations facilitating a crackdown on journalism – e.g. in Cameroon or Indonesia –, or cyber security laws, like those existing in Vietnam, requiring platforms such as Facebook to delete content at the government's request[120].

Following a completely different approach, Canada and France, for instance, introduced laws to improve tech platforms' transparency on political advertising, requiring social media companies to create ad repositories. However, while French legislation also enables its broadcasting agency to suspend or terminate broadcasters under the influence of foreign states if they spread false information likely to undermine electoral integrity, in Canada the government has created a non-partisan panel taking decisions on disinformation independently[121]. In Uruguay, political parties signed an agreement proposed by the national press association against disinformation during the elections of 2019[122]. The European Union's approach in this regard is presented in Chapter 6.

Beyond governmental actions, much has been undertaken by international organisations, not only in establishing legislative frameworks but also in engaging with different initiatives. A selection of these multilateral responses is presented below.

In 2018, the G7 established the Rapid Response Mechanism (RRM) following Canada's initiative. The RRM developed out of the *Charlevoix Commitment on Defending Democracies from Foreign Threats*[123] and aims at coordinating information sharing and responses to 'malign and evolving threats to G7 democracies'. The RRM is not limited to disinformation, but also covers threats to democracy more generally. The RRM is designed inter alia to: respond to foreign interference; share lessons and best practices that promote 'free, independent and pluralistic media', and freedom of expression; and engage with private actors to better tackle 'malicious misuse of information'.

[117] K. Bontcheva and J. Posetti, 2020.
[118] Daniel Funke and Daniela Flamini, A guide to anti-misinformation actions around the world, *Poynter*, 2019.
[119] Maria Repnikova, China's Lessons for Fighting Fake News, *Foreign Policy*, Published on 6 September 2018.
[120] Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.
[121] Chris Tenove, Protecting democracy from disinformation: normative threats and policy responses, *The International Journal of Press/Politics*, Vol 25(3), May 2020, pp 517-537
[122] Silvia Higuera, Por iniciativa de asociación de periodistas, partidos políticos de Uruguay firmarán pacto contra la desinformación, *LatAm Journalism Review*, Published on 24 April 2019.
[123] Government of Canada, Charlevoix commitment on defending democracy from foreign threats, January 2019.

On 13 February 2019, the Committee of Ministers in the Council of Europe adopted a declaration on the manipulative capabilities of algorithmic processes[124], calling for more protection in respecting human rights, given 'significant impacts of the targeted use of data on societies and on the exercise of human rights more broadly'.

In 2019, the report of the UN Secretary-General's High-Level Panel on Digital Cooperation's put forward proposals to improve 'global digital cooperation architecture', providing some momentum for developments at UN level. On 11 June 2020, United Nations Secretary-General António Guterres presented a set of recommended actions for the international community to promote – inter alia – digital trust and ensure the protection of human rights in the digital era[125]. These included the creation of digital public goods to counter xenophobia and disinformation as well as advocacy for Digital Human Rights that helps in dealing with the spread of disinformation.

During 2020, in the wake of disinformation campaigns surrounding the COVID-19 pandemic, the United Nations launched the *Verified* initiative[126] to counter the spread of misleading information on the public health crisis. This is a public-private partnership relying on individuals as trusted community messengers. The UN and the WHO have also partnered with Facebook, WhatsApp and other messaging services to deliver accurate information about the pandemic (see also chapter 4 above).

In January 2020, the Parliamentary Assembly of the Council of Europe (PACE) expressed concern 'about the scale of information pollution in a digitally connected and increasingly polarised world, the spread of disinformation campaigns aimed at shaping public opinion, trends of foreign electoral interference and manipulation, as well as abusive behaviour and the intensification of hate speech on the internet and social media'[127].

5.2 Corporate activities

Since the UN Human Rights Council in 2011 endorsed by consensus the *UN Guiding Principles on Business and Human Rights*, it has been widely accepted that companies have a non-binding 'responsibility to respect' human rights[128]. The largest digital information platforms are now increasing their human rights expertise and focus, particularly regarding freedom of expression, as they combat disinformation and foreign interference. They are working with fact-checkers and have introduced increased transparency for political advertisements. However, rules and procedures regrettably differ among platforms, leaving loopholes for cross-platform disinformation campaigns[129]. Platforms lack transparency on how they collect personal data, what they do with it or how their algorithms work, thus creating openings for micro-targeting and disinformation.

Most actions taken by the largest social media platforms are related to content curation, as illustrated below in Table 4. For instance, in 2019 Facebook employed 15 000 staff to deal with content moderation. By contrast, WhatsApp for example uses end-to-end encryption, which in practice means that it cannot access the contents of messages. There are also dedicated disinformation sites (e.g. *Infowars*, *Q-Anon*) that do not apply these types of restrictions. Although content curation has been the most used measure,

[124] Committee of Ministers, Council of Europe, Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes', Decl(13/02/2019)1, 1337th meeting of the Ministers' Deputies, February 2019.
[125] United Nations, Roadmap for Digital Cooperation, June 2020.
[126] See Verified webpage.
[127] PACE, Democracy Hacked? How to Respond?, Resolution 2326 of the Parliamentary Assembly of the Council of Europe on 31 January 2020 (9th Sitting), January 2020.
[128] UN OHCHR, Guiding Principles on Business and Human Rights, New York and Geneva: United Nations Office of the United Nations High Commissioner for Human Rights, 2011.
[129] Karen Kornbluh and Ellen P. Goodman, Safeguarding Digital Democracy. Digital Innovation and Democracy Initiative Roadmap, *The German Marshall Fund of the United States DIDI Roadmap n 4*, March 2020.

digital platforms have also adopted other precautions to fight disinformation, such as working with fact-checkers.

Table 4. Actions taken by online platforms

Facebook	• Rules for political and issue advertisements: any advertiser who wants to run political or issue ads must be verified on the platform and include 'paid for' disclaimers with the advertising[130]. • Applying machine-learning to assist their response teams in detecting fraud and enforcing our policies against inauthentic spam accounts[131]. • Media literacy campaign launched with fact-checkers *FullFact*[132].
Twitter	• Labelling or removing false or misleading information intended to undermine public confidence in an election or other civic process (misleading information, disputed claim, unverified claim)[133].
YouTube	• Removing content that has been technically manipulated or doctored in a way that misleads users (beyond clips taken out of context) and may pose a serious risk of egregious harm[134]. • Terminate channels that attempt to impersonate another person or channel, or artificially increase the number of views, likes and comments on a video[135].
Instagram	• Labelling content that has been rated as false or partly false by a third-party fact-checker[136]. • If something is flagged as false or partly false on Facebook, it is label identical if it is posted on Instagram (and vice versa).
WhatsApp	• Limits on message forwarding[137].

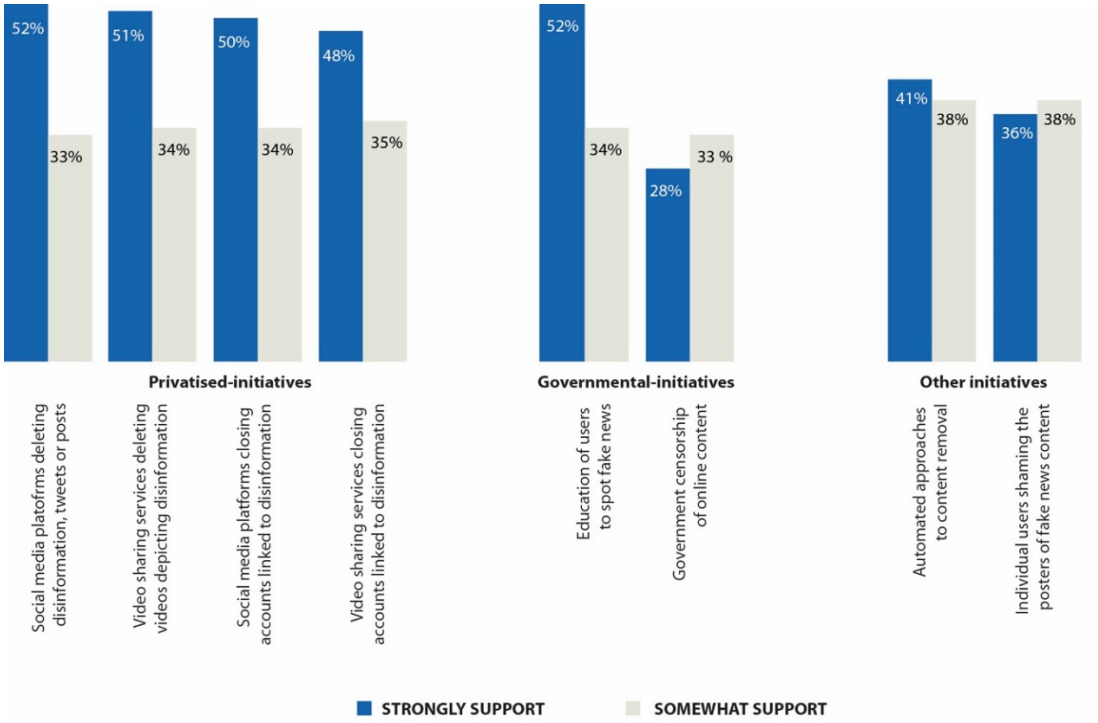
Ahead of the 2020 US elections, LinkedIn, Pinterest, Reddit, Verizon Media and the Wikimedia Foundation joined Google, Facebook, Twitter and Microsoft to coordinate with the US intelligence community in identifying disinformation campaigns. This led to several 'takedowns' of coordinated inauthentic behaviour, including the removal of a network linked to the Russian troll farm Internet Research Agency (IRA) from Facebook[138].

[130] See: https://wearesculpt.com/blog/political-ads-on-facebook/
[131] See: Facebook's policies on Facebook news webpage.
[132] FullFact, Report on the Facebook Third-Party Fact-Checking programme, December 2020.
[133] See Twitter's policies on Twitter webpage.
[134] See Youtube's policies on Youtube Help webpage
[135] See Youtube Help webpage
[136] See Instagram's policies on Instagram's webpage.
[137] Alex, Hern, WhatsApp to impose new limit on forwarding to fight fake news, *The Guardian*, published on 7 April 2020.
[138] Nimmo, Ben, Camille François, C Shawn Eib, Léa Ronzaud, IRA Again: Unlucky Thirteen. GRAPHIKA Report, September 2020.

The decision[139] of some of these digital platforms to silence President Donald Trump online presence temporarily – and ultimately permanently in the case of Twitter – following the storming of the US Capitol on 6 January 2021, raised new questions about limitations to freedom of expression and the role of social networks as gatekeepers of disinformation.

However – as Figure 3 below shows –, according to an Ipsos poll private sector initiatives to counter disinformation garner support from the majority of citizens.

