

Freedom of Expression and the Regulation of Disinformation in the European Union

András Koltay

6.1 INTRODUCTION

The issue of mass disinformation on the Internet is a long-standing concern for policymakers, legislators, academics and the wider public. Disinformation is believed to have had a significant impact on the outcome of the 2016 US presidential election.¹ Concern about the threat of foreign – mainly Russian – interference in the democratic process is also growing.² The COVID-19 pandemic, which reached global proportions in 2020, gave new impetus to the spread of disinformation, which even put lives at risk.³ The problem is real and serious enough to force all parties concerned to reassess the previous European understanding of the proper regulation of freedom of expression.

The author would like to thank all those who read and commented on earlier versions of the manuscript at various conferences and workshops, especially Eduardo Bertoni, Joanna Botha, John Charney, Mark Cole, Michael Epstein, Domingos Farinho, Andrew Kenyon, Ron Krotoszynski, Michael Losavio, Péter Nádori, Bernát Török, Louis Virelli, Russell Weaver, Christopher Yoo, Vincenzo Zeno-Zencovich and Zsolt Zódi.

¹ Alexandre Bovet and Hernán A. Makse, 'Influence of Fake News in Twitter during the 2016 US Presidential Election' (2019) 10(7) *Nature Communications*, <https://doi.org/10.1038/s41467-018-07761-2>.

² See, e.g., Richard Sakwa, *The Russia Scare: Fake News and Genuine Threat* (Abingdon: Routledge, 2022); Foreign Threats to the 2020 U.S. Federal Elections. Intelligence Community Assessment ICA 2020-00078D. National Intelligence Council, 10 March 2021, www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf; Christopher Paul and Miriam Matthews, 'The Russian "Firehose of Falsehood" Propaganda Model: Why It Might Work and Options to Counter It', RAND, 2016, www.rand.org/pubs/perspectives/PE198.html.

³ Hendrik Bruns, François J. Dessart and Myrto Pantazi, 'Covid-19 Misinformation: Preparing for Future Crises' EUR 31139 EN, Luxembourg, Publications Office of the European Union, 2022, <https://publications.jrc.ec.europa.eu/repository/handle/JRC130111>.

This chapter reviews the measures taken by the European Union and its Member States to limit disinformation, mainly through regulatory instruments. After a clarification of the concepts involved (Section 6.2), I will review the options for restricting false statements which are compatible with the European concept of freedom of expression (Section 6.3), and then examine the related tools of media regulation (Section 6.4). This will be followed by a discussion of the regulation of online platforms in the EU (Section 6.5), and by a presentation of EU (Section 6.6) and national (Section 6.7) measures which specifically address disinformation. Finally, I will attempt to draw some conclusions with regard to possible future regulatory responses (Section 6.8).

6.2 DEFINITIONAL ISSUES

Not only are the categories of fake news, disinformation and misinformation not precisely defined in law, but their exact meaning is also disputed in academia. Since the 2016 US presidential election campaign, the use of the term ‘fake news’ has spread worldwide. It is usually applied to news published on a public platform, in particular on the Internet, that is untrue in content or misleading as to the true facts, and which is not published with the intention of revealing the truth but with the aim of deliberately distorting a democratic process or the informed resolution of a public debate.⁴ According to Hunt Allcott and Matthew Gentzkow, fake news is news that is ‘intentionally and verifiably false, and could mislead readers’,⁵ meaning that intentionality and verifiable falsehood are important elements of it. However, in principle, fake news could also include content that is specifically protected by freedom of expression, such as political satire, parody and subjective opinions, a definition which would certainly be undesirable in terms of the protection of freedom of expression.⁶ Since President Trump, after his successful campaign in 2016, mostly applied the term to legacy media that was critical of him, it has gradually lost its original meaning and has fallen out of favor in legal documents.⁷

The EU has for some time preferred the term ‘disinformation’ to describe the phenomenon. Of course, fake news and disinformation are in fact two categories with a significant overlap, and as Bj rnstjern Baade points out, the former will not disappear from the public sphere either, so legislators, law enforcers and public

⁴ Alessio Sardo, ‘Categories, Balancing, and Fake News: The Jurisprudence of the European Court of Human Rights’ (2020) 33(2) *Canadian Journal of Law & Jurisprudence* 435–60, 451.

⁵ Hunt Allcott and Matthew Gentzkow, ‘Social Media and Fake News in the 2016 Election’ (2017) 31(2) *Journal of Economic Perspectives* 211–36, 213.

⁶ Rachael Craufurd Smith, ‘Fake News, French Law and Democratic Legitimacy: Lessons for the United Kingdom?’ (2019) 11(1) *Journal of Media Law* 52–81, 57.

⁷ Iri  Katsirea, ‘“Fake News”: Reconsidering the Value of Untruthful Expression in the Face of Regulatory Uncertainty’ (2019) 10(2) *Journal of Media Law* 159–88, 162.

policymakers will have to continue dealing with it.⁸ Tarlach McGonagle cuts the Gordian knot by defining fake news as content that is ‘disinformation that is presented as, or is likely to be perceived as, news’.⁹ The Code of Practice on Disinformation, in line with several EU documents, defines disinformation as ‘false or misleading content that is spread with an intention to deceive or secure economic or political gain and which may cause public harm’.¹⁰ Thus, intentional deception and the undue gain accrued or harm it causes are also conceptual elements here. By comparison, misinformation is ‘false or misleading content shared without harmful intent though the effects can still be harmful, e.g. when people share false information with friends and family in good faith’.¹¹ However, the inclusion of intentionality as an essential characteristic of disinformation may also raise concerns. It is inherently problematic to limit speech on the basis of a speaker’s intent, and not merely on the basis of the effect achieved. Furthermore, while there is a consensus that satire and parody, being protected opinions, cannot be considered disinformation, they can also be published in bad faith, distorting the true facts, which can have an effect similar to that which attempts to suppress disinformation seek to prevent.¹²

The Code of Practice approved by the EU focuses primarily on curbing disinformation in political advertising, which, according to the European Council and Parliament’s proposal for a regulation, ‘means the preparation, placement, promotion, publication or dissemination, by any means, of a message: (a) by, for or on behalf of a political actor, unless it is of a purely private or a purely commercial nature; or (b) which is liable to influence the outcome of an election or referendum, a legislative or regulatory process or voting behavior’ (on the Code, see Section 6.6.1).

The distinction between disinformation and misinformation, or in other words, the difference between falsehoods made with the intent to harm and untruths that are communicated in good faith but are likely to cause harm, is indeed important and each problem warrants different levels of action and intervention. Prosecuting both disinformation and misinformation with equal force might lead to an

⁸ Björnstjern Baade, ‘Don’t Call a Spade a Shovel: Crucial Subtleties in the Definition of Fake News and Disinformation’, *Verfassungsblog*, 14 April 2020, <https://verfassungsblog.de/dont-call-a-spade-a-shovel>.

⁹ Tarlach McGonagle, ‘“Fake News”: False Fears or Real Concerns?’ (2017) 35(4) *Netherlands Quarterly of Human Rights* 203–9, 203.

¹⁰ 2022 Strengthened Code of Practice on Disinformation, I. Preamble, <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>. See also 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. Brussels, 3.12.2020, COM(2020)790 final, Art. 17(4).

¹¹ *Ibid.*

¹² Maria L. Stasi and Pier L. Parcu, ‘Disinformation and Misinformation: The EU Response’ in Pier L. Parcu and Elda Brogi (eds.), *Research Handbook on EU Media Law and Policy* (Cheltenham: Edward Elgar, 2021) pp. 407–26, p. 408.

unfortunate situation in which citizens who wish to participate in the debate on public affairs, but who do not have the means to verify the truth of a piece of information or communication, who have no malicious intent and seek no direct personal gain, would suffer a disproportionate restriction on their freedom of expression. The most dangerous form of disinformation is that which comes from governments and public bodies. Tackling this is a separate issue, which allows for the use of more robust instruments.¹³ For example, in March 2022, the Council of the EU banned certain Russian television broadcasters on that basis following the outbreak of the Russian–Ukrainian war.¹⁴

To summarize the above brief conceptual overview, the current approach in the EU is to consider as disinformation content that: (a) is untrue or misleading; (b) is published intentionally; (c) is intended to cause harm or undue gain; (d) causes harm to the public; (e) is widely disseminated, typically on a mass scale and (f) is disseminated through an internet content service. Points (c) and (f) are not conceptual elements but refer to the usual characteristics of disinformation. Consequently, distorted information resulting from possible bias in the traditional media is not considered disinformation, nor is the publication of protected opinions (satire, parody). Since a specific characteristic of disinformation is that it is spread mainly on the Internet, in particular on social media platforms, attempts at preventing it focus especially on these services.

