

Disinformation in the Perspective of Media Pluralism in Europe – the role of platforms¹

Elda Brogi, Konrad Bleyer-Simon

Abstract: Media freedom and media pluralism are recognised as pillars of contemporary democracies. Technological advancements have not only created new opportunities to boost media freedom and media plurality, but prompted new sources of risks. One of them is the scale and the impact of disinformation on public opinion. Although not illegal, it may pose a growing threat, for instance, to the integrity of elections, including manipulation, as well as to efforts to respond to the COVID-19 pandemic. This paper aims to provide an overview of problems and opportunities, based on experiences from projects the authors are involved in. It provides a brief overview of multi-country data coming from country experts of the EU-wide Media Pluralism Monitor data collection, and describes the challenges and opportunities of European measures to fight disinformation, based on the work of the European Digital Media Observatory.

Keywords: disinformation, misinformation, Code of Practice on Disinformation, European Union, EU, platform regulation, EDMO, Media Pluralism Monitor

Chapter 1. Introduction

In the past years, disinformation has become another challenging issue in content moderation online, next to hate speech; although it is often not illegal, it can cause public harm. This led to the rethinking and re-interpretation of the rationale of the liability exemption for online platforms, as well as the governance of the digital environment.

The policy discussion is ongoing at both the EU and the member state level. In Germany, the Bundestag passed the so-called Network Enforcement Act (NetzDG, also referred to as the Facebook Act) in 2017, which

1 The opinions and views expressed in this chapter are those of the Authors.

required social media providers to proactively remove certain types of criminal content. This requirement was criticised early on by civil society for possibly damaging freedom of expression and freedom of the press. Many other EU member states decided to also react to the problem in one way or another in their national regulatory systems. In the EU, a key policy tool is the Code of Practice on Disinformation (the Code), which brought together online platforms, such as Google, Facebook, Twitter and TikTok, as well as a number of other stakeholders in an initial effort to fight disinformation in the context of a self-regulatory framework, applying in the framework of existing laws, including the e-Commerce Directive 2000/31/EC, with specific reference to articles 12 to 15 on exemption of liability.²

While both the EU members and the Commission took interesting steps in defining a policy against disinformation, for now, there has been limited concrete progress in defining an effective governance strategy. This is in part due to a lack of understanding of the criteria used by online platforms in their content moderation and the design of their recommendation systems. In addition, considering the trends in EU regulation, there is no defined methodology to assess how and with which consequences the platforms act in order to limit the spread of disinformation.

In this paper, the Authors reflect on the measures taken in order to tackle disinformation at EU level, starting from the results of the Media Pluralism Monitor project, and following with the early research carried out under the newly established European Digital Media Observatory (EDMO). The latter multi-disciplinary project brings together fact-checkers, media literacy experts and academic researchers with the aim of understanding and analysing the disinformation phenomenon. The Authors' contribution in EDMO focuses on research and policy analysis on disinformation, including suggestions on a methodology to assess to what extent the Code's implementation impacts the overall disinformation phenomenon. The Authors describe the current shortcomings of the Code, which have a lot to do with the text's lack of detailed practical guidance for its signatories. To help overcome this deficit, the Commission has announced its intent to transform the Code into a co-regulatory instrument. As already visible in the Commission's guidelines for strengthening the

2 European Commission, "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 2000/31/EC", COM(2020) 825 final, December 15, 2020. <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031>.

Code, compliance with the commitments would then be assessed with the help of well-defined key performance indicators (KPIs) at service-level, while the overall impact of the Code would be assessed with a set of structural KPIs. Finally, we focus on the issue of “trustworthiness” as a key attribute of online contents and content publishers, which can help guide online platforms in their efforts to improve the health and the plurality of online media landscapes. But here again, current measures are lagging behind, while current initiatives that work on providing trustworthiness indicators for platforms may risk creating unintended side-effects to media pluralism.