Figure 3. Citizens' support for different measures to tackle 'fake news' and disinformation[140]



Source: Authors' own elaboration based on CIGI-Ipsos, 2019 CIGI-Ipsos Global Survey on Internet Security and Trust., 2019; CIGI-Ipsos, Internet security and Trust. Part 3, 2019.

With regards to promoting and protecting the right to freedom of opinion and expression, the UN Special Rapporteur has repeatedly called for alignment of digital platforms' policies on content moderation with freedom of expression standards[141], given that transparency and accountability in relation to curatorial actions are essential for protecting this right.

Digitalisation has brought about a profound redistribution of power, but this transformation has not been accompanied by equally robust mechanisms for documenting responsibilities. For a long time, digital platforms – along with other private entities – have been setting their own standards. One of the main

[139] Sonnemaker, Tyler, 'Facebook and Instagram have blocked Trump for 24 hours after the president published posts spouting misinformation as his supporters violently stormed the US Capitol', *Insider*, published on 7 January 2021. [Accessed 20 February 2021]:

[140] The percentage refers to people interviewed in 25 different countries (Australia, Brazil, Canada, China, Egypt, France, Germany, Great Britain, Hong Kong (China), India, Indonesia, Italy, Japan, Kenya, Mexico, Nigeria, Pakistan, Poland, Russia, South Africa, Republic of Korea, Sweden, Tunisia, Turkey and the United States) accounting for 1.000+ individuals in each country during December 2018 and February 2019. See the original source (IPSOS CIGI) for more information about the methodology.

[141] UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Promotion and protection of human rights: human rights questions, including alternative approaches for improving the effective enjoyment of human rights and fundamental freedoms, UN Doc A/73/348, August 2018.

problems in holding them accountable for the spread of disinformation is determining who is responsible for the content posted.

The proposed EU Digital Services Act (see Chapter 6) tries to address this issue by requiring digital platforms to address systemic risks by publishing annual risk assessments. Digital platforms will have to implement measures to mitigate the identified risks, which could imply those associated with illegal content, impact on human rights such as freedom of expression along with any kind of malicious interference involving for instance public health issues and elections[142].

5.3 Civil Society

Civil society is at the heart of identification responses (monitoring, fact-checking and investigative journalism). This dimension is just one layer of rights-based defences against disinformation, which dovetails with human rights-oriented approaches to the problem. This approach seeks to instil greater empowerment within the users or target audiences of disinformation to heighten their resistance to its pernicious effects. As Chapter 7 will show, good-quality journalism and media pluralism are key in building societal resilience to disinformation.

6 EU responses to disinformation

'What convinces masses are not facts, and not even invented facts, but only the consistency of the system of which they are presumably part.'

Hannah Arendt in *The Origins of Totalitarianism*

Key takeaways

- In recent years, the European Union – along with many democratic countries and international organisations – has made progress in building its toolbox to orchestrate a coordinated defence of democratic values against the threat of disinformation.
- With the European Democracy Action Plan and the Digital Services Act, the EU upgrades its regulation of internet platforms and its commitment to tackling disinformation.
- The balance between regulation and freedom of expression raises difficult and sensitive questions that need full and open political debate in the EP plenary. The EP needs not only to be the advocate for global human rights but also the forum where these complex issues are tackled.
- The Action Plan on Human Rights and Democracy 2020-24 gives the EP a reference point as well as a platform from which to exert stronger influence over the EU's commitments to human rights in the fight against disinformation.

The European Union has gradually put in place a wide-ranging toolbox for countering disinformation. This toolbox contains many elements, internal and external, that overlap with one another. In the context of

[142] James R. Carroll Brian W. Duwe David C. Eisman Patrick Fitzgerald Todd E. Freed Marc S. Gerber Richard J. Grossman Michael E. Leiter Stuart D. Levi William Ridgway Jason D. Russell David E. Schwartz Ingrid Vandenborre Helena J. Derbyshire Jessica N. Cohen Peter Luneau Jamie S. Talbot Eve-Christie Vermynck, Privacy & Cybersecurity Update, *Skadden*, January 2021.

this study, we are concerned with those elements which not only speak most directly to the human rights dimensions of EU policies, but also could be relevant for addressing disinformation challenges abroad. The protection and the promotion of human rights combine to form one of the general principles of European law[143].

In its Communication on Tackling Online Disinformation: a European Approach[144], the European Commission acknowledges that democratic societies, or societies that aspire to live with the same or similar values as enjoyed in the European Union, 'depend on the ability of citizens to access a variety of verifiable information so that they can form a view on different political issues. In this way, citizens can participate in public debates from an informed position and express their will through free and fair political processes'.

The European Council's Strategic Agenda 2019-2024[145] includes disinformation under a chapter on 'Protecting citizens and freedoms' as it is deemed a menace to citizens' fundamental rights and freedoms, as well as democratic values and the rule of law. It includes disinformation alongside cybercrime, on the assumption that these two phenomena come from hostile state and non-state actors. The Council stipulates that efforts to protect democratic institutions from hybrid threats (like disinformation) must respect fundamental rights such as the freedom of expression.

The EU has developed a varied toolkit, intensifying its efforts particularly since March 2015 when the European Council 'stressed the need to challenge Russia's ongoing disinformation campaigns', following the annexation of Crimea and destabilisation of eastern Ukraine[146]. Figure 4 below presents a general overview of actions and initiatives take at EU level, clearly showing the acceleration and intensification of EU responses to the threat in recent years.

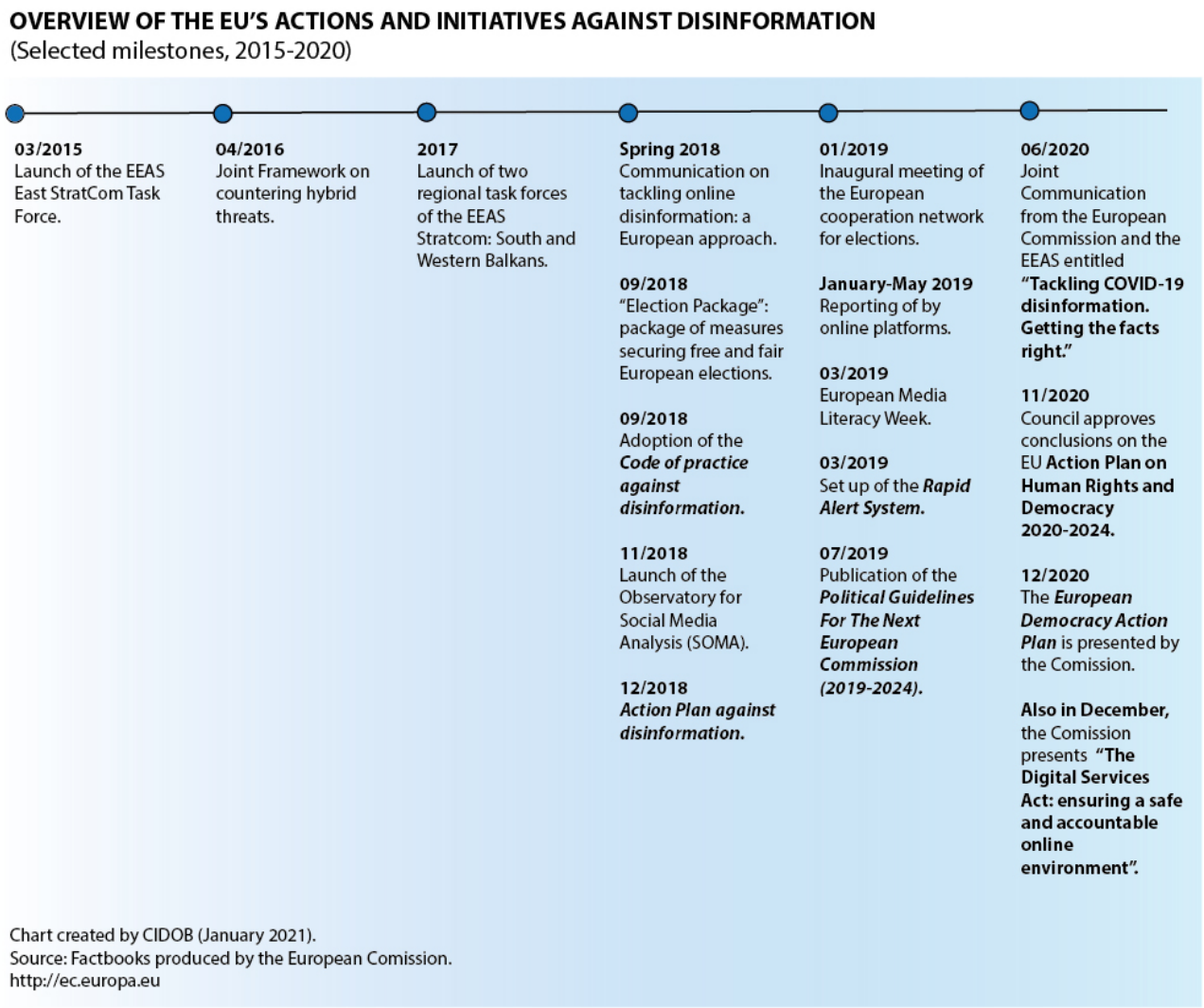
[143] TEU (art 2) states that: 'The Union is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights, including the rights of persons belonging to minorities. These values are common to the Member States in a society in which pluralism, non-discrimination, tolerance, justice, solidarity and equality between women and men prevail.'

[144] European Commission, 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, April 2018.

[145] European Council, A new strategic agenda for the EU 2019-2024, June 2020.

[146] European Council, European Council Conclusions on external relations (19 March 2015), March 2015.

Figure 4. EU's Actions and initiatives against disinformation



Source: Chart elaborated by CIDOB in January 2021 based on Factbooks produced by the European Commission.

6.1 The EU's policy framework and instruments focusing on disinformation and European democracy

6.1.1 The EEAS Strategic Communication Division

In March 2015, the European External Action Service (EEAS) was tasked with countering disinformation. The EEAS lies at the heart of the EU's external strategy related to foreign disinformation campaigns and is responsible for implementing the Action Plan Against Disinformation and the Rapid Alert System (see further below). The Task Forces focussing on the eastern and southern neighbourhoods and the Western Balkans are the main units dealing with proactive communication activities to counter disinformation.

The East StratCom Task Force unit (ESCTF) was initially formed as the core of EU efforts against a disinformation challenge directly linked with Russia's efforts to destabilise EU electoral processes and political debate. Under its activities, monitoring, identifying and debunking have become important aspects of the EU's strategy in countering disinformation. However, newer StratCom Units working in the Western Balkans and Southern neighbourhood have also broadened their scope of work, as they look beyond merely identifying disinformation to strengthening the overall resilience of societies against this threats.