6.3 THE RESTRICTION OF FALSE STATEMENTS IN THE EUROPEAN FREE SPEECH DOCTRINE

The European approach to disinformation, unlike that of the United States, allows for a broad restriction of certain false statements. The US Supreme Court in *United States v. Alvarez* held that the falsity of an allegation alone is not sufficient to exclude it from First Amendment protection,¹⁵ but that does not mean that untrue statements of fact, if they cause harm, cannot be restricted, albeit within a narrower range than that of the EU.¹⁶ While the extent to which and under what circumstances disinformation is restricted in Europe is a matter for national law, normative considerations generally take into account the following three requirements when assessing a restriction on speech: the principle of legality (that the restriction is

¹³ See Katie Pentney, ‘Tinker, Tailor, Twitter, Lie: Government Disinformation and Freedom of Expression in a Post-Truth Era’ (2022) 22(2) *Human Rights Law Review* 1–29.

¹⁴ Council Regulation (EU) 2022/350 of 1 March 2022 amending Regulation (EU) No 833/2014 Concerning Restrictive Measures in View of Russia’s Actions Destabilising the Situation in Ukraine. See also Section 6.6.1.

¹⁵ *United States v. Alvarez*, 567 US 709 (2012).

¹⁶ See Leslie Gielow Jacobs, ‘Freedom of Speech and Regulation of Fake News’ (2022) 70(1) *American Journal of Comparative Law* 278–311.

provided for by an appropriate rule, preferably codified law), the principle of necessity (that the restriction is justified in a democratic society) and the principle of proportionality (that the restriction does not go beyond the legitimate aim pursued).¹⁷

Within the framework of the protection of freedom of expression in Europe, according to the current doctrine, deliberate lies (intentional publication of untruthful information) may not be subject to a general prohibition. This does not mean that it is not permissible in certain circumstances to prohibit false factual statements but that a general prohibition is usually understood to be incompatible with the doctrine of freedom of speech. The special circumstances in which speech may be prohibited can be grouped into several areas.

First, defamation law and the legal protection of reputation and honor seek to prevent unfavorable and unjust changes being made to an individual's image and evaluation by society. These regulations aim to prevent an opinion published in the public sphere concerning an individual from tarnishing the 'image' of an individual without proper grounds, especially when it is based upon false statements. The approaches taken by individual states to this question differ noticeably, but their common point of departure is the strong protection afforded to debates on public affairs and the correspondingly weaker protection of the personality rights of public figures compared to the protection of freedom of speech.¹⁸

Second, the EU Council's Framework Decision on combating racism and xenophobia in the Member States of the EU¹⁹ places a universal prohibition on the denial of crimes against humanity, war crimes and genocide. Most Member States of the EU have laws prohibiting the denial of the crimes against humanity committed by the Nazis before and during World War II, or the questioning of those crimes or watering down their importance.²⁰

Third, a number of specific rules apply to false statements made during election campaigns. These can serve two purposes. On the one hand, communication in the campaign period enjoys robust protection: political speech is the most closely guarded core of freedom of expression, and what is spoken during a campaign is as closely linked to the functioning of democracy and democratic procedures as any

¹⁷ Rebecca H. Helm and Hitoshi Nasu, 'Regulatory Responses to "Fake News" and Freedom of Expression: Normative and Empirical Evaluation' (2021) 21(2) *Human Rights Law Review* 302–28, 308.

¹⁸ See *Lingens v. Austria*, app. no. 9815/82, decision of 8 July 1986, and the many cases decided by the ECtHR, www.echr.coe.int/Documents/FS_Reputation_ENG.pdf.

¹⁹ Council Framework Decision 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law.

²⁰ See the French 'Gayssot Act' (Loi no. 90-615 du 13 juillet 1990 tendant à réprimer tout acte raciste, antisémite ou xénophobe, amending the Law on the Freedom of the Press of 1881, by adding a new Article 24) and the German Criminal Code (Strafgesetzbuch), Art. 130(3).

speech can be. On the other hand, these procedures must also be protected so that no candidate or community party distorts the democratic decision-making process and ultimately damages the democratic order.²¹

Fourth, commercial communication can be regulated in order to protect consumers from false (misleading) statements. The European Court of Human Rights (ECtHR), in *Markt Intern and Beerman v. Germany*,²² declared that advertisements serving purely commercial interests, rather than contributing to debates in the public sphere, are also to be awarded the protection of the freedom of speech.²³ Nevertheless, this protection is of a lower order than that granted to ‘political speech’.

Fifth, in some jurisdictions, ‘scaremongering’ – that is, the dissemination of false information that disturbs or threatens to disturb public order or peace – may also be punishable.²⁴

Another example of an indirect ban on untrue statements is tobacco advertising. The EU has a broad ban on the subject,²⁵ which may be further strengthened by national regulations. The advertising ban includes, by definition, the positive portrayal of tobacco, while the publication of opinions other than advertising arguing for the potential positive effects of tobacco is obviously not banned from the public discourse.

6.4 EUROPEAN AND NATIONAL MEDIA REGULATION

Hate speech can also be tackled through media regulation. The Audiovisual Media Services Directive requires Member States to prohibit incitement to violence or hatred directed against a group of persons or a member of a group on the grounds of race, sex, religion or nationality as well as public provocation to commit terrorist offences in linear and nonlinear, television and other audiovisual media services (Article 6). Member States have transposed these provisions into their national legal systems. Under the Directive, only the authority of the state in which the media service provider is broadcasting has jurisdiction to verify whether the conduct in question constitutes hate speech, and to ensure that the broadcasts of the media

²¹ See, e.g., the UK’s Representation of the People Act 1983, s. 106 (False statements as to candidates); Austrian Penal Code (Strafgesetzbuch), § 264 (‘Spreading fake news during an election or referendum’).

²² *Markt Intern and Beerman v. Germany*, app. no. 10572/83, judgment of 20 November 1989.

²³ *Ibid.* [36].

²⁴ This is the case in Hungary (2012 Criminal Code, s. 337); see András Koltay, ‘On the Constitutionality of the Punishment of Scaremongering in the Hungarian Legal System’ in *Hungarian Yearbook of International Law and European Law* (The Hague: Eleven Publishing, 2021) pp. 23–42.

²⁵ Directive 2003/33/EC of the European Parliament and of the Council of 26 May 2003 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the advertising and sponsorship of tobacco products (Text with EEA relevance).

service provider do not contain incitement to hatred or violence. If a media service provider is not established in an EU Member State, it is not subject to the provisions of the Directive, and the national authorities can take action against it under their own legal systems. According to the well-established case law of the Court of Justice of the EU and the ECtHR, a television broadcaster which incites terrorist violence cannot itself claim freedom of expression.²⁶

Other (indirect) measures can also be applied against disinformation in media regulation. Access to the content of a media service provider is granted by the legislator based not on an external condition but on the right of reply, in response to content published previously by the service provider. The Audiovisual Media Services Directive prescribes that EU Member States should introduce national legal regulations with regard to television broadcasting that ensure adequate legal remedies for individuals whose personality rights have been infringed through false statements.²⁷ Such regulations are applied throughout Europe and typically impose obligations not only on audiovisual media but also on both printed and online press,²⁸ and the granting of the right of reply is also suggested in the EU High Level Expert Group's report on disinformation (see Section 6.6.1) as a possible tool to combat disinformation.²⁹ The promotion of media pluralism may involve a requirement for impartial news coverage, on the basis of which public affairs must be reported impartially in programs which provide information on them. Regulation may apply to television and radio broadcasters, and it has been implemented in several states in Europe.³⁰

In July 2022, the British media regulator Ofcom published its decisions on twenty-nine programs that were broadcast on Russia Today (RT) between 27 February 2022 and 2 March 2022. The licence for the RT service was, at the time of broadcast, held by Autonomous Non-Profit Organization TV-Novosti. The programs had raised

²⁶ Cases C-244/10 and C-245/10 *Mesopotamia Broadcast A/S and Roj TV A/S v. Bundesrepublik Deutschland*, judgment of the Court (Third Chamber) of 22 September 2011; *Roj TV A/S v. Denmark*, app. no. 24683/14, judgment of 17 April 2018.

²⁷ Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services ('AVMS Directive'), Art. 28.

²⁸ Kyu Ho Youm, 'The Right of Reply and Freedom of the Press: An International and Comparative Perspective' (2008) 76(4) *George Washington Law Review* 1017–64; Andrei Richter, 'Fake News and Freedom of the Media' (2018–19) 8(2) *Journal of International Media & Entertainment Law* 1–34, 14–19; András Koltay, 'The Right of Reply in a European Comparative Perspective' (2013) 54(1) *Hungarian Journal of Legal Studies – Acta Juridica Hungarica* 73–89.

²⁹ Final Report of the High Level Expert Group on Fake News and Online Disinformation (2018), <https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>, pp. 15–16.