Chapter 2. Disinformation and the threat to media pluralism

In light of the challenges posed to European democracy by the spread of disinformation, the European Commission expressed the need for a pan-European response, and in January 2018 established the High Level Expert Group on Fake News.³ This group was made up of industry representatives, civil society, policy makers and scholars, aiming to provide advice on policy initiatives to tackle the problems of online disinformation on the European level. It produced a report in March of the same year, which recommended a multidimensional approach to increase the transparency of online news, the promotion of media literacy, the development of tools to empower users, to safeguard the diversity and sustainability of the news ecosystem in Europe, as well as to promote research on the issue of disinformation.⁴ Ahead of the 2019 European elections, the EU followed up by sponsoring a “European approach”⁵ to tackle disinformation. This led to the signing of the Code of Practice on Disinformation (the Code),⁶ the first major initiative developed at EU level to fight disinformation.

3 It was later renamed to High Level Expert Group on Fake News and Online Disinformation.

4 High Level Expert Group on Fake News and Online Disinformation, “A multi-dimensional approach to disinformation - Report of the independent High level Group on fake news and online disinformation”, European Commission, 2018, <https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation>.

5 European Commission, “Tackling online disinformation: a European Approach” COM/2018/236 final, 2018. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=C_ELEX%3A52018DC0236.

6 European Commission, “Code of Practice on Disinformation”, 2018. <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>.

It followed the Expert Group's recommendations and encouraged online platforms to self-regulate, ensure the transparency of political advertising and restrict the automated spread of disinformation in the European Economic Area.

In parallel with these developments, the Media Pluralism Monitor⁷, an EU-wide data collection administered by the Centre for Media Pluralism and Media Freedom (CMPF) at the European University Institute (EUI), has tried to measure the risks for media pluralism stemming from disinformation. Overall, the findings of these assessments show that disinformation is seen as a serious threat all over the EU, but the debate around disinformation and its regulation is still in its early phases; there is a need to find a common language and common policies, compliant with the rule of law.

Looking at the years 2018-2019, the CMPF has introduced elements in its questionnaire that investigate the transparency of online political advertising. The variables aim to define the role and the limits of online platforms' activity, as well as the procedures for their accountability when dealing with political content online. Opacity in political advertising is seen as a key enabler of the rapid spread of disinformation. Results of the MPM2020 sub-indicator "rules on political advertising online" show, as expected, considering the novelty of this debate, that in 25 countries, parties and candidates were not fully transparent about the spending and methods they used in their social media campaigns. In 18 countries some issues were noted in relation to the implementation of the Code, with regard to clear labelling and registering of political and issue-based advertising, and in terms of indicating who paid for it. Overall, there is very little regulation of political advertising online, largely due to a lack of understanding of the criteria used by online platforms in content moderation and the design of recommendation systems.⁸ Not a lot has changed in the year 2020, when the CMPF asked again its country teams about

7 The Media Pluralism Monitor (MPM) is a tool that was developed by the CMPF to assess the risks for media pluralism in a given country. It is based on the prototype of the MPM that was designed by the 2009 Independent Study on Indicators for Media Pluralism in the Member States – Towards a Risk-Based Approach carried out by KU Leuven, JIBS, CEU, Ernst & Young, and a team of national experts, https://ec.europa.eu/information_society/media_taskforce/doc/pluralism/pfr_report.pdf.

8 Elda Brogi, Roberta Carlini, Iva Nenadic, Pier Luigi Parcu and Mario Viola de Azevedo Cunha, *Monitoring media pluralism in the digital era: Report 2020* (Florence: European University Institute, 2020), <https://cadmus.eui.eu/bitstream/handle/1814/67828/MPM2020-PolicyReport.pdf;sequence=5&isAllowed=y>, 77-81.

the assessment of the situation. Still 17 EU countries saw problems related to the implementation of the Code and its effectiveness in the national context.⁹

Chapter 3. The Code of Practice on Disinformation

The Code of Practice on Disinformation was the first major initiative developed at EU level to define a policy on disinformation online, within the current EU legislative framework. In its text, disinformation is understood as 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 (in line with the definition provided by the Commission Communication on tackling on-line disinformation, 2018).¹⁰ The text adds that deceptive content is disseminated either for economic gain (monetisation) or with the intent of deceiving the public. It also emphasises the component of “public harm” as disinformation is a threat “to democratic political and policymaking processes as well as public goods such as the protection of EU citizens’ health, the environment or security”.¹¹ The Commission’s guidelines for a new, strengthened version of the Code, which were published on 26. May 2021, just a few days before submitting our revised manuscript, extend this definition to include some forms of misinformation¹² as well, meaning harmful content that is spread unintentionally.¹³

9 Elda Brogi et al. *Monitoring media pluralism in the digital era: Report 2021* (Florence: European University Institute, forthcoming 2021).