At the same time, the EEAS has been dealing with other security aspects linked to these unconventional challenges to the EU's political and societal resilience. For instance, an EU Hybrid Fusion Cell within the EU Intelligence and Situation Centre (EU INTCEN) of the EEAS has been focussing on analysing external aspects of hybrid threats such as massive information campaigns, the recruitment of radicals or the use of proxy actors to conduct certain acts.

All these tools have placed the EEAS in a central coordinating position within the EU's strategy against disinformation, combining monitoring, analysis, public diplomacy and strategic communications and involving 140 EU delegations and offices around the world.

6.1.2 The Rapid Alert System

In preparation for the European elections during May 2019, the EU established a Rapid Alert System (RAS), implemented in the EEAS to coordinate with Member States and aiming to:

- facilitate information sharing;
- expose disinformation in real time;
- and coordinate with other multilateral efforts by the G-7 Rapid Response Mechanism and NATO.

Despite the deployment of these tools, isolated cyber-attacks, data protection and other elections-related complaints were still reported[147].

The acceleration of COVID-19 *infodemics* has prompted the EU to upgrade its toolkit and, crucially, more tightly link its strategy against disinformation to new human rights and democracy commitments.

6.1.3 The Action Plan Against Disinformation and the Code of Practice on Disinformation

In Autumn 2018, the European Commission developed an Action Plan Against Disinformation and concluded agreements for a Code of Practice on Disinformation with major social media companies. The Code was 'an experiment in voluntary self-regulation by the tech industry'[148]. Some civil society organisations have criticised the Code for theoretically allowing – and even incentivising – restrictions on the freedom of speech that are claimed to be technically lawful[149]. The European Commission complained about the initial compliance reports produced by these platforms, with former Commissioner Julian King describing them as 'patchy, opaque and self-selecting'[150].

These initiatives for tackling disinformation did not come with precise impact measurement indicators. They did not carry quantitative goals to prove how EU actions have helped to combat disinformation. Hence, the European Court of Auditors' decision to probe into the impact of the 2018 Action Plan against Disinformation reflected concerns over its efficacy[151]. The European Democracy Action Plan (presented in Section 6.1.5) offers the opportunity to strengthen the Code of Practice.

[147] European Commission, Commission reports on 2019 European elections: fostering European debates and securing free and fair elections, June 2020.
[148] James Pamment, The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework, *Carnegie Endowment for International Peace Future Threats Future Solutions series*, n 2, September 2020.
[149] Aleksandra Kuczerawy, Fighting Online Disinformation: Did the EU Code of Practice Forget about Freedom of Expression?, in Georgios Terzis, Dariusz Kloza, Elzbieta Kuzelewska and Daniel Trottier (eds.), 'Disinformation and Digital Media as a Challenge for Democracy', *European Integration and Democracy Series*, Vol. 6, June 2020.
[150] Stolton, Samuel 'EU Commission hits out at Facebook's disinformation report'. *EurActiv*, Published on 20 January 2019. [Accessed 07 January 2021]
[151] European Court of Auditors, Auditors look into the EU's fight against disinformation, ECA Press Release of 17 March 2020.

6.1.4 The European Digital Media Observatory

Strengthening journalism and supporting fact-checking to counter disinformation have been prioritised in building resilience inside the EU. Following the 2018 Communication on Tackling online disinformation: a European approach[152], the Social Observatory for Disinformation and Social Media Analysis (SOMA) was launched with support from the European Commission. It aimed to connect researchers, fact-checkers and media organisations. Building on that idea of a multidisciplinary approach and on a governance structure independent from public authorities, the European Digital Media Observatory (EDMO)[153] opened in June 2020 to facilitate closer coordination amongst fact-checking organisations, the scientific community, media practitioners and teachers with technological platforms and public authorities. The EDMO will offer new funding for targeted research on tackling disinformation and, although aimed at strengthening the EU media ecosystem, it could also set a model for building robust regional journalistic networks internationally.

6.1.5 The European Democracy Action Plan and the Digital Services Act

At the end of 2020 the Commission presented the European Democracy Action Plan (EDAP) together with its proposal for an updated e-commerce directive, the Digital Services Act (DSA). Both initiatives take an expansive view of digital regulatory policy by proposing to introduce legally binding tools, especially with regards to the accountability and transparency of digital platforms. These measures seek to enhance the EU's democratic resilience and regulatory toolbox. On the one hand, the EDAP pledges to revamp the Code of Practice on disinformation and reinforce the EU policy framework more broadly. On the other, the DSA's promise to develop 'systemic rules for the online ecosystem'[154] could offer a template for global digital governance – as sought by the European Parliament[155]. Taken together, these two initiatives offer the prospect of a more ambitious European Commission effort to protect fundamental rights.

6.2 Key elements of the EU's external Human Rights and Democracy Toolbox

6.2.1 EU human rights guidelines

In May 2014 the Foreign Affairs Council adopted the EU Human Rights Guidelines on Freedom of Expression Online and Offline, one of the earliest and most significant element of its human rights toolbox. It provided officials and staff with practical guidance on how to contribute in preventing potential violations affecting freedom of opinion and expression. The document establishes six priority action areas:

- Combating violence, persecution, harassment and intimidation of individuals, including journalists and other media actors exercising their right to freedom of expression online and offline, as well as combating impunity for such crimes;
- Promoting laws and practices that protect freedom of opinion and expression;
- Promoting media freedom and pluralism together with fostering an understanding among public authorities of dangers from unwarranted interference with impartial/critical reporting;

[152] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Tackling Online Disinformation: a European Approach, COM(2018) 236 final, April 2018.
[153] See EDMO webpage.
[154] Access Now, DSA: European Commission delivers on the first step toward systemic rules for online platforms, Published on 15 December 2020 [Accessed 07 January 2021]
[155] European Parliament, Digital: The EU must set the standards for regulating online platforms, say MEPs, EP Press release of 20 October 2020.

- Promoting and respecting human rights in cyberspace as well as other information and communication technologies;
- Promoting best practices by companies;
- Promoting legal amendments and practices aimed at strengthening data protection and privacy online/offline;

These guidelines were at the core of specific programmes such as *Media4Democracy*, a project aimed at enhancing EU Delegations' capacity for strategic advocacy, media sector engagement as well as promoting freedom of expression and media pluralism around the world[156]. *Media4Democracy* was established by the European Commission's Directorate-General for International Cooperation and Development (DEVCO) and was backed by a consortium of civil organisations working on media development and supporting freedom of expression such as Article19, Deutsche Welle Akademie, European Partnership for Democracy, Free Press Unlimited and the Thomson Foundation. This constitutes one clear example of EU mechanisms and instruments being used to address the phenomenon of information disorder.

6.2.2 EU engagement with civil society and human rights dialogues

During 2014, the EU also started elaborating Roadmaps for engagement with civil society in external relations so as to promote a meaningful participation of CSOs in the domestic policies of partner countries[157]. The EU Roadmaps for engagement with CSOs, introduced in 2012 by the Commission Communication on the roots of democracy and sustainable development, were meant to ensure structured dialogue and strategic cooperation with civil society and international actors, thereby increasing the consistency and impact of EU actions. These roadmaps could also play an important role in supporting civil society responses to disinformation.

Shrinking democratic space for civil society has led the European Union to strengthen policies and instruments against new challenges to democratic processes inside and outside the EU. Another important initiative in this context has been the Human Rights Defenders Protection Mechanism – known as *ProtectDefenders.eu* – , which was established in 2010 to support human rights defenders facing imminent threats.[158]

Civil society organisations are also demanding a more meaningful role in the Human Rights Dialogues that, since December 2001, form one of the EU's non-coercive foreign policy tools to promote human rights policies in third countries or with other regional organisations[159]. The EU Action Plan for Human Rights and Democracy (see section 6.2.4 below) incorporates the priority of reinforcing the political, human rights and sectoral policy dialogues as one of the effective tools for this plan's implementation.

6.2.3 Election observation and democracy support

In October 2019, the European Council adopted the 'Council Conclusions on Democracy' with its commitment to increase EU efforts in democracy-building capacities within third countries. This includes the promotion of inclusive and credible electoral processes through EU election observation and support

[156] See Media4Democracy webpage.
[157] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on The roots of democracy and sustainable development: Europe's engagement with Civil Society in external relations, COM(2012) 492 final, September 2012.
[158] See ProtectDefenders.eu
[159] The Human Rights and Democracy Network recommendations for the revision of the EU guidelines on human rights dialogues with third countries, announced in December 2020, regret that CSOs are very rarely allowed to participate directly in the dialogues and demand more ambition and a stronger communication.

to domestic electoral observers[160]. In this context, the EU's elections observation missions have developed a methodology to monitor online political campaigns with the aim of identifying and tackling the challenging effects of disinformation, manipulation and hate speech in their undermining of democratic processes. This methodology has already been tested by electoral missions in Peru, Sri Lanka and Tunisia. The Action Plan for Human Rights and Democracy also supports 'the development of policy frameworks that apply offline rules on elections and democratic processes to the online context, and assist to build capacities to implement them'[161]. Election observation is the starting point for further cooperation with third countries, also at parliamentary level.

6.2.4 The Action Plan for Human Rights and Democracy for 2020-2024 and funding tools

Adopted by the Council at the end of November 2020, the Action Plan for Human Rights and Democracy lists among its priorities: 'Promoting efforts to counter disinformation, hate speech, extremist and terrorist content, including online media literacy and digital skills; Supporting independent fact-checking and research, investigative reporting and quality journalism, including at local level'. This 2020-2024 iteration of the Action Plan stresses these digital issues to a much greater extent than was evident in previous versions covering the 2010s. The Action Plan has a very political ambition to enhance EU leadership in promoting and protecting human rights as well as democracy worldwide and to improve coherence, unity and efficiency not only between member states but also in all areas of EU external action. The plan is also important as an umbrella strategy that nominally guides EU funding and other decisions on the ground in countries around the world. The priority now attached to disinformation opens up the prospect of higher levels of funding for this issue in the EU's external action over the years ahead.

In line with this Action Plan's remit, the European Instrument for Democracy and Human Rights (EIDHR) is the EU's flagship funding instrument dedicated to democracy and human right issues. Under the 2014-2020 multiannual financial framework, the EIDHR enjoyed an annual budget of around EUR 160 million, higher than the previous budget period. Unlike parts of the toolbox outlined above, the EIDHR is not concerned specifically with disinformation. Yet, it has increasingly funded initiatives in third countries related to digital concerns, including disinformation. The EIDHR's approach is somewhat indirect in this sense: it funds civil society organisations seeking to tackle local disinformation problems. In recent years this has comprised a range of media literacy initiatives, alongside training and capacity-buildings for CSOs on digital protection and rights.