³⁰ See, e.g., the German regulations (Rundfunkstaatsvertrag, ss. 25–34) and the UK regulation (ss. 319(2)(c) and 319(2)(d), 319(8) and 320 of the Communications Act 2003, and s. 5 of the Broadcasting Code).

issues warranting investigation under the due impartiality rules.³¹ Under Section 3(3) of the Broadcasting Act 1990 and of the Broadcasting Act 1996, Ofcom ‘shall not grant a licence to any person unless satisfied that the person is a fit and proper person to hold it’ and ‘shall do all that they can to secure that, if they cease to be so satisfied in the case of any person holding a licence, that person does not remain the holder of the licence’. Taking into account a series of breaches by RT of the British broadcasting legislation concerning the due impartiality and accuracy rules, Ofcom revoked these licences.³²

The 1936 International Convention on the Use of Broadcasting in the Cause of Peace and the 1953 Convention on the International Right of Correction would also provide for action against communications from state bodies that have a detrimental effect on international relations, but they are hardly applicable generally to disinformation or misinformation.³³

6.5 PLATFORM REGULATION IN THE EUROPEAN UNION

False claims are spreading across different online platforms at an unprecedented rate and at the same time to a massive extent. In particular, disinformation is being distributed on social media platforms which consciously focuses on electoral campaigning, for political reasons (involving political parties with conflicting interests, other states acting against a particular state and so on). Initially, the platforms defended themselves by claiming that they were neutral players in this communication.³⁴ It became increasingly obvious, however, that the platforms themselves are actively able to shape the communication on their services, and that they have an economic interest in its vigor and intensity, and hence that the spread of false news is not necessarily contrary to their interests.³⁵ Under EU law, online platforms are considered a type of host providers, whose liability for infringing content which appears in their services is limited, but by no means excluded.

6.5.1 *Directive on Electronic Commerce*

According to the Directive on Electronic Commerce, if these platforms provide only technical services when they make available, store or transmit the content of others

³¹ Ofcom Broadcasting Code, s. 5.

³² Notice of a decision under s. 3(3) of the Broadcasting Act 1990 and s. 3(3) of the Broadcasting Act 1996 in Respect of Licences TLCS 000881, TLCS 001686 and DTFS 000072 held by ANO TV Novosti, www.ofcom.org.uk/_data/assets/pdf_file/0014/234023/revocation-notice-ano-tv-novosti.pdf.

³³ See Björnstjern Baade, ‘Fake News and International Law’ (2019) 29 *European Journal of International Law* 1357–76.

³⁴ Andrew Marantz, ‘Facebook and the “Free Speech” Excuse’, *The New Yorker*, 31 October 2019, www.newyorker.com/news/daily-comment/facebook-and-the-free-speech-excuse.

³⁵ According to George Soros, Facebook was working directly to re-elect President Trump. See George Soros, ‘Remove Zuckerberg and Sandberg from Their Posts’, *Financial Times*, 18 February 2020, www.ft.com/content/88f6875a-519d-11ea-90ad-25e377c0ee1f.

(much like a printing house or a newspaper stand), then it would seem unjustified to hold them liable for the violations of others ('illegal activity or information'), as long as they are unaware that such violations have occurred. However, in the European approach, gatekeepers may be held liable for their own failure to act after becoming aware of a violation (if they fail to remove the infringing material).³⁶ The Directive requires all types of intermediaries to remove such materials after they become aware of their infringing nature (Articles 12–14). In addition, the Directive stipulates that intermediaries may not be subject to a general monitoring obligation to identify illegal activities (Article 15).

While this system of legal responsibility should not necessarily be considered outdated, things have certainly changed since 2000, when the Directive was enacted: there are fewer reasons to believe that today's online platforms remain passive with regard to content and do nothing more than store and transmit information. While content is still produced by users or other independent actors, the services of gatekeepers select from and organize, promote or reduce the ranking of such content, and may even delete it or make it unavailable within the system. This notice and takedown procedure applies to the disinformation that appears on the platforms, but resorting to the actual removal of content is reserved for disinformation that is unlawful under the legal system of the state in question (slander, terrorist propaganda, denials of genocide and so on). Generally speaking, false claims are not subject to the removal obligation as they are not illegal. Similarly, even if a piece of content is infringing, but no one reports it to the platform, there is no obligation to remove it.

The notion of 'illegal activity or information' raises an important issue, as the obligation to remove offending content is independent of the outcome of any possible court or official procedure that may establish that a violation has been committed, and the host provider is required to take action before a decision is passed (provided that a legal procedure is actually initiated). This means that the provider has to decide on the illegality of content on its own, and its decision is free from any legal guarantee (even though it may have an impact on the freedom of expression). This rule may encourage providers to remove content to escape possible liability, even in highly questionable situations. It would be comforting (but probably inadequate, considering the speed of communication) if the liability of an intermediary could not be established unless the illegal nature of the content it has not removed is established by a court.³⁷

Although continuous, proactive monitoring of infringing content is not mandatory for platforms, the European Court of Justice opened up a loophole for it well

³⁶ Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on Certain Legal Aspects of Information Society Services, in Particular Electronic Commerce, in the Internal Market ('Directive on Electronic Commerce'), Art. 14.

³⁷ Christina M. Mulligan, 'Technological Intermediaries and Freedom of the Press' (2013) 66 *SMU Law Review* 157–88, 175.

before the recent Regulation banning Russian media outlets, in 2019, in *Glawischnig-Piesczek v. Facebook Ireland*.³⁸ The decision in that case required the platform to delete defamatory statements that had been reported once and removed but which had subsequently reappeared. Likewise, the hosting provider may be obliged to ‘remove information which it stores, the content of which is identical to the content of information that was previously declared to be unlawful, or to block access to that’. This is only possible through the use of artificial intelligence, which is encouraged by this decision and even implicitly made mandatory. Putting that decision in a broader context, it seems that platforms are required to act proactively against unlawful disinformation (or any unlawful content), even given the purported continued exclusion of monitoring obligations. The legality of the content is determined by algorithms, which seems quite risky for freedom of speech.³⁹

6.5.2 *Digital Services Act*

The EU’s Digital Services Act (DSA),⁴⁰ which aims to regulate online platforms in a more detailed and nuanced way, and is applicable from 2023 and 2024, respectively, keeps the most important foundations of European regulation of online platforms in place. The response of the EU to the problem of disinformation is to legislate for more societal responsibility for very large online platforms, while still leaving it to the discretion of the platforms themselves to decide if and how to deal with any systemic risks to freedom of expression.

The DSA retains the essence of the notice and takedown procedure, and platforms still cannot be obliged to monitor user content (Articles 6 and 8), but if they receive a notification that a certain piece of content is illegal, they will be obliged to remove it (Article 6), as set out also in the Directive on Electronic Commerce. The DSA will also seek to protect users’ freedom of expression. It requires users to be informed of the content removed by platforms and gives them the possibility to have recourse to dispute resolution mechanisms in their own country, as well as to the competent authorities or courts if the platform has infringed the provisions of the DSA. These provisions seek to strengthen the position of users, in particular by providing procedural guarantees (most importantly through greater transparency,

³⁸ Case C-18/18 *Eva Glawischnig-Piesczek v. Facebook Ireland Ltd.*, judgment of the CJEU of 3 October 2019, <https://curia.europa.eu/juris/document/document.jsf?text=&docid=218621&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=7924>.

³⁹ Elda Brogi and Marta Maroni, ‘*Eva Glawischnig-Piesczek v Facebook Ireland Limited: A New Layer of Neutrality*’, CMPF, 17 October 2019, <https://cmpf.eu.eu/eva-glawischnig-piesczek-v-facebook-ireland-limited-a-new-layer-of-neutrality>.

⁴⁰ Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and Amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance).

the obligation to give reasons for deletion of a piece of content or for the suspension of an account, and the right of independent review).⁴¹

The democratic public sphere is protected by the DSA (Article 14(4)), which states that the restrictions in contractual clauses (Article 14(1)) must take into account freedom of expression and media pluralism. Article 14(4) states that:

Providers of intermediary services shall act in a diligent, objective and proportionate manner in applying and enforcing the restrictions . . . with due regard to the rights and legitimate interests of all parties involved, including the fundamental rights of the recipients of the service, such as the freedom of expression, freedom and pluralism of the media, and other fundamental rights and freedoms as enshrined in [the Charter of Fundamental Rights of the European Union].

Where platforms do not act with due care, objectivity and proportionality in applying and enforcing restrictions when deleting user content, taking due account of the rights and legitimate interests of all interested parties, including the fundamental rights of the users of the service, such as to freedom of expression, freedom and pluralism of the media, and other fundamental rights and freedoms as set out in the Charter of Fundamental Rights of the European Union (CFR), the user may have recourse to the public authorities. In regard to very large online platforms in Europe, this will most often be the designated Irish authority, to which other national authorities must also refer complaints they receive concerning these platforms, for which the European Commission has also reserved certain powers (it is for the Commission to decide whether to act itself or to delegate this power to the Irish authority).