10 European Commission, *Code of Practice*; European Commission, *Tackling online disinformation*.

11 European Commission, *Code of Practice*.

12 Claire Wardle and Hossein Derakhshan differentiate between three key forms of information disorders: misinformation (when the information is not true, but it is not created and shared with the intent of doing harm), disinformation (when the untrue content was created and shared with the intent of doing harm) and malinformation (when the information is factually true, but it is shared in a way that it can cause harm). See: Claire Wardle and Hossein Derakhshan, “Thinking about ‘information disorder’: formats of misinformation, disinformation, and mal-information,” in *Journalism, ‘Fake News’ & Disinformation. Handbook for Journalism Education and Training*, eds. Cherilyn Ireton and Julie Posetti, 44-56. Paris: UNESCO, 2018. <https://unesdoc.unesco.org/ark:/48223/pf0000265552/PDF/265552eng.pdf.multi>

13 “European Commission Guidance on Strengthening the Code of Practice on

The Code provides commitments to the online platforms that sign up to it in five main areas (which are referred to as the five pillars):

- A. “Scrutiny of ad placements” is about preventing providers of disinformation from monetising their content (“reduce revenues”), mainly by urging online platforms to exclude them from their advertising services.
- B. “Political advertising and issue-based advertising” aims to make sure that political advertisement can be clearly identified by users. For this aim, users should also have an understanding of why they were targeted by a given advertisement.
- C. “Integrity of services” refers to two linked issues, fake accounts and automated bots, that play a key role in the spread of disinformation. The text calls for effective efforts to close fake accounts and to publicly issue policies on what “constitutes impermissible use of automated systems”.
- D. “Empowering consumers” aims at making access to “trustworthy” sources of information easier, by nudging consumers to access sources that are less likely to spread disinformation, as well as by providing them with tools to reliably assess the credibility of sources and content, and easily report those that spread disinformation.
- E. “Empowering the research community” affirms signatories’ willingness to allow research on disinformation and political advertisement on their platforms, and support efforts to track disinformation by giving some access to “privacy protected datasets” and supporting joint projects.

The Code was first signed by advertisers, the software developer Mozilla, as well as the online platforms Facebook, Google and Twitter in October 2018. Microsoft joined in May 2019, while TikTok signed the Code in June 2020. After agreeing on the Code, signatories have pledged to report on the actions taken in order to further the goals that were identified.¹⁴

Disinformation” European Commission, 4-5, <https://ec.europa.eu/newsroom/dae/redirection/document/76495>

14 See: “Annual self-assessment reports of signatories to the Code of Practice on Disinformation 2019,” European Commission, <https://digital-strategy.ec.europa.eu/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019>.

Chapter 4. Limited impact

The Code of Practice is an interesting step in defining a policy against disinformation, as its signatories have adhered to self-commitments that currently are not required from them by law, while its implementation and assessment can be read as a pilot test of the Digital Services Act, a 2020 proposal to update the EU's legal framework with a safer digital space in mind.¹⁵ The impact of the Code is nevertheless limited, for the time being. Problems can be traced back to three groups of issues: lack of guidance, limited compliance and the small number of signatories.

First, the Code does not provide detailed practical guidance for its signatories, just a set of vaguely defined commitments that the platforms are expected to achieve with whatever means they see fit, defining, in the end, some potentially good practices. Terms used in the commitments can be either misinterpreted, or they provide grounds for platforms to selectively comply with their obligations. Many of them lack proper definitions. For example, the commitments are aimed at “Relevant Signatories”, but who the relevant signatories are is not specified, and platforms thus have the freedom to determine for themselves what commitments they will comply with. In addition, signatories are expected to use “commercially reasonable efforts” – but this criterion is not detailed either. As some of the platforms were financially profiting from the activities of purveyors of disinformation, in their case it is not necessarily “commercially reasonable” to give up revenues. It is against this background that the Digital Services Act (DSA) proposal announced the establishment of a powerful framework for transparency and clear accountability, which enables democratic oversight over online platforms.¹⁶ The Commission's guidance would turn the Code of Practice into a “Code of Conduct” and would only allow signatories to opt out of commitments in case they provide relevant and public justification.