Geographically defined aid instruments have offered some supplementary funding for similar programmes. One example of this comes from the European Neighbourhood Instrument's funds for counter-disinformation in Eastern Partnership countries. This work is focused on the analysis and exposure of state-driven pro-Kremlin disinformation narratives. Funded initiatives focus on strengthening resilience and increasing awareness within society. The EU supports media outlets and projects aimed at increasing media literacy as well as enhancing citizen-led journalism. It also funds civil society actors and fact-checking initiatives to expose online disinformation stories. Local fact-checking organisations in states like Georgia and Ukraine contribute to the *EUvsDisinfo* platform and public database, which produces and stores a weekly disinformation newsletter (*Disinforeview*) summarising major Pro-Kremlin trends. This platform allows policymakers working on disinformation to improve the tailoring of their strategies and action plans to counter disinformation both EU wide and in neighbouring countries. The EU also works

[160] Council of the European Union, Democracy: EU adopts conclusions, Published on 14 October 14 2019. [Accessed on 07 January 2021]
[161] European Commission, Joint communication to the European Parliament and the Council, 'Action Plan on Human Rights and Democracy 2020-2024', March 2020.

with state officials to improve their strategic communication skills in counteracting the push-back against disinformation.

These trends are highly relevant within a human rights-centred approach. While Stratcom entails the EU itself rebutting disinformation, the EIDHR and other instruments fund local rights-oriented groups to build their own capacities to push back against digital control and manipulation as part of their human rights work. At the time of writing, the precise structure and details of democracy and human rights funding for the 2020-2027 period are still being finalised. However, it seems that the level of funding allotted to countering disinformation will increase modestly. A new catch-all Neighbourhood, Development and International Cooperation Instrument (NDICI) will act as an umbrella for thematic programmes, including one for democracy and human rights.

6.2.5 Restrictive measures

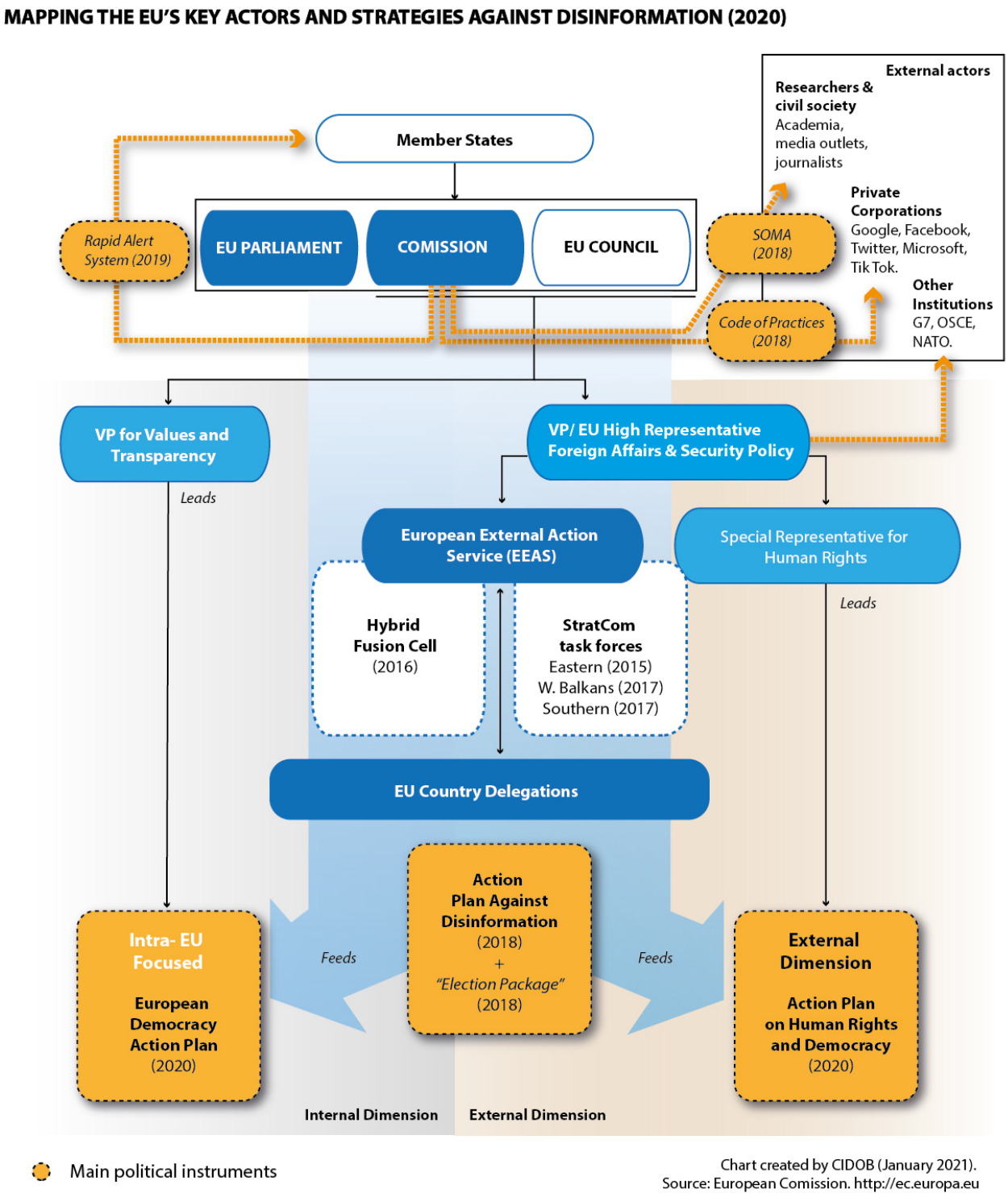
The EU can to use sanctions and conditionality in relation to human rights and democracy. These are not specific to disinformation, but are a crucial part of the Union's external toolbox for human rights concerns. All EU agreements with third countries have since long included essential element clauses for human rights and democracy. The EU has in recent years taken steps to impose restrictive measures against individuals from countries such as Belarus, Iran, Myanmar, Venezuela and Zimbabwe involved in human rights abuses. It has also withheld some aid and trade preferences to countries suffering democratic backsliding and human rights problems. These are generally not measures imposed against regime disinformation as such, but they have targeted human-rights abuses carried out by some of the regimes most heavily implicated in the use of disinformation.

In December 2020 the EU adopted its long-awaited Global Human Rights sanctions regime and applied measures under this framework in several instances in early 2021. This will make it easier for the EU to impose sanctions on individuals deemed to be guilty of serious human rights abuses, separately from the Union's country strategies. While disinformation may not be a direct, explicit focus of this regime, the EU may be able to more readily target government officials where they abuse the risk of disinformation to suppress media freedom. The sanctions regime includes freedom of expression as one of the rights whose breach will justify restrictive measures. This is one of the rights most threatened by disinformation (see Chapter 3). Accordingly, the EDAP already mentions the EU's need for further development of tools which will impose 'costs on the perpetrators' of foreign interference and influence operations[162].

The Commission has floated the possibility of developing a sanctions regime specifically for disinformation as part of steps to take forward the European Democracy Action Plan, although it is not yet clear whether this would still apply only to disinformation's impact on EU citizens or also cover human rights infringements within third countries[163].

[162] Council of the European Union, Cyber attacks: EU ready to respond with a range of measures, including sanctions, Press release of 19 June 2017.
[163] Alexandra Brzozowski, Commission floats sanctions regime for disinformation offenders, *EurActiv*, Published on 3 December 2020.

Figure 5. Key actors and strategies against disinformation



6.3 The European Parliament's role

The balance between regulation and freedom of expression raises difficult and sensitive questions that call for full and open political debate in the EP plenary. Accordingly, the European Parliament must not only continue to advocate global human rights when it comes to tackling disinformation, but also provide the assembly within which these complex issues can be tackled. Indeed, the EU Action Plan on Human Rights and Democracy 2020-24 highlights the European Parliament's key role in supporting human rights, which includes fighting disinformation. During interviews conducted specifically for this study, MEPs indicated

their full awareness of the EP's role as a speaker for human rights defenders. Additionally, these interviews revealed common political ground between different ideological perspectives in tackling disinformation through a human rights lens. MEPs need to continue collaborating in monitoring the precise follow-through and results from activities involving various parts of the EU toolbox, as described in Section 6.2.

The EP also has a role in working with other parliaments around the world to strengthen global standards. It could take the lead together with parliaments from like-minded countries such as Australia, Japan and Canada to push for a UN Convention on Universal Digital (Human) Rights[164,165]. To move matters forward, the European Parliament will have a role in cooperating with the Council of Europe, the body which houses much of the requisite expertise, along with civil society organisations that have already attempted to work up drafts of such a new document. Hence, parliamentary delegations to third countries and the EU delegation to the UN should form the platforms from which actions to counter disinformation can be initiated.

The European Parliament has put in place a special committee on 'Foreign Interference in all Democratic Processes in the European Union including Disinformation' (INGE). Through its regular contact with other Parliaments around the world, the EP is able to exchange ideas on best practices with third countries regarding parliamentary processes in this field. For example, the Inter-Parliamentary Union has held initial discussions on the issue of disinformation and is forming its own conclusions[166]. Others have formally debated the increasing threat posed by disinformation and the British parliament[167], for instance, has created the International Grand Committee. This is the first of its kind to promote further cross-border co-operation in tackling the spread of disinformation. The EP is a natural partner for these other chambers which seek to approach disinformation from the perspectives of democracy and human rights.

In our interviews, MEPs stressed how important it is for the EP to develop a more proactive role in the Action Plan on Human Rights and Democracy 2020-2024. They also highlighted the value of inter-parliamentary delegations in facilitating exchanges of information between MEPs and their peers. The input from inter-parliamentary delegations allows MEPs to play a critical role in giving voice to global citizens' rights. Indeed, the Action Plan emphasises the importance of support to parliamentary institutions. This gives the EP a reference point and platform from which to exert stronger influence over the EU's external toolbox and ensure that this gives adequate protection inter alia to human rights in the fight against disinformation.

The EP can also advocate funding increases for projects aimed at counteracting disinformation. Interviews with MEPs have revealed that they are already aware of civil society organisations' need for more resources if they are to prevail in the fight against this problematic and worrying issue.

[164] The IO Foundation defines Digital Rights as an extension of Human Rights in the digital space; both the public spaces on the Internet and private networks.
[165] See the result of the workshop Future-proofing our digital rights; and Article 19's 'Universal Declaration of Digital Rights', Internet Rights & Principle's Coalition's 'Charter of Human Rights and Principles for the Internet'; or ZEIT-Stiftung's 'Charter of Fundamental Digital Rights of the European Union' which was presented to the European Parliament in 2016
[166] Inter-Parliamentary Union,'Key conclusions from the April 2019 expert hearing on disinformation and 'fake news'', *Innovation tracker*, Issue 2, June 2019.
[167] UK Parliament Digital, Culture, Media and Sport Committee , Disinformation and 'fake news': Final Report, *House of Commons*, Eighth Report of Session 2017–19, 2019.