Under the DSA, the authorities do not explicitly take action against disinformation, only doing so if it constitutes an infringement (war propaganda, which can be conducted through disinformation, can of course constitute an infringement). However, since disinformation alone does not constitute an infringement in national jurisdictions, the DSA does not introduce any substantive change in this respect. Furthermore, very large online platforms and very large online search engines must identify and analyze the potential negative effects of their operations (in particular their algorithms and recommendation systems) on freedom of expression and on ‘civil discourse and electoral processes’,⁴² and must then take appropriate and effective measures to mitigate these risks (Article 35). In addition, the DSA’s rules on codes of conduct also encourage the management of such risks and promote the enforcement of codes (including, for example, the Code of Practice on Disinformation, which predates the DSA). These tools also provide an indirect means of tackling disinformation. One of the main purposes of the DSA is to protect users’ freedom of speech, but users’ speech can also contain dis- or misinformation.

⁴¹ Ibid. Arts. 17, 21 and 24.

⁴² Ibid. Arts. 34(1)(b) and 34(1)(c).

It will be difficult to reconcile these conflicting interests when applying the regulation.

Article 36 of the DSA introduces a new ‘crisis response mechanism’. Crisis in this legislation means ‘extraordinary circumstances’ that ‘lead to a serious threat to public security or public health in the Union or in significant parts of it’ (Article 36(2)). Very large online platforms will need to assess to what extent and how the functioning and use of their services significantly contribute to a serious threat, or are likely to do so and then to identify and apply specific, effective and proportionate measures, to prevent, eliminate or limit any such contribution to the serious threat identified (Article 36(1)).

6.6 THE EUROPEAN UNION’S EFFORTS TO CURB DISINFORMATION ON ONLINE PLATFORMS

European jurisdictions allow actions against disinformation, defined as action on the grounds of defamation or the violation of the prohibition of hate speech or scaremongering, while platforms, being ‘host providers’, can be required to remove infringing content. However, these measures in and of themselves seem inadequate to deal with such threats in a reassuring manner. Concerns of this nature have been addressed by the EU in various documents it has produced since 2017.

6.6.1 *Communications, Recommendations and Co-Regulation*

The first relevant EU Communication, issued in 2017,⁴³ concerns tackling illegal content, so it only indirectly addresses the issue of disinformation. It mentions that ‘[t]here are undoubtedly public interest concerns around content which is not necessarily illegal but potentially harmful, such as fake news or content that is harmful for minors. However, the focus of this Communication is on the detection and removal of illegal content.’⁴⁴ The Communication introduced a requirement for platforms to take action against violations in a proactive manner and even in the absence of a notice, even though the platforms are still exempted from liability.⁴⁵ The Recommendation that followed the Communication reaffirmed the requirement to apply proportionate proactive measures in appropriate cases, which thus permits the use of automated tools to identify illegal content.⁴⁶

⁴³ Communication on Tackling Illegal Content Online: Towards an Enhanced Responsibility of Online Platforms, 28 September 2017, COM(2017) 555 final.

⁴⁴ *Ibid.* s. 2.

⁴⁵ *Ibid.* s. 3.3.1.

⁴⁶ Commission Recommendation of 1.3.2018 on Measures to Effectively Tackle Illegal Content Online, C(2018) 1177 final, ss. 12 and 18.

The High Level Expert Group on Fake News and Online Disinformation published a report in 2018.⁴⁷ The report defines disinformation as ‘false, inaccurate, or misleading information designed, presented and promoted for profit or to intentionally cause public harm’.⁴⁸ While this definition might be accurate, the report refrains from raising the issue of government regulation or co-regulation, and is limited to providing a review of the resources and measures that are available to social media platforms and which they may apply voluntarily. The Communication⁴⁹ issued following the report of the High Level Expert Group already recognized the need for more concrete action, not only by online platforms but also by the European Commission and Member States. The document called for a more transparent, trustworthy and accountable online ecosystem. It foresaw the reinforcement of the EU bodies concerned and the creation of a rapid alert system that would identify in real time, through an appropriate technical infrastructure, any disinformation campaign.

Later in 2018, online platforms, leading technology companies and advertising industry players agreed, under pressure from the European Commission, on a code of conduct to tackle the spread of online disinformation. The 2018 Code of Practice on Disinformation was designed to set out commitments in areas ranging from transparency in political advertising to the demonetization of disinformation spreaders. The Code may appear to be voluntary in form – that is, a self-regulatory instrument – but it is in fact a co-regulatory solution that was clearly imposed on the industry players by the European Commission. Its primary objectives are to deprive disseminators of disinformation of advertising revenue from that activity, to make it easy to identify publishers of political advertising, to protect the integrity of the platform’s services (steps against fake accounts and bots) and to support researchers and fact-checkers working on the subject. The Code actually further exacerbates the well-known problem of private censorship (the recognition of the right of platforms to restrict the freedom of expression of their users through rules of their own making),⁵⁰ by putting decisions on individual content in the hands of the platforms, which raises freedom of expression issues.⁵¹

⁴⁷ Final Report of the High Level Expert Group on Fake News and Online Disinformation (2018), <https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>.

⁴⁸ *Ibid.* p. 10.

⁴⁹ 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, 26 April 2018, COM(2018) 236 final.

⁵⁰ See András Koltay, ‘Private Censorship of Internet Gatekeepers’ (2020–2021) 59(2) *Louisville Law Review* 255–304; András Koltay, ‘The Protection of Freedom of Expression from Social Media Platforms’ (2022) 73(2) *Mercer Law Review* 523–89.

⁵¹ Matteo Monti, ‘The EU Code of Practice on Disinformation and the Risk of the Privatisation of Censorship’ in Serena Giusti and Elisa Piras (eds.), *Democracy and Fake News: Information Manipulation and Post-Truth Politics* (Abingdon: Routledge, 2016) pp. 214–25, 220–21.

The Code of Practice was signed in October 2018 by online platforms such as Facebook, Google, Twitter and Mozilla, as well as advertisers and other players in the advertising industry, and was later joined by Microsoft and TikTok. The online platforms and trade associations representing the advertising industry submitted a report in early 2019 setting out the progress they had made in meeting their commitments under the Code of Practice on Disinformation. In the first half of 2019, the European Commission carried out targeted monitoring of the implementation of the commitments by Facebook, Google and Twitter, with a particular focus on the integrity of the European Parliament elections. The Commission published its evaluation of the Code in September 2020, which found that the Code provided a valuable framework for structured dialogue between online platforms, and ensured greater transparency and accountability for their disinformation policies. It also led to concrete actions and policy changes by relevant stakeholders to help combat disinformation.⁵²

The Joint Communication of the European Parliament and of the Council on an Action Plan against Disinformation foresees the same measures as in the previous Communication in 2018.⁵³ The Communication called upon all signatories of the Code of Practice to implement the actions and procedures identified in the Code swiftly and effectively on an EU-wide basis. It also encouraged the Member States to launch awareness-raising initiatives and support fact-checking organizations. While this document reaffirms the primacy of means that are applied voluntarily by platform providers, it also displays restraint when it comes to compelling the service providers concerned to cooperate (in a forum convened by the European Commission). If the impact of voluntary undertakings falls short of the expected level, the necessity of action of a regulatory nature might arise.⁵⁴

The arrival of the COVID pandemic in Europe in early 2020 gave a new impetus to the mass spread of disinformation, this time directly threatening human lives. Therefore, the EU bodies issued a new document proposing specific measures to be taken by platforms to counter disinformation about the epidemic, but did not actually broaden the scope of the general measures on disinformation previously set out.⁵⁵ Section 4 of the European Democracy Action Plan also specifically addresses the fight against disinformation and foresees the reinforcement of the

⁵² Code of Practice on Disinformation (2018), <https://ec.europa.eu/newsroom/dae/redirection/document/87534>.

⁵³ Joint Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, Action Plan against Disinformation. Brussels, 5.12.2018, Join(2018) 36 final.

⁵⁴ *Ibid.* 9, Action 6.

⁵⁵ Joint Communication to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, Tackling COVID-19 Disinformation – Getting the Facts Right. Brussels, 10.6.2020, Join (2020) 8 final.

2018 Code of Practice, the addition of further commitments and the establishment of a monitoring mechanism.⁵⁶

In 2021, EU bodies issued a new Communication,⁵⁷ which foreshadowed the content of the updated Code of Practice. Subsequently, a review of the Code was launched, leading to the signing of the Strengthened Code of Practice on Disinformation by thirty-four signatories in June 2022.⁵⁸ The updated and strengthened Code aims to deliver on the objectives of the Commission's guidance,⁵⁹ presented in May 2021, by setting out a broader range of commitments and measures to combat online disinformation. While the Code has not been officially endorsed by the Commission, the Commission set out its expectations in its Communication, and has indicated that it considers that the Code meets these expectations overall. Since this guidance sets out the Commission's expectations in imperative terms ('the Code should', 'the signatories should', and so on), it is not an exaggeration to say that the fulfilment of the commitments is seen as an obligation for the platforms, which, if fulfilled, could avoid the imposition of strict legal regulation. Consequently, it is correct to consider the Code not as a self-regulatory instrument, but as a co-regulatory mechanism, which is not created and operated purely by the free will of industry actors but by a public body (in this case, the EU Commission) working in cooperation with industry players.