-
- 15 Elda Brogi and Iva Nenadic, European plan to increase transparency and accountability of the gatekeeper online platforms to protect democracy: EDMO's role in the Commission's digital policy approach <http://www.medialaws.eu/european-plan-to-increase-transparency-and-accountability-of-the-gatekeeper-online-platforms-to-protect-democracy-edmos-role-in-the-commissions-digital-policy-approach/>
- 16 European Commission, “The Digital Services Act package”, 2020. <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>.

Secondly, the European Regulators Group for Audiovisual Media Services (ERGA)¹⁷ has pointed to problems related to compliance and transparency, as the Code relies on self-reporting, and thus statements of platforms cannot be verified. There is an absence of standards for its evaluation and for reporting, lack of oversight on compliance, lack of sanctions for non-compliance, and lack of data against which to check the statements and reports created by platforms themselves.¹⁸ In fact, ERGA has found in its cooperation with national regulators that the reported achievements of platforms are not as successful as the platforms themselves make them sound.¹⁹ The Commission itself highlighted the most serious deficiencies in the attempts to demonetise purveyors of disinformation. Thus, the new Guidelines ask for a common reporting template and a set of key performance indicators (KPIs) to more effectively measure signatories' compliance.

Thirdly, the number of signatories is small. Although the initial group of signatories was extended, among others by TikTok, there are still many online platforms missing. A debate is open on whether messaging platforms should sign the Code²⁰, as experience from the past years shows that messaging services such as Messenger, Telegram or WhatsApp are among the amplifiers of the spread of disinformation content.²¹ The new guidelines would address this by introducing different reporting requirements for small and large signatories, depending on their market share in Europe; and encouraging private messaging services as well as representatives of the advertising sector to join.

17 ERGA, "ERGA Report on Disinformation: Assessment of the Implementation of the Code of Practice", 2020. <https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf>.

18 Iva Nenadić, "Unpacking the 'European approach' to tackling challenges of disinformation and political manipulation," *Internet Policy Review* 8, No. 4 (2019): 1-22.

19 ERGA, "ERGA Report on Disinformation", 17-18.

20 Messaging apps are aimed for personal communication, as they entail a sender and a recipient for a private message. Nonetheless, they are currently used to convey content to big groups of recipients and offer easy functions to share messages from one group to another, thus reaching a mass audience.

21 Samuel Woolley, "Encrypted messaging apps are the future of propaganda," *Brookings*, May 1, 2020. <https://www.brookings.edu/techstream/encrypted-messaging-apps-are-the-future-of-propaganda/>.

Chapter 5. Some suggestions to address the shortcomings

Continuous monitoring of platforms' compliance and the independent assessment of their activities to limit disinformation are key to the success of Code of Practice and are part of a new regulatory toolbox that is somehow already sketched by the proposal of the DSA and the guideline for the Code 2.0. It is important, therefore that a particular focus is devoted to the way in which an independent assessment of the Code's implementation can be done, looking at what standards to use when evaluating the compliance of the platforms with the Code's obligations, and the effects of the Code implementation, and what kind of governance to foresee in order to create an oversight mechanism that is effective and respectful of the rule of law. This is the focus of the research carried out (amongst many other activities) under the EDMO project. It aims at contributing to the definition of an assessment methodology that includes standards for platforms' reporting that enable the verification of platforms' compliance with the measures taken when implementing the Code. This methodology will be complemented with the definition of indicators that enable assessing the Code's impact in limiting the spread of disinformation and on the health of the digital information environment.

To enable a comprehensive evaluation, the Code's overall methodology designed encompasses (a) a service-level and (b) a structural assessment. The first assessment looks at platforms' compliance while the second one is interested in the Code's wider impact. What we consider reasonable is to develop and test a methodology that is: inclusive (considering current and potential future signatories of the Code); feasible (capable of being implemented on a regular basis under different forms of regulatory regime); mixed-methods based (combining quantitative and qualitative indicators); and data informed (relying on an increased transparency of platforms and functional data access).