7 Rights-based initiatives against disinformation: identifying best practices

'Whenever there is repression, there is resistance'

Hong Kong activist Nathan Law[168]

Key takeaways

- EU external policies can helpfully learn from practices emerging around the world that tackle disinformation through a human rights lens.
- Such 'best practices' offer useful policy ideas and pointers, but far greater support is needed if they are to be extended and strengthened.
- The EU needs to offer such support for the different levels that this report has identified as crucial to holistic policy approaches, both within governments and civil society.
- While policies aimed at governmental and regulatory measures are important, it is vital that more bottom-up approaches are not overlooked, as the development of civil capacity to deal with disinformation is particularly crucial for the EU to support human rights.

As the scale of disinformation has increased in recent years, so have efforts to tackle it. There are many different practices across the world that the EU can draw from in facing the challenge to counter disinformation in ways that are consistent with and indeed further human rights. Whilst many emerging approaches to disinformation do not sit easily with human rights commitments, others have already begun to map strategies that fuse together disinformation and human rights concerns.

This is now a vast field as policy responses and new initiatives have increased exponentially over recent years. We do not aim here to offer any kind of exhaustive account of best practices, but rather concentrate on a select number of lessons that might be of specific relevance to EU external support for human rights and democracy in line with its counter disinformation aims. There are several practices emerging around the world that can and should be included within the range of EU policy instruments outlined in previous sections of this report. In line with the distinction drawn between different actors referred to in Chapter 5, these practices are evident among both governments and civil society bodies.

7.1 Government and parliamentary responses

At the level of governments and public authorities, a number of dimensions from emerging practices are relevant for human rights considerations[169]. In some countries, governments and state authorities have begun to monitor the human rights impacts of their new legal restrictions against disinformation. An increasing number of states have also introduced measures specifically around the time of elections, when countering disinformation is integrally linked to the protection of core democratic rights. There are best practice examples of electoral commissions working to increase transparency of information in election campaigns. Furthermore, broader government responses have targeted the mechanisms of disinformation with the aim of giving citizens more effective rights to free choice and information. Authorities in many

[168] Marc Perelman, Hong Kong activist Nathan Law: 'Whenever there is repression, there is resistance', France24, Published on 28 July 2021.
[169] Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.

countries have worked hard on: pressing platforms to introduce more effective and transparent algorithms; curating practices that improve editorial processes and online community standards; and demonetisation rules that stop disinformation being profitable.

A number of countries have already introduced policies that appear to acknowledge human rights considerations within their strategies to counter disinformation. A selection of these can be mentioned, chosen on the basis that they have nested counter-disinformation efforts within a focus on core democratic processes, rights protection and citizens' engagement.

Taiwan is often cited as offering best practice to the extent that its public authorities have developed a digital strategy which reflects the concerns and criticisms of rights-oriented bodies. The country is considered to be a leader in developing digital policies predicated on societal participation and democratic legitimacy; its best practice resides in the way that formal government commitments are joined together with civil society input[170]. This core approach has been successful during the coronavirus pandemic, when the health emergency has further deepened the government's commitment to rights-led and participative approaches to disinformation[171]. The underlying rationale is that of a collective societal involvement in pushing back against disinformation[172].

The Canadian government has developed a range of good practices, including its Digital Citizen Initiative. Canada offers a best practice example of a government obliging tech companies to increase transparency in political content, particularly around the time of elections. Its Critical Election Incident Public Protocol[173], established for the 2019 general election, is a strong example of effective action. Australia has developed a similar Electoral Integrity Assurance Taskforce[174].

Other examples where governments and parliaments have been highly attentive to freedom of speech concerns include Japan, New Zealand and South Korea. In the latter country, legal rights protection has had an impact, as courts have overturned many cyber libel cases on freedom of speech grounds. Counter-disinformation has been approached increasingly through a narrative of defending democratic rights in the wake of President Park Gun-Hye's impeachment and the election of her replacement. Supported by a range of public authorities and multiple coordination efforts, fact-checking around South Korea's elections provides an example of good practice in tightly linking counter-disinformation to the defence of core democratic rights and process[175].

In Eastern Europe, Ukraine presents an example of significant policy shift. The Ukrainian government was widely criticised for an overly draconian approach to Russian disinformation when it set up a Ministry for Information Policy in 2014. More recently it has focused more on empowering citizens to recognise disinformation and be more discerning generally in their consumption of apparently genuine information[176].

Many examples from Latin America offer potential good practice lessons. For instance, the Argentinian government has developed strong provisions against misleading political advertising[177]. The incumbent

[170] Rorry Daniels, Taiwan's unlikely path to public trust provides lessons for the US, *Brookings*, 2020.
[171] Kelsie Nabben, Hacking the pandemic: how Taiwan's digital democracy holds COVID-19 at bay, *The Conversation*, Published on 11 September 2020.
[172] Elizabeth Lange, How One Social Media App Is Beating Disinformation, *Foreign Policy*, published on 23 November 2020.
[173] Government of Canada, Cabinet Directive on the Critical Election Incident Public Protocol, 2019.
[174] See AEC Electoral Integrity Assurance Taskforce webpage.
[175] Lim Boyoung, What's behind South Korea's fact-checking boom? Tense politics, and the decline of investigative journalism, Poynter, published on 16 June 2017.
[176] Olga Robinson, Alistair Coleman and Shayan Sardarizadeh, A report of Anti-Disinformation Initiatives, Oxford Internet Institute, 2019.
[177] Ruth Levush, Government Responses to Disinformation on Social Media Platforms: Comparative Summary, Law Library of Congress, September 2019.

president there and other presidential candidates also cooperated on a campaign against disinformation in the 2019 elections and support a wider, multi-stakeholder coalition, *Reverso*, in developing a wider set of measures[178]. In Uruguay, political parties went a step further by signing a formal 'ethical pact', committing not to engage in disinformation during election campaigning[179].

Brazil's Superior Electoral Court ran a counter disinformation campaign to defend the integrity of elections in 2018, linking the issue specifically to democratic rights. Mexico's National Electoral Institute (INE) worked with tech companies on a similar initiative to shore-up democratic elections[180]. This was the first election management body in the world to have Memoranda of Understanding with Facebook, Twitter and Google, as well as engaging with organisations such as the Venice Commission. The Mexican case has been considered a good practice model, covering both disinformation and campaign finance regulation, by other countries such as Tunisia[181].

This short selection of cases is not meant to imply any overarching general conclusions about the direction of travel in government policies around the world. Nevertheless, the aforementioned examples could provide helpful inspiration and lessons for the EU's external disinformation and human rights policies.

Advanced democracies such as Canada, South Korea or Taiwan may not need direct EU support, but could certainly offer useful templates for the EU to build on and fund in other third countries. For countries where governments are broadly democratic and proactively committed to human rights, the best practice examples could usefully be built into EU-supported partnerships and programmes run with governments and state authorities.

In other countries where public authorities have tended to undercut more than protect human rights and democratic processes, the EU will instead need to exert critical diplomatic pressure to push governments to accept best practices. This applies probably to most countries where the EU has high-priority and challenging foreign-policy agendas. This focus will often entail the EU using its diplomatic leverage to persuade other states into meeting the standards of human rights best practice by building freedom of expression guarantees relevant to disinformation into their legislation. Of course, the EU does at the same time need to exercise caution as authoritarian regimes such as those in Cambodia, China, Russia and Thailand have already launched counter disinformation initiatives that clearly serve political interests.

The EU has made important steps forward in building counter-disinformation measures into its Election Observation Missions, closely resonating with the kind of best practice examples which combine democracy-protection and counter-disinformation. Developing this emerging area of work further would be extremely valuable in broadening the EU's disinformation policies into a more political direction. Where public authorities are themselves implicated in election-related disinformation – as is the case in a sizeable and growing number of countries around the world – this will require more robust EU responses to governments' manipulation of democratic processes. In deepening an already promising focus on election-related disinformation, the EU could gainfully draw upon the kinds of positive examples being demonstrated in Canada and some Latin American countries.

Building on many of these emergent best practices, governments are now involving parliaments in working against disinformation, thereby offering important templates and reference points for the future

[178] For more information about Reverso see here.
[179] UNDP, Partidos políticos uruguayos firmaron pacto ético contra la desinformación, Press release of 26 April 2019.
[180] Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.
[181] Beata Martin-Rozumiłowicz and Rasto Kužel, 'Social Media, Disinformation and Electoral Integrity', International Foundation for Electoral Systems, Working Paper, August 2019.

in the fight against disinformation. The parliamentary dimension will be very much central to a rights-based approach.

7.2 Civil society pathways

Notwithstanding the opportunities for formal government and parliamentary engagement, the EU needs to pay equal attention to supporting non-governmental actors in its external policies. Best practice from a human rights angle often benefits from a more bottom-up approach that focuses on strengthening civic capacity against disinformation.

However, such civil society dimensions cannot easily be employed as some kind of panacea. In highly repressive regimes it may be impossible to pursue effective civic initiatives. Furthermore, not all civil society is liberal, pro-democratic and supportive of human rights; indeed, research has shown that in recent years certain sections of civil society around the world have become part of the disinformation problem rather than its antidote[182]. Yet, this kind of support has been effective in many contexts as an element of EU democracy and human rights support; the EU itself has acknowledged the core role of civil society in all its key documents on democracy and human rights over many years. It is one important strand that should not be overlooked, especially as this kind of right-based empowerment is appropriate in a practical sense to the funding instruments and toolbox that the EU has at its disposal in many third countries.

This civic-centred approach certainly fits with UNESCO's call for responses that strengthen societal cohesion around a new social contract, within which citizens are able to exercise their rights in pushing back against disinformation[183]. It is an approach that is aligned with external EU human rights and support for democracy, so much so that it works on the assumption that a more effective exercise of democratic rights is one vital part of counter-disinformation strategies.

This civil society dimension is of particular importance in responding to this study's human rights remit, particularly in cases where governments themselves generate and deploy disinformation, typically as part of policies which are deliberately designed to limit democratic checks and balances. The UNESCO report suggests that this means many best practice regulations or voluntary codes cannot be used elsewhere in the world where fewer checks and balances exist[184].

As shown in Table 3 earlier (see Chapter 5), some civil society best practice responses are about revealing and monitoring disinformation, such as fact-checking. Others are about generating alternative, reliable sources of information. The logic here is that free and active journalism increases the exertion of a core democratic right in efforts against disinformation, while government bans on certain sites and publications undercut it. Finally, some practices are about education and media literacy, encouraging a more critical use of platforms from citizens.

It is important to stress that such local rights-building strategies go well beyond the strong tendency to focus on fact-checking within standard news outlets. There are now hundreds of initiatives run by large media organisations, international NGOs, international institutions and tech companies aimed at tracking and exposing disinformation, finding ways to flag this online, developing standards about sources and the like. These approaches will naturally continue to be at the forefront of efforts against disinformation.