The Strengthened Code of Practice on Disinformation includes 44 commitments and 128 concrete measures in the areas of demonetization (reducing financial incentives for the disseminators of disinformation), transparency of political advertising (provisions to allow users to better identify political ads through better labeling), ensuring the integrity of services (steps against manipulative behavior such as the use of spam or disinformation), and the protection of the integrity of services (for example, measures to curb manipulative actions such as fake accounts, bot-driven amplification, impersonation and malicious deep spoofing), empowering users through media literacy initiatives, ensuring greater transparency for platforms' recommendation systems, supporting research into disinformation, and strengthening the fact-checking community. These measures will be supported by an enhanced monitoring framework, including service-level indicators to measure the implementation of the Code at EU and Member State level. Signatories submitted their first reports on the implementation of the Code to the Commission in early 2023. Thereafter, very large online platforms (as defined in the DSA) will report every

⁵⁶ Communication from the Commission on the European Democracy Action Plan.

⁵⁷ Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions; European Commission Guidance on Strengthening the Code of Practice on Disinformation. Brussels, 26.5.2021, COM(2021) 262 final.

⁵⁸ 2022 Strengthened Code of Practice on Disinformation.

⁵⁹ Communication from the Commission; Guidance on Strengthening the Code of Practice on Disinformation.

six months, while other signatories will report annually. The Strengthened Code also includes a clear commitment to work towards the establishment of structural indicators to measure the overall impact of the Code's requirements. The 2022 Strengthened Code focuses on political advertising, but also refers to other 'malicious actors' beyond those who commission political campaigns containing disinformation.⁶⁰ However, it only covers other speakers beyond the political sphere (citizens interested in public affairs and participating in debates) and misinformation without malicious intent more narrowly. Moreover, as with previous documents, it leaves the most important question open: who decides what constitutes disinformation? More precisely, it leaves the decision to the platform moderators and, to a lesser extent, to the fact-checkers.

The first baseline reports on the implementation of the Code were published in February 2023.⁶¹ According to their reports, the service providers that signed the Code have taken a number of measures: for example, Google deprived disseminators of disinformation of €13 million in advertising revenue in the third quarter of 2022, while TikTok removed 800,000 fake user accounts, followed by a total of 18 million users during the same period and, on Facebook, 28 million fact-checking tags were added to different posts.

An ongoing legislative procedure is also worth noting in this regard. In 2021, the European Parliament and the Council proposed a regulation on the transparency and targeting of political advertising.⁶² The regulation, if adopted, would be uniformly binding on all Member States, covering the identification of the customers of political advertising, their recognizability, measures against illegal political advertising and the requirements for targeting specific users.

The EU's approaches are in many respects forward-looking and can help to achieve several objectives, although they have also faced a number of criticisms. We may perceive a certain lack of sincerity on the part of both Member States and the EU when it comes to disinformation. All the related documents avoid a clear assessment of the question of whether the dissemination of disinformation falls within the scope of freedom of expression. Following the prevailing European doctrine, one cannot but conclude that a significant proportion of communications containing disinformation is content protected by freedom of expression, so their restriction by instruments with a not entirely clear legal status, such as a co-regulatory code of practice, may be cause for concern. These communications relate to matters of public interest and are therefore subject to the strongest

⁶⁰ 2022 Strengthened Code of Practice on Disinformation, pp. 8–9, 15–16, 19 and 37.

⁶¹ European Commission, 'Signatories of the Code of Practice on Disinformation Deliver Their First Baseline Reports in the Transparency Centre', 9 February 2023, <https://digital-strategy.ec.europa.eu/en/news/signatories-code-practice-disinformation-deliver-their-first-baseline-reports-transparency-centre>; for the reports, see <https://disinfocode.eu/reports-archive/?years=2023>.

⁶² Proposal for a Regulation of the European Parliament and of the Council on the Transparency and Targeting of Political Advertising. Brussels, 25.11.2021, COM(2021) 731 final.

protection of freedom of expression, with the exception of unlawful content, the publication of which is prohibited by specific rules (see Section 6.3). This also applies to content generated or sponsored by governments. However, communications involving untrue statements of fact may not be considered particularly valuable in the European approach, and could actually be restricted by the imposition of further prohibitions. In other words, Member States are free to introduce prohibitions against intentional disinformation that harms society, if this is necessary and proportionate, as this is not within the EU's competence.⁶³ Member States must take the ECtHR's case law into account when restricting freedom of expression, and this applies equally to disinformation.⁶⁴ Even so, the production and transmission of disinformation can justify restrictions on freedom of expression. However, other content beyond the sufficiently narrow prohibitions thus defined may still claim protection of freedom of expression, so measures taken against them by online platforms – based either on voluntary commitments or on the co-regulatory Code of Practice, but which are not based on the law as it stands – may be unjustified or disproportionate.⁶⁵

The EU approach also reveals a kind of hidden elitism. While the EU focuses on political advertising and intentional disinformation campaigns, some of the measures it enforces on platforms also cover misinformation and communications by citizens. Ian Cram argues that the ECtHR's jurisprudence privileges traditional (institutional) media over citizen journalists, imposing standards of 'responsible journalism' on the latter.⁶⁶ It follows from this that the obligations of the media and other speakers are, where conceptually possible, the same. According to Cram, this is a kind of elitist approach, linked to a – democratically contradictory – perception of media freedom that seeks to create an 'enlightened public opinion' even vis-à-vis 'the people' (that is, individual speakers, who may be unbridled, perhaps foul-mouthed, and may lack the resources of the institutional media to uncover reality or create informed opinions).⁶⁷ The same is true for the obligations

⁶³ Donato Vese, 'Governing Fake News: The Regulation of Social Media and the Right to Freedom of Expression in the Era of Emergency' (2022) 13 *European Journal of Risk Regulation* 477–513.

⁶⁴ Sardo, 'Categories, Balancing, and Fake News'; Ethan Shattock, 'Fake News in Strasbourg: Electoral Disinformation and Freedom of Expression in the European Court of Human Rights (ECtHR)' (2022) 13(1) *European Journal of Law and Technology* 1–25; Paolo Cavaliere, 'The Truth in Fake News: How Disinformation Laws are Reframing the Concepts of Truth and Accuracy on Digital Platforms' (2022) 3(4) *European Convention on Human Rights Law Review* 481–523.

⁶⁵ Paolo Cavaliere also draws attention to the threat to freedom of expression, see Cavaliere, 'The Truth in Fake News' (n 64) 520–21.

⁶⁶ Ian Cram, *Citizen Journalists: Newer Media, Republican Moments and the Constitution* (Cheltenham: Edward Elgar, 2015) pp. 112–43; Ian Cram, *Liberal Democracy, Law and the Citizen Speaker: Regulating Online Speech* (Oxford: Hart, 2022) pp. 30–37, 144–86.

⁶⁷ Cram, *Liberal Democracy* (n 66) p. 30. Jacob Rowbottom points to a similar problem, arguing that the ECtHR 'tends to protect speech that is deemed to be of "high value", and therefore

imposed on platforms, which ultimately also restrict this kind of citizen participation in the public sphere. The EU thus turns to the ‘elite’ of the public arena, namely to the traditional media and fact-checking organizations, for help in judging disinformation.

The lack of honesty is also reflected in the interpretation of the Code of Practice, a formally self-regulatory instrument, which in reality is co-regulation imposed by the EU,⁶⁸ where coercion is not based on legislation but on informal agreements, and accompanied by concerns on the part of service providers about the risk of stricter regulation in the future. This co-regulatory nature is recognized by the reference in the Preamble of the Code: ‘This Code of Practice aims to become a Code of Conduct under Article 35 of the DSA’⁶⁹ (in this section, the DSA itself advocates the creation of codes of conduct that set out the various obligations of platforms). Of course, the concerns of service providers are not necessarily justified, given their economic interest in the spread of disinformation, as the 2021 leak by a former Facebook employee, Francis Haugen, starkly highlighted.⁷⁰ Disinformation, unfortunately, tends to attract users, who readily consume such content and interact with it heavily, which in turn generates financial benefits for the platforms. It is therefore also difficult to believe that the transparency required by the EU and committed to by the service providers in relation to the spread of disinformation – covering decision-making and all relevant information – will actually be achieved, and it is very difficult for an actor outside the platform to verify whether it has been.

Twitter announced in May 2023, under the leadership of Elon Musk, that it would leave the Code. Because of its formally self-regulatory nature, this was, of course, within its rights. In any case, Thierry Breton, a senior official of the European Commission, announced immediately after the decision that the Code would nevertheless be enforced, including against Twitter.⁷¹ This will be possible indirectly, if the Code becomes a kind of industry standard, and thus effectively binding, by applying Section 35 of the DSA.