At the time of writing, we cannot provide a complete list of service-level and structural KPIs. However, we find it important to emphasise that indicators and KPIs should be phrased in a way that prevents platforms to arbitrarily (re)interpret the questions. For this reason, we propose that each indicator be framed as a question and be complemented by a set of clearly defined guidelines. While answering the questionnaire, signatories should provide exact numbers related to the measures they have taken. This includes, among others, reporting on content sources removed or suspended due to being identified as untrustworthy by platforms or by fact-checkers (with detailed information on removals, suspensions, the length of suspensions, the number of reinstated accounts, number of

relapsing accounts), as well as information on the number of accounts reported by users and fact-checkers, the number of cases acted on, and number of complaints found justified (with a breakdown of reasons and grounds of intervention). This degree of detail is important to have a clear understanding of the extent of the problem and the exact nature of efforts taken by signatories. In the case of structural indicators, we suggest the use of audience samples to have a clear understanding of users' consumption of untrustworthy sources of information, while the methodology also should rely on the input of civil society members, fact checkers and other stakeholders who should provide additional qualitative and quantitative data in line with their expertise. The assessment of the impact of the Code implementation on the disinformation phenomenon should also rely on an analysis of the legal, economic, political and social context in which the Code has an effect. This allows us to consider all the agents that potentially could affect the spread of disinformation in a specific national media environment.

Chapter 6. Trustworthiness as a feature of the online information environment?

When signing the Code, Google, Facebook, TikTok and other signatories have committed to make changes to their algorithms based on so-called “trustworthiness indicators” which would reduce the risk that users get misled by shifty content. Thus, a first and key focus in our work under EDMO was the analysis of what could constitute the possible indicators that would allow online service providers to prioritise content that is informative and not likely to mislead or deceive users. The Code assigns great importance to the term “trustworthiness” when it comes to signatories' commitments. Pillar A (scrutiny of ad placements) highlights the importance of indicators of trustworthiness when identifying the sites where advertisement can be placed without (unintentionally) monetising purveyors of disinformation; and Pillar D (empowering consumers) mentions indicators of trustworthiness as the basis of content prioritisation and media literacy measures.

In the Code, the term “trustworthiness” refers first and foremost to content sources, and is often mentioned in connection with ownership transparency and the “verified identity” of content creators.²² Indicators of trustworthiness are expected to provide the basis for platforms for im-

22 European Commission, “Code of Practice”.

proving the findability of trustworthy content sources and “diluting” visibility (downranking) of their non-trustworthy counterparts.²³ “Ranking”, “prioritising” or “pushing up” trustworthy content is often mentioned in related documents and assessments as the method that makes the best use of the indicators.²⁴

As such, in the current context, we define “trustworthiness” as a term that refers to the source or publisher of a piece of information. A publisher of information can be regarded as trustworthy (or credible) when the users’ chance of being exposed to false or misleading content (dis- or misinformation) by that source is relatively low. Moreover, it is expected that a trustworthy publisher has a procedure in place to make sufficient and timely corrections, for the case that it publishes false or misleading content. A trustworthy source of information is, generally, transparent in its ownership, authorship and sourcing of information, in addition, it holds procedures in place to clearly label advertisement and paid content, as well as separating fact from opinion.

These considerations can be relevant when platforms have to make decisions that aim to contribute to a trustworthy online ecosystem. The European Commission points out that online platforms have supported the development of projects by independent third parties to design trustworthiness and credibility indicators, such as the Trust Project, the Cred-