However, our concern here is also broader, in that we need to examine how strategies against disinformation adopted by third countries can link more closely with local agendas on protecting human rights and strengthening democratic control. While locally-rooted fact-checking operations have an important role to play in this sense, a rights-led approach also requires a more pre-emptive *ethos*. The

[182] Richard Youngs, Civic Activism Unleashed: New Hope or False Dawn for Democracy?, Oxford: *Oxford University Press*, 2019.
[183] Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.
[184] K. Bontcheva and J. Posetti, 2020.

underlying rationale must be about how communities generate, consume and engage with information and not simply about correcting misleading information after-the-fact.

Over recent years, many examples of best practice initiatives have emerged and developed, seeking to adopt a rights-strengthening approach in tackling disinformation. While diverse, they entail a shared concern about building civic capacity, which in turn helps local communities gain stronger control and become more critically engaged with information flows. They can be defined as best practice in the fact that they seek to empower local communities in deliberating on the issue of disinformation and deciding on own priorities, while enhancing civic capacity to monitor and redress rights abuses in government digital strategies.

We offer here just a few illustrative civil society initiatives, selected on the basis of these criteria related to local civic engagement and monitoring capacities.

7.2.1 Middle East and North Africa

In Lebanon, *Megaphone* is a volunteer collective of local journalists that generates news narratives designed to reach young audiences affected by disinformation. It does not aim to teach what is good or bad information, but rather produces short-form news stories in ways that younger people can relate to, as an indirect way of providing them with stronger resilience against disinformation. Its basic approach is expressly to ensure an 'equality of rights' as a route into tackling information distortions[185].

Syrian Archive is a Syrian project that preserves and memorialises documentation of human rights violations and crimes for advocacy as well as accountability reasons. By doing so, the organisation fights disinformation around these conflict related issues that have profoundly damaged Syrian society. It is a project run by *Mnemonic*, a non-profit organisation dedicated to archiving disappearing digital material. They accept non-governmental funding and are fully independent. They have already reinstated 350 357 videos onto social media platforms[186]. *Verify Syria* is another ambitious and vital initiative that has been seeking to counter disinformation in the Syrian conflict through local engagement[187].

7.2.2 Asia

The *Digital Empowerment Foundation* (DEF) works across 500 villages of India, empowering communities with access to digital tools which can be used against disinformation that has become so widespread there in recent years. The organisation focuses on ensuring that community members have some understanding of accessing and evaluating information before it is consumed. DEF has a number of signature programmes, including on the digital empowerment of community organisations; the digital protection of CSOs; digital rural entrepreneurship; and a COVID-19 Digital Emergency Relief Programme[188].

7.2.3 Eastern Europe

In Eastern Europe, Russian and other external influence operations have prompted a wave of counter-disinformation civic initiatives. *FactCheck Georgia* founded in 2013 by Georgia's Reforms Associates (GRASS) runs capacity-building initiatives related to disinformation[189]. In Ukraine, *VoxCheck* was founded in 2015 by VoxUkraine with a similar approach[190]. *StopFake* was founded in 2014 by students and faculty members at the Kyiv Mohyla School of Journalism[191]. A notable feature in this region is how closely such initiatives are nested within broader CSO work to defend democracy and further human rights.

[185] France24, Megaphone, the independent news platform giving voice to Lebanon's uprising, Published on 16 November 2019.
[186] See Syrian Archive webpage.
[187] See Verify webpage.
[188] See Digital Empowerment Foundation webpage.
[189] See FactCheck Georgia webpage.
[190] See VoxCheck webpage.
[191] See StopFake webpage.

7.2.4 Latin America

Chicas Poderosas is a company started in 2013 and now operating in 13 Latin American countries. It aims at providing digital and media skills as well as leadership to women at community level. This company has created investigative journalism workshops and 'hackathons' to provide women with skills to facilitate their reporting on issues accurately, so they can hold to account people in positions of power. It oversees a New Ventures Lab that provides women with guidance and funding for entrepreneurial news or media ventures that seeks truth and information[192].

A Mexico-based company called *Animal Politico* runs an independent digital media site that is focused on grassroots investigative journalism with citizen-driven questions on political issues. It provides verifications, news and visualisations of issues, as a way of pushing back against disinformation from a rights perspective[193].

Comprova in Brazil and *Re-verso* in Argentina have been collaborative efforts between media companies and CSOs to fact-check around elections in these two countries[194].

7.2.5 Africa

The Centre for Innovation and Technology (CITE), is a project located in Zimbabwe that combines digital technology, fact-checking journalism and an emphasis on the promotion of social accountability at local government level. CITE produces reports, videos and podcasts testing claims made by public figures and institutions against hard evidence[195]. The logic is to generate pressure from local level against the effects of disinformation, something which has become increasingly pervasive in Zimbabwe, hindering democratic transition.

Pesa Check is a fact-checking indigenous project in East Africa that holds public figures accountable. It verifies financial and statistical numbers quoted by government leaders in Kenya, Tanzania and Uganda. It publishes clear articles on what information is true and what is false so that citizens are more informed on public information[196].

AfricaCheck is the main umbrella for fact-checking across Africa and has been functioning since 2012. It tests politicians' statements as requested by readers. The initiative has dozens of international funders[197].

7.2.6 Multi-regional projects

People in Need has a programme called *One World in Schools (*OWIS*)* that uses documentary films and associated activities to develop critical thinking skills as a foundation for developing citizenship skills. One important sphere of this OWIS programme is media literacy work. This is very much a grassroots, youth-oriented programme to counter disinformation, whilst at the same time providing the youth with critical media literacy skills[198].

Launched by Reporters Without Borders (RSF), *Tracker-19* is a tool whose name refers to both COVID-19 and Article 19 in the Universal Declaration of Human Rights, which protects freedom of expression and opinion. This project aims to evaluate the COVID-19 pandemic's impacts on journalism and especially the

[192] See Chicas Poderosas webpage.
[193] See Animal Politico webpage.
[194] Kalina Bontcheva and Julie Posetti (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, September 2020.
[195] See CITE webpage.
[196] See Pesa Check webpage.
[197] See AfricaCheck webpage.
[198] See One World in School.

way it has unleashed a wave of disinformation. The project documents state censorship and disinformation together with their impact on the right to reliable news and information. It makes recommendations on how to defend journalism. *Tracker-19* offers an interactive world map on the status of press freedom, constant coverage of developments and analyses of key issues. The tool explicitly combats disinformation through a focus on freedom of speech[199].

Meedan has created *Check*, a product that helps with the newsroom process of identifying misinformation and automates responses. It is a human and machine system that helps journalists focus on more high-level work instead of fact-checking. Since October 2019, *Meedan* has used *Check* for a global fact-checking project via WhatsApp in 5 countries: India, Brazil, South Africa, Kenya and Nigeria. The project has been supported by both WhatsApp and Facebook. *Check* can be used for elections, open-source investigations, translations and research/analysis on misinformation. The focus is on community-building and political engagement for journalists, students, CSOs, human rights defenders and others[200].

There are hundreds of similar initiatives currently in place across the world and it is beyond the scope of this report to offer a comprehensive survey of all such innovative approaches. However, it is important to stress that we are not concerned here with an assessment of every element of disinformation responses, but that the focus is on one very specific part of this field. Namely, the types of rights-oriented responses to disinformation that might be relevant to and built into the EU's external support programmes.

Several of the EU's policy documents and funding instruments outlined in Chapter 5 already refer to the key importance of active citizenship in the sphere of digital information. This is because all these efforts to approach disinformation through the promotion of more active citizenship give centre-stage to core citizen rights. They are about putting rights over information more clearly into citizens' hands – almost the opposite of legal or regulatory approaches that set limits to what is deemed acceptable information. The best practice examples mentioned here show the kinds of opportunities that exist for the EU actively to support such approaches. This is a current area of growth in global civil society, but it need higher levels of support to keep growing and maintain its incipient momentum.

On fact-checking, an important point to consider is that there has been a dramatic proliferation of large fact-checking initiatives in the US and European countries. The EU should be careful when it funds them to undertake fact-checking in other countries, as they can sometimes do so in a way that actually undercuts local actors' democratic empowerment.

General human rights work has evolved towards a focus on building local capacities and the EU has played a valuable role in advancing this trend. If this is true in the generic sphere of human rights, then it is also a lesson that needs to be applied more specifically in relation to the fight against disinformation's pernicious impacts. Better internet standards or regulatory controls over platforms are unlikely to provide far-reaching solutions in contexts where a wider respect for human rights and democratic norms remains weak and is declining, as is the case with many – if not most – of the EU's third-country partners.

In this sense, there are key lessons to be borne in mind and already-existing templates that can feed into a new EU effort to approach disinformation through a human rights lens. The EU can use these best practices to inform its human rights external aid programming. Moreover, beyond such funded projects, the EU will also need to use its full diplomatic weight and leverage to protect these third-country civil initiatives that routinely find themselves on the receiving end of government restrictions and harassment.

[199] See Tracker19 webpage
[200] See Meedan webpage.

8 Conclusions and recommendations

The EU must approach the concept of disinformation in a fully comprehensive manner in order to counter it effectively. This means accepting and adopting the ways in which this challenge has become more complex and multi-faceted over recent years. It also means tackling not only disinformation itself, but the many tactics of manipulation that accompany and amplify its pernicious effects. It entails targeting the multiple instigators and agents that drive disinformation strategies and unpacking the different motivations behind this phenomenon, whether they be political, financial or reputational. In its external relations, the EU needs to push back more systematically against both disinformation and a broader range of deceptive influence strategies.

The relationship between disinformation and human rights is double-edged. Disinformation infringes a range of core rights. These include: the freedom of thought; the right to privacy; the right to participation; as well as economic, social and cultural rights. It also diminishes broader indicators of democratic quality by: weakening trust in democracy; interfering with elections; as well as feeding digital violence and repression. However, counter-disinformation initiatives also carry risks for human rights and democracy. In many countries around the world, measures against disinformation have constricted human rights. The EU needs to support counter-disinformation efforts in its external relations, but at the same time be attentive to the ways in which these may cut across its human rights objectives.

The COVID-19 pandemic has intensified these trends and problems. It has triggered disinformation campaigns driven by political and profit-driven motivations. In the last year, a huge amount of disinformation has spread through social media and the internet, thereby sharpening debate on the governance of social media. Meanwhile, non-democratic regimes have made use of the pandemic to crack down on political opposition by restricting freedom of expression and freedom of the media.

Different responses have been initiated to tackle disinformation. Legislative and executive bodies have tried to regulate the spread of disinformation. Responses have gone from elaborating codes of practice and best practice guides to enabling verification networks that debunk disinformation. Corporations have also launched some initiatives to contain disinformation in their cyber-spaces, although it has proved very difficult to pursue all disinformation on the internet. Civil society has also been mobilised in the fight against disinformation and the protection of human rights online.