A problem that goes hand in hand with the spread of disinformation is the breakdown of traditional media. The media are gradually losing the trust of the public,⁷² but their economic foundations and, in turn, their professionalism are also

does little to protect much internet content’. See Jacob Rowbottom, ‘To Rant, Vent and Converse: Protecting Low Level Digital Speech’ (2012) 71(2) *Cambridge Law Journal* 355–83.

⁶⁸ Cavaliere, ‘The Truth in Fake News’ (n 64) 490.

⁶⁹ 2022 Strengthened Code of Practice on Disinformation, I. Preamble, para. (i).

⁷⁰ Scott Pelley, ‘Whistleblower: Facebook Is Misleading the Public on Progress against Hate Speech, Violence, Misinformation’, CBS News, 4 October 2021, www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021-10-03/; Stasi and Parcu, ‘Disinformation and Misinformation’ (n 12) p. 410.

⁷¹ ‘Twitter Withdraws from EU Disinformation Code, Commissioner Says’, *Time*, 27 May 2023, <https://time.com/6283183/twitter-withdraws-from-eu-disinformation-code-commissioner-says>.

⁷² Megan Brennan, ‘Americans’ Trust in Media Remains Near Record Low’, Gallup, 18 October 2022, <https://news.gallup.com/poll/403166/americans-trust-media-remains-near-record-low.aspx>.

under threat, not least because of the proliferation of internet services. Some EU documents mention the role and importance of the traditional media, although they can hardly offer solutions to these problems. Similarly, only at the level of a mere mention does the EU, including the DSA, address the issue of filter bubbles,⁷³ which reinforce social fragmentation, such as the ‘Daily Me’ content offer,⁷⁴ customized for each user, which contributes significantly to the spread of disinformation among susceptible users.⁷⁵ It would not be inconceivable to adopt some of the approaches taken in the regulation of traditional media, such as the right of reply, which would allow disinformation to be accompanied immediately by a reaction containing true facts, or an appropriate adaptation of the obligation of balanced coverage, which would allow a controversial issue to be presented in several readings, immediately visible to the user. This is also hinted at in the Code of Practice, which seeks to steer users towards reliable sources. Measure 22 (7) of the Code states that ‘Relevant Signatories will design and apply products and features (for instance, information panels, banners, pop-ups, maps and prompts, trustworthiness indicators) that lead users to authoritative sources on topics of particular public and societal interest or in crisis situations.’ The right to information from multiple sources is the objective of both the right of reply and the obligation to provide balanced information, meaning that even if the means differ, the objectives may be similar in the regulation of traditional media and platforms.

Finally, another problem with the EU’s approach that has been identified is that it is up to platforms and fact-checkers to judge content in the fight against disinformation. This is understandable, since the EU did not want to set up a kind of Orwellian Ministry of Truth, as it would consider it incompatible with freedom of expression for state bodies, courts and authorities to decide on the veracity of a claim. However, it is also doubtful whether leaving such decisions up to private individuals is capable of facilitating informed, fair and unbiased decision-making and whether it does not itself pose a threat to freedom of expression. The very term ‘fact-checking’ is unfortunately Orwellian, and the fact-checkers – and the platform moderators – can themselves be biased, as well as wrong.⁷⁶ Human cognitive mechanisms themselves make fact-checking difficult,⁷⁷ and its credibility is easily

⁷³ Eli Pariser, *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think* (London: Penguin, 2011).

⁷⁴ Nicholas Negroponte, *Being Digital* (New York: Alfred A. Knopf, 1995).

⁷⁵ The 2022 Code of Practice only requires platforms to provide information on recommendation systems, although, under the DSA, users will be able to prohibit the operation of recommendation systems based on profiling by changing their settings (see Section 6.5.2).

⁷⁶ Otávio Vinhas and Marco T. Bastos, ‘Fact-Checking Misinformation: Eight Notes on Consensus Reality’ (2022) 23(4) *Journalism Studies* 448–68; ‘6 Ways Fact Checkers Are Biased’, AllSides, 23 February 2022, www.allsides.com/blog/6-ways-fact-checkers-are-biased.

⁷⁷ Sungkyu Park et al., ‘The Presence of Unexpected Biases in Online Fact-Checking’ (2021) 2(1) *Harvard Kennedy School Misinformation Review*, https://misinforeview.hks.harvard.edu/wp-content/uploads/2021/01/park_unexpected_biases_online_fact_checking_20210127.pdf.

undermined, as ‘fact-checkers ... disagree more often than one might suppose, particularly when politicians craft language to be ambiguous’.⁷⁸ An empirical study found that ‘fact-checkers are both less likely to fact-check ideologically close entities and more likely to agree with them’.⁷⁹ Fact-checkers are not accountable to society, even less so than the traditional media (through legal regulation or ethics-based self-regulation). Their activities are neither necessarily transparent, nor do they have guarantees of independence. In many cases, such as EU-funded organizations, they operate using public money, which makes these shortcomings problematic. If the traditional media are increasingly losing people’s trust, what reason would people have to trust fact-checking organizations, which face similar credibility problems? While fact-checkers share similar problems with traditional media, their emergence is an interesting development and, if they can bridge the institutional problems, it is not inconceivable that they could be a useful contributor to the public sphere.⁸⁰ It is noteworthy that those fact-checkers who work on behalf of or with the approval or support of social media platforms, and who check the veracity of users’ posts on those sites, bring social media closer to traditional media in terms of the way they operate, as these verifiers have a specific editorial role.

6.6.2 *Banning Russian Media Outlets in the Context of the Russian–Ukrainian War*

Shortly after the outbreak of the Russian–Ukrainian war, on 1 March 2022, the Council of the EU adopted a Decision⁸¹ pursuant to Article 29 of the Treaty of the European Union (TEU) and a Regulation⁸² pursuant to Article 215 of the Treaty on the Functioning of the European Union (TFEU) under which it is prohibited for:

operators to broadcast or to enable, facilitate or otherwise contribute to broadcast, any content by the legal persons, entities or bodies listed in Annex XV [RT – Russia Today English, RT – Russia Today UK, RT – Russia Today Germany, RT – Russia Today France, RT – Russia Today Spanish, and Sputnik news agency], including through transmission or distribution by any means such as cable, satellite, IP-TV, internet service providers, internet video-sharing platforms or applications, whether new or pre-installed.

(Article 1(1))

⁷⁸ Chloe Lim, ‘Checking How Fact-Checkers Check’ (2018) 5(3) *Research & Politics* 1–7.

⁷⁹ Charles Louis-Sidois, ‘Both Judge and Party? An Analysis of the Political Leaning of Fact-Checkers’ (2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4030887.

⁸⁰ See Michael M. Epstein, ‘Sustaining Country-Specific Fact-Checking Remedies, the Sierra Leone Experience’ (manuscript, on file with author).

⁸¹ Council Decision (CFSP) 2022/351 of 1 March 2022 Amending Decision 2014/512/CFSP Concerning Restrictive Measures in View of Russia’s Actions Destabilising the Situation in Ukraine.

⁸² Council Regulation (EU) 2022/350 of 1 March 2022.

All broadcasting licences or authorization, transmission and distribution arrangements with RT and Sputnik were suspended. (Later, these measures were extended to other Russian media outlets.) These sanctioning rules derive directly from the TEU. The Council of the EU used the prerogatives under Title V of the TEU concerning the general provisions on the EU's External Action and the specific provisions on the Common Foreign and Security Policy.⁸³ According to a leaked letter, the Regulation should be applied to any links to the internet sites of the media outlets, as well as to their social media accounts.⁸⁴ As a result, the ban is a departure from the general monitoring ban in Article 15 of the E-Commerce Directive.⁸⁵ This provision makes it clear that any state-imposed orders on social media platforms (referred to in the Directive as host services) to monitor users' content are not compatible with European law. Later, a lawsuit was initiated by RT France against the Regulation, but the Court of Justice of the EU dismissed RT France's application.⁸⁶

According to the Recitals of the Decision and the Regulation, the Russian Federation 'has engaged in a systematic, international campaign of media manipulation and distortion of facts in order to enhance its strategy of destabilization of its neighboring countries and of the Union and its Member States'.⁸⁷ The recitals indicate two reasons for the ban: disinformation and propaganda.⁸⁸ Under Article 52 (1) of the CFR, any such interference must pursue 'objectives of general interest recognized by the Union'. Considering this, the restriction targeting disinformation and propaganda might be in line with the CFR.⁸⁹ However, according to Baade, the EU should not invoke the prohibition of disinformation or propaganda as a legitimate aim, as they may be protected expressions. An alternative aim would be to stop propaganda for war specifically.⁹⁰ The prohibition of propaganda for war is enshrined in Article 20 of the International Covenant on Civil and Political Rights. As all the EU Member States have ratified the Covenant, this prohibition

⁸³ Francisco J. Cabrera Blázquez, *The Implementation of EU Sanctions against RT and Sputnik* (Strasbourg: European Audiovisual Observatory, 2022), <https://rm.coe.int/note-rt-sputnik/1680a5dd5d>.