-
- 23 In parallel with the discussion of trustworthiness of online content sources, it must be acknowledged that the EU audiovisual policy is facing the challenges of defining standards for the online environment and is proposing, since its most recent revision in 2018, not only “prominence” of European works as an obligation for all on-demand AVMS (Article 13(1), Recital 35 AVMSD), but also that “Member States may take measures to ensure the appropriate prominence of audiovisual media services of general interest” (Article 7(a), Recital 25 AVMSD). Member States are still in the process of adopting national prominence frameworks and approaches significantly vary from country to country. Some built on long standing traditions regarding PSM, others consider the use of ‘quality labels’. See Eleonora Maria Mazzoli and Damian Tambini, “Prioritisation uncovered. The discoverability of public interest content online”, Council of Europe, 2020. <https://rm.coe.int/publication-content-prioritisation-report/1680a07a57>.
- 24 European Commission VVA, “Study on the assessment of the Code of Practice against Disinformation SMART 2019/0041”, 2020. <https://digital-strategy.ec.europa.eu/en/library/study-assessment-implementation-code-practice-disinformation>; ERGA, “ERGA Report in Disinformation”; European Commission, “Staff Working Document (SWD (2020)180 Final Assessment of the Code of Practice on Disinformation” 2020. <https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>.

ibility Coalition or the Journalism Trust Initiative (JTI).²⁵ However, an evaluation by VVA highlighted that there is no detailed information available on the integration of these indicators in platforms' search services and recommender systems.²⁶ There is also no mention in the documents of the criteria the platforms use to determine one source's trustworthiness, aside from recommendation by fact checkers, and Microsoft's partnership with large, "vetted" sources.²⁷

In order to assist the operationalisation of trustworthiness indicators, the CMPF has looked at the three above-mentioned initiatives – the Credibility Coalition, the JTI and the Trust Project – as well as the Newsguard²⁸ browser extension. Their indicators focused, among others, on the following key areas:

- a. Past conduct of publisher: looking at whether or not it was found repeatedly publishing verifiably false information;
- b. The sourcing of articles: focusing on the diversity of sources used in published items, the transparent sourcing of articles (existence and quality of references, hyperlinks, quotes from identified sources), the openness of methods used to acquire information, reliance on reader feedback and the logical soundness of content published
- c. Correction and labelling: looking at whether errors and inaccuracies were corrected or clarified on time, advertising and sponsored content was clearly labelled, fact was separated from opinion.
- d. Transparency of funders and content creators: the emphasis is on the disclosure of ownership and financing of the media organisation, as well as the disclosure of authors, including their contact details.

In its Staff Working Document, the Commission indicates a preference for *ex ante* measures, e.g. when recommending the following option: "Ex ante approval by ad-placement service providers of websites selling advertisement space, possibly based on trustworthiness indicators agreed with advertisers (a 'white list' approach)".²⁹ This *ex ante* approach and white list is in line with the Code's attempts to classify content producers or content sources as trustworthy and untrustworthy, and the effort can be

25 "Journalism Trust Initiative". <https://www.journalismtrustinitiative.org/>, "The Trust Project," <https://thetrustproject.org/>, "Credibility Coalition". <https://credibilitycoalition.org/>.

26 VVA, "Study on the assessment of the Code of Practice".

27 European Commission, "Staff Working Document," 6.

28 "NewsGuard". <https://www.newsguardtech.com/>.

29 European Commission, "Staff Working Document," 8.

supported by the above listed indicators as well. Some of these indicators can be checked automatically (e.g. existence of a masthead, owner information, as well as additional indicators, such as being registered with the country's media authority, or checking the average number of outside links, corrections, etc.) or provide the basis of self-reporting (such as the machine-readable, detailed questionnaire of JTI). Others need the active work of users and fact checkers (such as reporting suspicious contents by users and then flagged by fact-checkers).

However, this approach may raise some concerns. Even if the Commission's guidance emphasises that users can decide for themselves whether they want the services provided to them to be curated by trustworthiness indicators or not,³⁰ tools that rely solely on these indicators when determining trustworthiness of content sources may create a media environment in which established players gain further competitive advantage, while new players will face unprecedented barriers to entry. This can lead to serious problems for media pluralism and can distort the media market in a way that news players will find their access to the advertising market or other revenue sources further limited. The overreliance on these indicators can also silence diverging or non-mainstream voices.