European institutions have begun to develop a series of instruments to fight disinformation, both internally and externally. The EU's rights-based culture has helped embed a human rights approach in its internal actions. In parallel, the EU has begun to build stronger human rights and democracy considerations into its external actions against disinformation and deceptive influence strategies. The EU's policy instruments have improved in this regard over recent years.

Despite this progress, EU efforts to tackle disinformation outside Europe still need to be infused with a stronger human-rights focus and *ethos*. The basic challenge remains to find a way of building strategies against disinformation more fully into the EU's overarching approach towards human rights internationally. While the EU's activity in the field of disinformation has expanded, it has done so as a stand-alone area of policy rather than as one element integral to polices aimed at rights infringements across the world. In its external relations, the EU has tended to approach disinformation mainly as a geopolitical problem – other powers using it to weaken the EU – instead of measuring the human-rights impacts within third countries. Furthermore, while rhetorically the EU insists that it is committed to ensuring that disinformation does not weaken global human rights, in practice it has deepened commercial and security partnerships with many regimes guilty of using disinformation to abuse human rights and undermine democratic checks and balances.

The European Parliament has an important role to play in ensuring that the human rights dimension to counter-disinformation is more prominent in EU external actions. It holds influence over EU laws, external funds and third-country agreements. It is the appropriate place to have more political discussions around disinformation. Invitations to intervene in the DROI subcommittees and the INGE Special Committee could be extended to more national MPs (also from outside the EU), human rights defenders and activists as well as CSO representatives to learn from other contexts and experiences. Engagement among these actors could eventually lead to the drafting of a UN Convention on Universal Digital (Human) Rights. The EP also needs to use its public diplomacy function, giving voice to those who do not have it in their own countries, improving capacity building and exchanging best practices with other parliaments on how to fight disinformation.

The EU's guiding principle must be to combat disinformation without infringing freedom of expression. This means being attentive to counter-disinformation tactics that undermine human rights and rather finding other ways to build up the incentives and capacities to lessen disinformation's reach around the world. Content regulation can lead to censorship, internet shutdowns and the prosecution of dissenting voices all ostensibly in the name of fighting disinformation. A human rights perspective needs to be built into debates on regulations, not only within the EU but also with regard to EU actions in third countries.

The EU should work to prevent the spread of disinformation through holding digital platforms more accountable for their impacts on democratic rights. The EU should use its diplomatic weight and leverage to push tech platforms and other governments to coordinate in applying accountability and transparency standards in third countries – including the US which is the largest companies' home base. The European Parliament can also play a role in persuading platforms to implement better standards worldwide, not just in Western states. More proactive parliamentary diplomacy – through EP relations with parliaments all over the world but also with private stakeholders – would boost the EU's capacity to lead on improving global governance, both in terms of regulations on illegal content that more firmly protect human rights and a more positive commitment from large tech companies.

The EU could also explore the use of restrictive measures as part of a rights-based approach. Experts have started debating how penalties could be imposed on those who intentionally seek to manipulate information against the EU and its Member States[201]. This could in the future lead to some sort of sanctions regime specifically for disinformation. As outlined above, the focus here has been not on human rights abuses within third countries, but the use of disinformation against the Union. The EU has generally been cautious in its use of sanctions and punitive approaches are unlikely to be at the forefront of EU external actions. Yet, the EU should be open to exploring subtle and careful ways of beginning to tighten the pressure on regimes guilty of systematic disinformation as part of their authoritarian playbooks.

8.1 Empowering Societies against Disinformation

Chapter 5 of the study suggests that responses to disinformation are needed at different levels: laws and regulations; corporate actions; and civil society. As well as concentrating on online techniques regarding various regulations and platforms, the EU should also strengthen its work at civil society level. It should offer formal and more generous support to third country local communities in their efforts to combat disinformation. This approach is clearly most consistent with a human rights approach and also resonates with the kind of policy instruments that the EU has at its disposal in third countries. We put forward five recommendations that could be useful at grassroots level:

[201] Michael Peel and Max Seddon, EU imposes sanctions on 6 Russian officials over Navalny poisoning, *Financial Times*, Published on 15 October 2020; RFE/RL, U.S. Imposes New Sanctions Targeting Russian 'Troll Farm,' Owner Prigozhin, *Radio Free Europe Radio Liberty*, Published on 30 September 2019.

8.1.1 Supporting local initiatives addressing disinformation

The EU should increase support for local initiatives that seek to give actors within third countries their own tools and capacities to push back against disinformation. The EU has supported such initiatives and officials running its aid policies have come to adopt a narrative of 'human-centric digitalisation'. However, there remains much scope for such initiatives to become a more central element of the EU's approach to disinformation. The EU could stipulate that more of its funding will in the future go to these rights-oriented initiatives. This would help contextualise disinformation as part of local communities' broader struggles for political influence and control. The EP should push to influence EU programming in this direction.

8.1.2 Enhancing support to media pluralism within disinformation strategies

EU policies have, of course, stressed the importance of media pluralism for many years, yet this is arguably still not a prominent priority within the Union's overarching external action. The EU could do more to support a wider range of media sources within third countries, recognising that this is likely to be more valuable in the longer term than, for instance, funding fact-checking projects with the largest organisation in a particular third-country partner. Media pluralism might also be made more of a critical prerequisite for the EU when certain kinds of trade, aid and security benefits are being offered to third countries. The EP should use its leverage over trade agreements to this end.

8.1.3 Responding rapidly to disinformation surges

The EU could ring-fence a certain amount of funding for quick release when disinformation campaigns spike to dangerous levels. This often happens around the time of elections, when long-standing conflicts flare up, during important international summits or at times of health emergencies. The EU should be able to provide funds quickly to local actors with a strong presence at community level where the impacts of disinformation spikes are likely to be most acute.

The EU has admirably supported flexible funding to protect individual human rights defenders when they face attacks from state authorities. This has been one area in which EU rights policies have grown in strength over recent years. It could be complemented with funding aimed not so much at individual activists in danger, but rather at preparing communities to engage in information-related activities around certain events when societies as a whole will need stronger resilience to combat spikes in disinformation.

8.1.4 Empowering small-scale deliberative forums targeting disinformation

The EU should support participative deliberation initiatives as a means of tackling disinformation. Its support is beginning to increase for small-scale forums as part of democracy and human rights strategies. Yet, its approach to disinformation is not linked with such initiatives. The EU should make more effort to combine these two important strands of its external actions.

Participative deliberation that gives citizens a specific mandate for discussing disinformation could be a vital and as yet unexplored strand within a more human-rights and empowerment-oriented approach to this challenge. It should go hand-in-hand with more training for independent community-level journalists and for political party officials in third countries as they struggle with disinformation in upholding democratic rights. Participation needs to be part of the way in which disinformation is tackled.

8.1.5 Developing human rights training

In recent years, the EU and other donors have run a large number of programmes to train human rights and democracy NGOs in digital literary and tech skills. This kind of training will continue to be important. Conversely there has been a significant absence of training tech-oriented communities in wider issues of democracy and human rights. There are huge numbers of civil society initiatives around the world now run by tech experts and focused on very technical aspects of content control and the like. Often these 'tech NGOs' are disconnected from the human rights community and commonly admit that they engage little

on the question of how their tech work relates to democratic quality and rights issues. The EU could make an express and concerted attempt to correct this imbalance. The need today is not so much to have even more standard tech training, but rather for increased capacity-building on the nexus between tech and democracy.

8.2 Global dialogue

EU policymakers recognise the need to do more at global level to tackle these challenges and are open to building on the approaches they have begun to develop in recent years. Yet they stress that this will need a deeper shift to complement regulation-based initiatives with rights-based thinking on disinformation. The EP in particular should promote a broad global forum among democracies, specifically aimed at ensuring that as new regulations, laws and standards regarding online content move forward, these are accompanied with additional efforts to strengthen democratic capacity over disinformation in third countries. Crucially, this forum for global dialogue would specifically and operationally focus on human rights and democracy concerns. The EP could be a lead sponsor in this regard, helping to raise its profile specifically on the human rights dimension of counter-disinformation. The EP should also make use of parliamentary diplomacy tools to create a best practices forum among legislatures. These efforts should be dovetailed into the civil society roadmaps (and other tools outlined previously) that would bring civil actors into exchanges with parliamentarians.

In conclusion: these ideas would help shift disinformation from a stand-alone policy area to becoming a core strand of the EU's external human rights and democracy policies. While the EU has made significant steps forward in efforts to limit disinformation, there is still room to give greater prominence to its rights-oriented dimension. While the challenges are severe, the EU can draw from many encouraging best practices which are now taking shape across the world. A fully developed human rights approach would imply that rights do not simply need to be protected online from overly draconian counter-disinformation measures, but rather be perceived as a more positive imperative in combatting disinformation. It would move from a defensive position to an empowerment-oriented, proactive stance on the relationship between disinformation and external human rights support. By offering third countries greater assistance in human rights protection, the EU would thereby also deliver a tremendous increase in their resilience to disinformation.

Bibliography

Amnesty International, Out of Control: Failing EU Laws for Digital Surveillance Export, 2020.

Avaaz, 'Facebook's Algorithm: A Major Threat to Public Health', Avaaz report, 2020.

Bentzen, Nadja, At a glance: Understanding propaganda and disinformation, European Parliament Research Services (EPRS), European Parliament, 2015.

Bontcheva, Kalina and Posetti, Julie (eds.), Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression, UNESCO Broadband Commission Report, 2020.

Bradshaw, Samantha and Howard, Philip N., The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation, University of Oxford Working Paper 2019(3), 2019.

Brady, Madeline, Deepfakes: a new disinformation threat?, *Democracy Reporting International*, 2020.

Brennen, Scott J., Simon, Felix, Howard, Philip N. Nielsen, Rasmus Kleis, Types, Sources, and Claims of COVID-19 Misinformation, Reuters Institute for the Study of Journalism, 2020.

Center for Countering Digital Hate, The Anti-vaxx industry. How Big-Tech powers and profits from vaccine misinformation, in Burki, T. (2020), The online anti-vaccine movement in the age of COVID-19, *The Lancet*, Vol. 2, 2020.

Chaccour, Carlos and Vilasanjuan, Rafael, Infodemic: has the epidemic of misinformation affected the response to COVID-19?, ISGlobal, Barcelona Institute for Global Health, 2020.

Colomina, Carme, Techno-multilateralism: The UN in the age of post-truth diplomacy, in Bargués, P., *UN@75: Rethinking multilateralism*, CIDOB Report, Vol. 6, 2020.

Council of the European Union, Council conclusions on Complementary efforts to enhance resilience and counter hybrid threats, 14972/19, 2019.

Council of the European Union, Council conclusions on strengthening resilience and countering hybrid threats, including disinformation in the context of the COVID-19 pandemic, 14064/20, 2020.

Council of the European Union, Council decision concerning restrictive measures against cyber-attacks threatening the Union or its Member States, 7299/19, May 2019.

Council of the European Union, Democracy: EU adopts conclusions, Press Release of 14 October 2019.