⁸⁴ A letter from the European Council to undisclosed recipients, see https://lumendatabase.org/file_uploads/files/5061360/005/061/360/original/6-9267000032260.pdf?1646430529&access_token=w6YC_Vjutt85UvZMLExMag, p. 2.

⁸⁵ Directive 2000/31/EC on Electronic Commerce.

⁸⁶ Case T-125/22, *RT France v. Council of the European Union*, judgment of the General Court (Grand Chamber) on 27 July 2022.

⁸⁷ Council Regulation (EU) 2022/350 of 1 March 2022, Recital 6; Council Decision (CFSP) 2022/351 of 1 March 2022 Recital 6.

⁸⁸ Council Regulation (EU) 2022/350 of 1 March 2022, Recitals 3–10; Council Decision (CFSP) 2022/351 of 1 March 2022, Recitals 4–6 and 10.

⁸⁹ Igor Popović, 'The EU Ban of RT and Sputnik: Concerns Regarding Freedom of Expression', *EJIL: Talk!*, 30 March 2022, www.ejiltalk.org/the-eu-ban-of-rt-and-sputnik-concerns-regarding-freedom-of-expression.

⁹⁰ Baade, 'Don't Call a Spade a Shovel' (n 8).

can also be considered a generally accepted principle of EU law. As Baade notes, the justification for the ban imposed on RT and Sputnik in the current situation cannot be based solely on the character of their content as ‘propaganda’ and not even as disinformation.⁹¹ As already mentioned, propaganda and disinformation are generally protected by the freedom of expression, with certain exceptions.

After the Regulation came into force, the largest social media companies relaxed the enforcement of their rules involving threats against Russian military personnel in Ukraine.⁹² According to a leaked internal letter, Meta allowed Facebook and Instagram users to call for violence against the Russian and Belarusian leaders, Vladimir Putin and Alexander Lukashenko, so long as the violence was nonspecific (without referring to an actual plot), as well as violence against Russian soldiers (except prisoners of war) in the context of the Ukraine invasion, which involves a limited and temporary change to its hate speech policy.⁹³ Twitter also announced some changes in its policies related to the war, although the company did not amend its generally applicable hate speech policies.⁹⁴

The right of platforms to change the boundaries of free speech at will, without any constitutional guarantee or supervision, is an extremely dangerous development. Their propensity to make changes in a less transparent way, avoiding any meaningful public debate on the proposed changes, only increases the risk to freedom of expression.

6.7 ATTEMPTS TO REGULATE DISINFORMATION AT THE NATIONAL LEVEL

In order to strengthen the obligations of online platforms, some European countries have adopted rules, in line with common European law, to compel platforms to remove illegal content more quickly and effectively. The corresponding Act in German law (effective as of 1 January 2018) is a paramount example of this trend.⁹⁵ According to the applicable provisions, all platform providers within the scope of the Act (that is, platform providers with over 2 million users from Germany) must

⁹¹ Björnstjern Baade, ‘The EU’s “Ban” of RT and Sputnik: A Lawful Measure against Propaganda for War’, *Verfassungsblog*, 8 March 2022, <https://verfassungsblog.de/the-eus-ban-of-rt-and-sputnik>.

⁹² Munsif Vengattil and Elizabeth Culliford, ‘Facebook Allows War Posts Urging Violence against Russian Invaders’, *Reuters*, 10 March 2022, www.reuters.com/world/europe/exclusive-facebook-instagram-temporarily-allow-calls-violence-against-russians-2022-03-10.

⁹³ Emerson T. Brooking, ‘Meta Meets the Reality of War’, *SLATE*, 17 March 2022, <https://slate.com/technology/2022/03/meta-facebook-calls-violence-invading-russians.html>.

⁹⁴ Sinéad McSweeney, ‘Our Ongoing Approach to the War in Ukraine’, *Twitter*, 16 March 2022, https://blog.twitter.com/en_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine.

⁹⁵ Act to Improve Enforcement of the Law in Social Networks 2017 (Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz)), Art. 1 G. v 01.09.2017 BGBl. I S. 3352 (Nr. 61).

remove all user content that commits certain criminal offences specified by the Act. Such offences include defamation, incitement to hatred, denial of the Holocaust and the spreading of scaremongering news stories.⁹⁶ Manifestly unlawful pieces of content must be removed within twenty-four hours after receipt of a notice, while any 'ordinary' unlawful content must be removed within seven days.⁹⁷ If a platform fails to remove a given piece of content, it may be subject to a fine of up to €50 million (theoretically, in cases of severe and multiple violations).⁹⁸ The German legislation does not go much further than the E-commerce Directive itself, or its successor, the DSA; it simply refines the provisions of the Directive, lays down the applicable procedural rules and sets harsh sanctions for platforms which violate them. Nonetheless, the rules are followed in practice, and Facebook seems eager to perform its obligation to remove objectionable content.⁹⁹ The German regulation shows how difficult it is to apply general pieces of legislation and platform-specific rules simultaneously, and it demonstrates how governments prefer to have social media platforms act as the judges of user-generated content.

Subsequently, France¹⁰⁰ and Austria¹⁰¹ adopted similar rules, although the French law ('Avia law') was annulled by the Constitutional Council because some of its provisions did not meet the constitutional requirements.¹⁰² France had introduced transparency and reporting obligations for platforms in a law adopted in 2018, prior to the Avia law, along with a fast-track judicial procedure to remove content disseminated during election campaigns and deemed misleading or inaccurate.¹⁰³ The law confers new powers on the Media Council (Conseil Supérieur de l'Audiovisuel), such as the ability to suspend or withdraw the licence of certain media services if, for example, a service under the control or influence of a foreign company is endangering France's fundamental interests, including the proper functioning of its institutions, by transmitting false information.¹⁰⁴ Under an amendment to the 1986 Freedom of Communication Act,¹⁰⁵ the Media Council can order

⁹⁶ *Ibid.* s. 1.

⁹⁷ *Ibid.* s. 3.

⁹⁸ *Ibid.* s. 4.

⁹⁹ 'Facebook Deletes Hundreds of Posts under German Hate-Speech Law', Reuters, 27 July 2018, www.reuters.com/article/us-facebook-germany-idUSKBN1KH21L.

¹⁰⁰ Law to Combat Hateful Content on the Internet (Loi no. 2020-766 du 24 juin 2020 visant à lutter contre les contenus haineux sur internet, JORF no. 0156 du 25 juin 2020 ('Avia Law')).

¹⁰¹ Federal Law on Measures for the Protection of Users on Communication Platforms (Bundesgesetz über Maßnahmen zum Schutz der Nutzer auf Kommunikationsplattformen (Kommunikationsplattformen-Gesetz-KoPl-G)).

¹⁰² Décision no. 2020-801 DC du 18 juin 2020, see www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm.

¹⁰³ Law on the Fight against the Manipulation of Information (Loi no. 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information, JORF no. 0297 du 23 décembre 2018); Craufurd Smith, 'Fake News, French Law and Democratic Legitimacy'.

¹⁰⁴ Law on the Fight against the Manipulation of Information, Art. 8.

¹⁰⁵ Law on Freedom of Communication (Loi no. 86-1067 du 30 septembre 1986 relative à la liberté de communication ('Léotard Law')).

the suspension of the electronic distribution of a television or radio service owned or controlled by a foreign state if the company deliberately transmits false information that could call into question the integrity of an election.¹⁰⁶ These powers may be exercised from the beginning of the three months preceding the first day of the month in which the presidential, general or European Parliamentary elections or referendums are held. The Constitutional Council found this law constitutional.¹⁰⁷

The German and French attempts to regulate disinformation have introduced rules imposing obligations on platforms to remove certain content quickly. At the same time, German legislation only imposes obligations on content that is in breach of the Criminal Code, hence it is only the 2018 French law that regulates disinformation that is not in any case illegal, during election campaigns. However, these approaches still leave the decision on content in the hands of the platforms, and do not attempt to limit the spread of disinformation in general.¹⁰⁸ In Germany, another important piece of legislation has been passed, which also addresses the issue of disinformation. In 2020, the Interstate Treaty on Media Services (Medienstaatsvertrag, MStV) was adopted, which provides for the transparency of algorithms, the proper labelling of bots and the easy findability of public service media content on the platforms on which it is available. The MStV obliges social media platforms, video-sharing platforms and search engines to be nondiscriminatory in terms of content and to prioritize public service content, while not restricting user preferences. On video-sharing platforms, available public broadcasting content should be especially highlighted and made easy to find. These intermediaries may not unfairly disadvantage (directly or indirectly) or treat differently providers of journalistic editorial content to the extent that the intermediary may potentially have a significant influence on their visibility.¹⁰⁹ These rules only indirectly limit the spread of disinformation, but they provide a good example of how regulation can try to steer users towards credible content, in line with the traditional approach to media regulation.¹¹⁰

In the fight against the COVID pandemic and the disinformation related to it, several European countries tried to curb the spread of false and dangerous information by tightening criminal laws. Hungary, for example, tightened its rules on scaremongering,¹¹¹ and Greece extended the scope of the existing offence of

¹⁰⁶ Ibid. Art. 33-1-1.