At discussions among stakeholders, representatives of publishers have also signalled that reporting about one's trustworthiness (or even auditing this reporting) based on indicators like the ones developed by JTI or the Trust Project cannot be made mandatory. Thus, they argue, media outlets should not be labelled untrustworthy simply for not being party to such a project or initiative. Not to mention that the Code itself highlights that measures should be consistent with Article 8 of the European Convention on Human Rights (right to respect of private and family life), the fundamental right of anonymity and pseudonymity, and the proportionality principle – these could all be violated by too stringent reporting requirements on, among others, ownership or authorship. In addition, the Code also highlights Article 10 of the European Convention on Human Rights (freedom of expression), as decisions on prioritisation might limit users' access to relevant ideas and information.

In light of the previously highlighted concerns, we recommend an approach that is built on carrots, but without evident sticks. This approach would mean that content sources with a large enough audience would be asked to provide sufficient information about their compliance with indicators. Although non-compliance would not be punished with downgrad-

30 European Commission, "Guidance", 16.

ing, compliance should be rewarded with upgrading (prioritising) one's content. In practice this would mean that non-compliant publishers could go on using the services according to the current terms, while compliant sites would receive a boost by being shown more frequently to users, and by being included on a list of trusted partners for advertisers.

In parallel, fact-checkers would monitor content, or react to reporting by users. Those content creators who are caught repeatedly publishing misinformation or disinformation would be downgraded in rankings. As the detailed assessment of trustworthiness is not feasible in the case of small or new players, they should get a chance to use the organic (not paid) services of social media to reach audiences without constraints, as long as there is no sign of malicious use of the content-sharing platforms. Social media platforms themselves have already introduced some checks and requirements for users or accounts that come into play once they aim to monetise their content or boost their messages;³¹ these requirements can also be used for quick trustworthiness checks to filter out which providers have to be subject to increased scrutiny.

Chapter 7. Conclusion

Disinformation is a challenging issue in content moderation, as it refers to content that is often not illegal, but can cause harm. As such, it reshapes the ways in which we think of the liability of online platforms and the governance of the digital environment. The findings of the EU-wide Media Pluralism Monitor data collection show that disinformation is increasingly seen as a risk to both media pluralism and democratic processes in EU member states, and the policy responses are often regarded as unsuitable to address the problem. This has been reiterated by our assessment of the Code of Practice. Even if signatories make a pledge for cooperation, the measures taken by online platforms under the commitments of the Code of Practice often fall short of what they have committed to. Problems can be traced back to three groups of issues: lack of guidance, limited compliance and the small number of signatories. Policymakers, on the

31 See Google.com, "YouTube Channel Monetization Policies," <https://support.google.com/youtube/answer/1311392?hl=en#zippy=%2Cfollow-adsense-program-policies>. Facebook.com, "Facebook Community Standards." https://www.facebook.com/business/help/185404538833362?id=2520940424820218&recommended_by=321041698514182.

other hand, might leave possible side-effects, such as limits to freedom of expression or a decrease of media pluralism, unattended.

There are some important attempts to strengthen the Code. The Commission came up with guidelines for a new Code 2.0 that functions as a co-regulatory framework, KPIs and indicators are developed to better assess platforms' compliance and the Code's impact, while independent third parties are working on indicators of trustworthiness to provide internet users with the necessary tools for informed online navigation. These efforts are still just taking shape, thus stakeholders, such as representatives of academia, civil society, fact-checking organisations, the media, the advertising sector and regulatory authorities, need to be ready to continue the deliberation and come up with proposals that address current shortcomings.

Bibliography

- Brogi, Elda, Roberta Carlini, Iva Nenadic, Pier Luigi Parcu and Mario Viola de Azevedo Cunha. *Monitoring media pluralism in the digital era: Report 2020*. Florence, 2020. <https://cadmus.eui.eu/bitstream/handle/1814/67828/MPM2020-PolicyReport.pdf?sequence=5&isAllowed=y>.
- Brogi, Elda and Nenadic, Iva. "European plan to increase transparency and accountability of the gatekeeper online platforms to protect democracy: EDMO's role in the Commission's digital policy approach", 2021. <http://www.medialaws.eu/european-plan-to-increase-transparency-and-accountability-of-the-gatekeeper-online-platforms-to-protect-democracy-edmos-role-in-the-commissions-digital-policy-approach/>.
- Brogi, Elda et al. *Monitoring media pluralism in the digital era: Report 2021*. Florence, forthcoming 2021.
- Credibility Coalition. "Credibility Coalition: Our goal: to understand the veracity, quality and credibility of online information." Accessed July 14, 2021. <https://credibilitycoalition.org/>.
- ERGA. "ERGA Report on Disinformation: Assessment of the Implementation of the Code of Practice", 2020. <https://erga-online.eu/wp-content/uploads/2020/05/ERGA-2019-report-published-2020-LQ.pdf>.
- European Commission. *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*. Brussels: European Commission, 2000. <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031>.
- European Commission. *Tackling online disinformation: a European Approach. COM/2018/236 final*, Brussels: European Commission, 2018. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0236>.