Daniels, Rorry, Taiwan's unlikely path to public trust provides lessons for the US, *The Brookings institute*, 2020.

EUROPOL, Catching the virus cybercrime, disinformation and the COVID-19 pandemic, Europol, 2020.

Diamond, Larry, The Democratic Rollback. The Resurgence of the Predatory State, *Foreign Affairs*, Vol. 87(2), 2008, pp 36–48.

Dobber, Tom, Metoui, Nadia, Trilling, Damian, Helberger, Natali, de Vreese, Claes, Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes, *The International Journal of Press/Politics*, Vol 26(1), 2021, pp 69-91.

Dobber, Tom; Ronan Ó Fathaigh and Frederik J. Zuiderveen Borgesius, The regulation of online political micro-targeting in Europe, *Journal on internet regulation*, Vol. 8(4), 2019.

Dubay, Carolyn. A., A Refresher on the Principle of Non-Intervention, *International Judicial Monitor*, Spring Issue, 2014.

Engler, Alex, Fighting deepfakes when detection fails, *The Brookings Institute*, 2019.

European Commission, Tackling online disinformation, 2021.

European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Region on the European Democracy Action Plan, COM(2020) 790 final, 2020.

European Commission, Joint Communication to the European Parliament, the European Council, the Council, The European Economic and Social Committee and the Committee of the Regions on Tackling COVID-19 disinformation - Getting the facts right, JOIN(2020) 8 final, 2020.

European Commission, Fourth set of reports – Fighting COVID-19 disinformation Monitoring Programme, 2020.

European Commission, Action Plan on disinformation: Commission contribution to the European Council, (13-14 December 2018), 2018.

European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Tackling online disinformation: a European Approach, COM/2018/236 final, 2018.

European Commission, Joint Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Action Plan against Disinformation, JOIN(2018) 36 final, 2018.

European Commission, Code of Conduct on Countering Illegal Hate Speech Online, 2016.

European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on The roots of democracy and sustainable development: Europe's engagement with Civil Society in external relations, COM(2012) 492 final, 2012.

European Commission, No Disconnect Strategy Workshop: European Capability for Situational Awareness (ECSA), Directorate-General for Communications Networks, Content and Technology (DG CONNECT), 2012.

European Council, A New Strategic Agenda 2019-2024, 2019.

European Parliament, Resolution of 23 November 2016 on EU strategic communication to counteract propaganda against it by third parties, P8_TA(2016)0441, 2016.

European Partnership for Democracy and Netherlands Institute for Multiparty Democracy, Thinking democratically: recommendations for responding to the phenomenon of shrinking space, EPD Study, 2020.

European Partnership for Democracy, Louder than words? Connecting the dots of European democracy support, EPD Report, 2019.

Feldstein, Steven, How Artificial Intelligence Is Reshaping Repression, *Journal of Democracy*, Vol. 30, 2019, pp 40-52.

Flore, Massimo, Understanding Citizen's Vulnerabilities: form Disinformation to Hostile Narratives, JRC Technical Report, European Commission, 2020.

Fregoso, Juliana, Mexico's Election and the Fight against Disinformation, *European Journalism Observatory*, 2018.

Garside, Susanna, Democracy and Digital Authoritarianism: An Assessment of the EU's External Engagement in the Promotion and Protection of Internet Freedom, EU Diplomacy paper 01/2020, College of Europe, 2020.

Ghosh, Dipayan, What Is Microtargeting and what is it doing in our politics, *in Internet Citizen*, Harvard University, 2018.

Gillespie, Tarleton, The platform metaphor, revisited, High Science Blog, Institut fur Internet und Gesellschaft, 2017.

Hudock, Ann, The digital disruptions of Human Rights, *Counterpart International*, 2019.

Hwang, Tim, Digital disinformation: A primer, Atlantic Council Eurasia Group, 2017.

Independent High level Group on fake news and online disinformation, A multi-dimensional approach to disinformation, Report for the European Commission, 2018.

International Telecommunication Union, Measuring Digital Development: Facts and Figures 2019, ITU Report, 2019.

Inter-Parliamentary Union, Key conclusions from the April 2019 expert hearing on disinformation and fake news, *Innovation tracker*, Issue 2, 2019.

Ipsos Public Affairs and Centre for International Governance Innovation (CIGI), Global Survey Internet Security & Trust: 2019 Part 3: Social Media, Fake News & Algorithms, CIGI-IPSOS Global Survey Report, Vol 3, 2019.

Jeangène Vilmer, Jean-Baptiste, The #Macron Leaks Operation: a post-mortem, *Atlantic Council*, 2019.

Jones, Kate, Online Disinformation and Political Discourse: Applying a Human Rights Framework, Chatham House Research paper, 2019.

Jones, Kate, Persuasion or Manipulation? Limiting Campaigning Online, Chatham House Expert Comment, 2021.

Judit Bayer; Bitiukova, Natalija; Bárd, Petra; Szakács, Judit; Alemanno, Alberto; and Uszkiewicz, Erik 'Disinformation and propaganda – impact on the functioning of the rule of law in the EU and its Member States', Directorate General for Internal Policies of the Union (IPOL), European Parliament, 2019.

Kavanagh, Jennifer and Michael D. Rich, Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life. RAND Corporation Research Reports, 2018.

Knuutila, Aleksi, Aliaksandr Herasimenka, Hubert Au, Jonathan Bright and Philip N. Howard, The Spread of Misinformation Videos on Social Media and the Effectiveness of Platform Policies, Oxford Internet Institute, 2020.

Kornbluh, Karen and Goodman, Ellen P., Safeguarding Digital Democracy. Digital Innovation and Democracy Initiative Roadmap, The German Marshall Fund of the United States DIDI Roadmap n 4, 2020.

Kuczerawy, Aleksandra, Fighting Online Disinformation: Did the EU Code of Practice Forget about Freedom of Expression?, in Terzis, Georgios, Dariusz Kloza, Elzbieta Kuzelewska and Daniel Trottier (eds.), Disinformation and Digital Media as a Challenge for Democracy *European Integration and Democracy Series*, Vol. 6, 2020.

Lewis, Rebecca and Marwick, Alice, Taking the Red Pill: Ideological motivations for Spreading Online Disinformation, in *Understanding and Addressing the Disinformation Ecosystem*, Annenberg School for Communication, December 2017.

Marsden, Chris and Meyer, Trisha, Regulating disinformation with artificial intelligence, European Parliamentary Research Service (EPRS), European Parliament, 2019.

Martin-Rozumiłowicz, Beata and Rasto Kužel, Social Media, Disinformation and Electoral Integrity, *International Foundation for Electoral Systems*, Working Paper, 2019.

Mueller, Robert, Report on the Investigation into Russian Interference in the 2016 Presidential Election, U.S. Department of Justice, March 2019.

Ness, Susan, Freedom and Accountability: A Transatlantic Framework for Moderating Speech Online, Report, German Marshall Fund, 2020.

Newman, Nic, Richard Fletcher, Anne Schulz, Simge Andi and Rasmus Kleis Nielsen, Reuters Institute Digital News Report 2020, University of Oxford, 2020.

Nimmo, Ben, Camille François, C Shawn Eib, Léa Ronzaud, "IRA Again: Unlucky Thirteen". GRAPHIKA Report, 2020.

Nockleby, John T., Hate Speech. In Leonard W. Levy, Kenneth L. Karst, and Dennis J. Mahoney (editors), *Encyclopedia of the American Constitution*, Macmillan, 2000, pp 1277–1279.

OSCE, Joint declaration on freedom of expression and "fake news", disinformation and propaganda, 2017.

OSCE, Joint Declaration on Freedom of Expression and Elections in the Digital Age, 2020.

PACE, Democracy Hacked? How to Respond?, Resolution 2326 of the Parliamentary Assembly of the Council of Europe on 31 January 2020 (9th Sitting), January 2020.

Pamment, James, Nothhaft, Howard and Fjällhed, Alicia., Countering Information Influence Activities: The State of the Art, Lund University, 2018.

Pamment, James, The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework, Paper, *Carnegie Endowment for International Peace*, 2020.

Pauwels, Eleonore. The New Geopolitics of Converging Risks. The UN and Prevention in the Era of AI, United Nations University Centre for Policy Research, 2019.

Pomerantsev, Peter, Human rights in the age of disinformation, *Unherd*, 2020.

Rebello, K., Schwieter, C., Schliebs, M., Joynes-Burgess, K., Elswah, M., Bright, J. & Howard, P. N., "COVID-19 News and Information from State-Backed Outlets Targeting French, German and Spanish-Speaking Social Media Users". University of Oxford, 2020.

Repnikova, Maria, China's Lessons for Fighting Fake News, *Foreign Policy*, 2018.

Robinson, Olga, Coleman, Alistair and Sardarizadeh, Shayan, A report of Anti-Disinformation Initiatives, Oxford Internet Institute, 2019.

Schiener, Bruce, 8 Ways to Stay Ahead of Influence Operations, *Foreign Policy*, 2019.

Spaulding, Suzannne, Nair, Devi, and Nelson, Arthur, Beyond the Ballot: How the Kremlin Works to Undermine the U.S. Justice System, Center for Strategic and International Studies, 2019.

Tenove, Chris, "Protecting democracy from disinformation: normative threats and policy responses", *The International Journal of Press/Politics*, Vol. 25(3), 2020, pp 517-537.

Thomas, Elise, Thompson, Natalie, and Wanless, Alicia, The Challenges of Countering Influence Operations, Paper, Carnegie Endowment for International Peace, 2020.

UK Parliament Digital, Culture, Media and Sport Committee , Disinformation and 'fake news': Final Report, *House of Commons*, Eighth Report of Session 2017–19, 2019.

UN OHCHR, Guiding Principles on Business and Human Rights, New York and Geneva: United Nations Office of the United Nations High Commissioner for Human Rights, 2011.

UNHR, Monitoring Human Rights in the Context of Elections, in *Manual on Human Rights Monitoring*, Ch. 23, 2011.

Vosoughi, Soroush; Roy, Deb y Aral, Sinan, The spread of true and false news online. *Science*, vol. 359(6380), 2018, pp 1146-1151.

Wardle, Claire and Derakhshan, Hossein, Information Disorder: Toward an interdisciplinary framework for research and policy making, Council of Europe report DGI(2017)09, 2017.

WHO, Joint statement by WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, and IFRC on Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation, 2020.

Youngs, Richard, Civic Activism Unleashed: New Hope or False Dawn for Democracy?, Oxford: *Oxford University Press*, 2019.

PE 653.635
EP/EXPO/DROI/FWC/2019-01/LOT6/R/02

Print ISBN 978-92-846-8015-3 | doi:10.2861/677679 | QA-02-21-559-EN-C
PDF ISBN 978-92-846-8014-6 | doi:10.2861/59161 | QA-02-21-559-EN-N