¹⁰⁷ Conseil Constitutionnel, Décision no. 2018-773 de du 20 décembre 2018.

¹⁰⁸ Licia Cianci and Davide Zecca, 'Polluting the Political Discourse. What Remedies to Political Microtargeting and Disinformation in the European Constitutional Framework?' (2023) 10 *European Journal of Comparative Law and Governance* 1–46, 1.

¹⁰⁹ Interstate Media Treaty (Medienstaatsvertrag), especially § 18 Abs. 3; § 19; § 84; §§ 93 and 94.

¹¹⁰ On the national measures against disinformation, see Giovanni Pitruzella and Oreste Pollicino, *Disinformation and Hate Speech: A European Comparative Perspective* (Milan: Bocconi University Press, 2020) pp. 94–126.

¹¹¹ 2012 Criminal Code of Hungary, s. 337 (see Section 6.3).

dissemination of false information and introduced a prison sentence for those who spread disinformation on the Internet.¹¹²

8. ON THE POSSIBLE FUTURE SOLUTIONS: SOME CONCLUSIONS

The European states and the EU clearly assign primary responsibility for addressing disinformation issues to the platforms. Of course, the national governments and the European institutions have made a number of commitments themselves, but they leave it to the platforms to sort out the substantive issues, including compliance with their commitments under the Code. However, this is not a reason to give up on introducing further restrictions on free speech, as allowed by the European concept of freedom of expression. Even in the context of the US legal system, Cass Sunstein argues that intentional lies, if they cause at least moderate harm, may be constitutionally prohibited – and even negligent or mistaken misrepresentations can be restricted, if the harm incurred by them is serious.¹¹³ It is still better – at least in Europe, we typically think so – that the line between forbidden and permissible speech is drawn by the legislature and the courts, constrained by strict constitutional guarantees, rather than by private organizations (in this case, mainly social media platforms) operating without such guarantees. But each and every social media post of concern cannot be taken to court, because nowhere could the judicial system cope with such a workload. Therefore, the right of platforms to decide on user content is likely to remain necessary in the long-term future. However, the protection of content that is *not* prohibited under the regime of freedom of expression is an important consideration, even if it contains untruths.

Although the European approach is wary of considering the communication of untrue statements of fact to be of high value, freedom of expression, at least according to the traditional approach, is in a sense a ‘black and white’ issue. Either a particular piece of content falls within the scope of freedom of expression or it does not. In other words, once the sometimes-difficult question of whether a particular piece of content constitutes defamation, invasion of privacy, hate speech and so on has been successfully answered, the consequences are self-evident: the content will either be protected or it will not. ‘Pizzagate’,¹¹⁴ for example, could in

¹¹² Law 4855/2021: Amendments to the Penal Code, the Code of Criminal Procedure and other urgent provisions (Τροποποιήσεις του Ποινικού Κώδικα, του Κώδικα Ποινικής Δικονομίας και λοιπές επείγουσες διατάξεις του Υπουργείου Δικαιοσύνης), Art. 36.

¹¹³ Cass R. Sunstein, *Liars: Falsehoods and Free Speech in an Age of Deception* (Oxford: Oxford University Press, 2021) pp. 12–18, 128–30.

¹¹⁴ Amanda Robb, ‘Anatomy of a Fake News Scandal’, *Rolling Stone*, 16 November 2017, www.rollingstone.com/feature/anatomy-of-a-fake-news-scandal-125877.

principle have been dealt with under defamation law (at least if it would have happened in Europe, as under US defamation law it is more difficult to protect the reputation of a specific person against false allegations), and the false allegations made in the Brexit campaign¹¹⁵ could in principle also have been prohibited under the rules governing fair election or referendum campaigns. Of course, even in these cases, the permissibility of restricting speech is not clear-cut and requires a nuanced decision by a court. Furthermore, an otherwise patently untrue statement – for example, how much more money will be available for the National Health Service in the United Kingdom if the country leaves the EU – may not necessarily be clearly refutable in a legal proceeding. But the main point is that many untrue statements are actually protected by freedom of speech. This does not mean that the protected content has the right to reach an audience or to have its volume amplified by a particular service (for example, through the media), but rather that its restriction is not allowed. This traditional approach is being disrupted by online platforms, which, as is their general practice, also restrict content that is not legally prohibited, according to their own intentions and contractual terms. The same problem dogs the fight against (not legally prohibited) disinformation: the EU also encourages restrictions on content that is otherwise protected by freedom of expression, and the relevant documents do not attempt to resolve this contradiction.

It is also important to make a clear distinction between disinformation originating from governments and dis- or misinformation that comes from members of society, whether deliberate or in good faith, but in this respect the EU documents currently available are not fully consistent. Members of society should not be disproportionately restricted in their freedom of expression, even if they approach public debate with malicious intent, and certainly not if they are unaware of the falsity or potential for damage of the news they are spreading (the good faith transmission of government disinformation also falls into this category). Private speech controlled or promoted by the government should be taken into account and only the freedom of honest citizens' speech who are otherwise wrong should be strongly protected. The question is whether this separation is even possible. And if so, whose job is it, the legal regulators or purely the platforms? We do not have good answers to this dilemma at the moment.

Nor would it be inconceivable to regulate platforms more strictly, setting out their obligations vis-à-vis content not protected by freedom of expression, not in self- or co-regulatory instruments but in clearly prescribed legal rules. This would of course require close cooperation between Member States and the EU, as speech bans can only be imposed at Member State level, while platform regulation can only be effective at EU level.

¹¹⁵ Hannah Marshall and Alena Drieschova, 'Post-Truth Politics in the UK's Brexit Referendum' (2018) 26(3) *New Perspectives* 89–106.

Users need to be led out of the filter bubble imposed on them by the platforms, which would fundamentally affect their business model. In this regard, the provision of the option prescribed by the DSA to opt out of content recommendation based on profiling is a step in the right direction, but not big enough, because it puts the decision in the hands of the users, and it is questionable how many will take advantage of it, and that particular bubble can also be produced by means other than profiling. Data protection regulations can also be called upon to help in the fight, in particular by tightening up data-processing actions by platforms.¹¹⁶

It would be worth considering making the transmission of substantiated statements and opinions on public affairs to users mandatory, or providing easy access to divergent and dissenting views on specific issues, while maintaining the choice for users who do not wish to hear them, as exemplified by the regulation of traditional media. Such instruments include, in respect of television and radio, the right of reply and the obligation to provide balanced (impartial) news coverage, the mandatory publication of local or national content or the mandatory transmission of certain content of public interest by broadcasters. These duties could also be applied to social media, with some adaptation. In principle, social media could be required to make available, alongside a post on a contentious issue, posts that present the dissenting views on that issue. Algorithms might be able to do this, although the business model of the platforms might be adversely affected. Such a rule would be similar to the right-of-reply and impartial information obligations known from media regulation, except that it could be done automatically without a specific request. Strengthening nonlegislative approaches, raising awareness and supporting traditional media are also necessary tools – within the competence of Member States.

The fight against disinformation is a seemingly open-ended task that poses particular challenges for policymakers, both in terms of protecting freedom of expression and in defining new obligations for members of the public. It has become clear that traditional legal instruments, legislation and the imposition and enforcement of obligations by the relevant authorities can only partially address the problems it raises, and that the cooperation of all stakeholders is necessary. However, this should not lead to the ‘outsourcing’ of decisions by putting them fully in the hands of private companies. Member States and the EU must continue to play a leading role in shaping the rules. The EU has taken a number of important measures, and some Member States are trying to address some of the issues, but it is reasonable to fear that we are only at the beginning of the journey and that further technological

¹¹⁶ Joris van Hoboken and Ronan Ó Fathaigh, ‘Regulating Disinformation in Europe: Implications for Speech and Privacy’ (2021) 6(9) *UC Irvine Journal of International, Transnational, and Comparative Law* 9–36.

developments will bring new risks. Disinformation, as Paul Bernal has so eloquently demonstrated,¹¹⁷ is essentially the same age as public communication; there is nothing new under the sun, but we must be able to formulate new answers to old questions all the time. But the end result of any struggle of legal systems in this regard will be that responsible, informed participation in public debates will remain primarily the responsibility of the individual concerned, just as it has been in past centuries.

¹¹⁷ Paul Bernal, *The Internet, Warts and All: Free Speech, Privacy and Truth* (Cambridge: Cambridge University Press, 2018) pp. 230–34; Paul Bernal, 'Facebook: Why Facebook Makes the Fake News Problem Inevitable' (2018) 69(4) *North Ireland Legal Quarterly* 513–30, 516–19.