- European Commission. *Action Plan against disinformation*. Brussels: European Commission, 2018. <https://digital-strategy.ec.europa.eu/en/library/action-plan-against-disinformation>.
- European Commission. *Code of Practice on Disinformation*. Brussels: European Commission, 2018. <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>.
- European Commission. *European Democracy Action Plan: making EU democracies stronger*. Brussels: European Commission, 2020. https://ec.europa.eu/commission/presscorner/detail/en/IP_20_2250.
- European Commission VVA. *Study on the assessment of the Code of Practice against Disinformation SMART 2019/0041*. Brussels: European Commission, 2020. <https://digital-strategy.ec.europa.eu/en/library/study-assessment-implementation-code-practice-disinformation>.
- European Commission. *The Digital Services Act package*. Brussels: European Commission, 2020. <https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package>.
- European Commission. *Staff Working Document (SWD (2020)180 Final - Assessment of the Code of Practice on Disinformation*. Brussels: European Commission, 2020. <https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>.
- European Commission. *Annual self-assessment reports of signatories to the Code of Practice on Disinformation*. Brussels: European Commission, 2019. <https://digital-strategy.ec.europa.eu/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019>.
- European Commission. *European Commission Guidance on Strengthening the Code of Practice on Disinformation European Commission*. Brussels: European Commission, 2019. <https://ec.europa.eu/newsroom/dae/redirection/document/76495>.
- High Level Expert Group on Fake News and Online Disinformation. *A multi-dimensional approach to disinformation - Report of the independent High level Group on fake news and online disinformation*, Brussels: European Commission, 2018. <https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation>.
- Google. "YouTube Channel Monetization Policies" Accessed July 14, 2021. <https://support.google.com/youtube/answer/1311392?hl=en#zippy=%2Cfollow-adsense-program-policies>. Facebook.com, "Facebook Community Standards. https://www.facebook.com/business/help/185404538833362?id=2520940424820218&recomended_by=321041698514182.
- Journalism Trust Initiative. "Journalism Trust Initiative. Rewarding trustworthy Journalism" Accessed July 14, 2021. <https://www.journalismtrustinitiative.org/>.
- Mazzoli, Eleonora Maria and Damian Tambini. *Prioritisation uncovered. The discoverability of public interest content online*. Strasbourg: Council of Europe, 2020. <https://rm.coe.int/publication-content-prioritisation-report/1680a07a57>.
- Nenadić, Iva. "Unpacking the 'European approach' to tackling challenges of disinformation and political manipulation," *Internet Policy Review* 8, No. 4 (2019): 1-22.

- NewsGuard. “NewsGuard: Das Vertrauens-Tool fürs Netz,” Accessed July 14, 2021. <https://www.newsguardtech.com/>.
- The Trust Project. “The Trust Project: How do you know which news stories you can trust?” Accessed July 14, 2021. <https://thetrustproject.org/>.
- Wardle, Claire and Hossein Derakhshan. “Thinking about ‘information disorder’: formats of misinformation, disinformation, and mal-information,” in *Journalism, ‘Fake News’ & Disinformation. Handbook for Journalism Education and Training*, eds. Cherilyn Ireton and Julie Posetti, 44-56. Paris: UNESCO, 2018. <https://unesdoc.unesco.org/ark:/48223/pf0000265552/PDF/265552eng.pdf.multi>.
- Woolley, Samuel. “Encrypted messaging apps are the future of propaganda,” *Brookings*, May 1, 2020, <https://www.brookings.edu/techstream/encrypted-messaging-apps-are-the-future-of-propaganda/>.

