THE ROLE OF SELF REGULATION AND REGULATION IN COMBATING DISINFORMATION

Ekin Balkan
Sinan Ülgen

INTRODUCTION

In the era of digitalization and digital communication, globally, societies are increasingly confronted with challenges stemming from and/or amplified by disinformation. The trend of digital media and digitalization in the information ecosystem, which has a profound impact on social, political and economic dynamics, not only affects societies but is also shaped by the specific societal dynamics. At a time when the information crisis has emerged as a major challenge, there are increasing demands and expectations for social media platforms to take responsibility for combating disinformation and for ensuring a safe online environment.

This research paper examines the internal policies and self-regulation models of social media platforms in combating disinformation in the current context, taking into account the digital media landscape and specific dynamics in Turkiye and the globally evolving approaches to self-regulation in social media (especially in the EU).

The paper consists of 3 chapters in order to analyze the main characteristics and dynamics of digital media and social media in Turkiye, the changing context regarding the regulation of social media platforms and the different internal policies adopted by the platforms.

First chapter focuses on the context and the background by analyzing and examining the role and effects of the social media platforms both globally and in Turkiye. Moreover, this chapter analyzes the specific dynamics of the digital landscape (polarization, fragmentation, distrust and censorship) in Turkiye.
The second chapter discusses and compares the approaches to combating disinformation on social media platforms in Turkiye and the EU. Here, in the first section the main legislations enacted in Turkiye in recent years and the effects of these legislations are examined. This section is followed by an analysis of the EU Code of Practice, (the world's first selfregulatory framework to combat disinformation), and the Digital Services Act and Digital Markets Act that followed the EU Code of Practice.
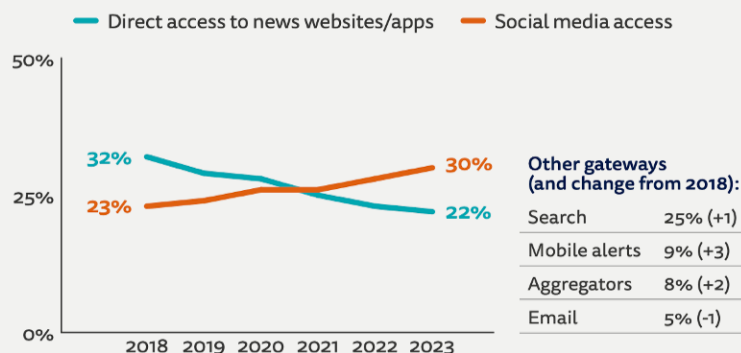
The last chapter deals with the internal policies and self-regulation models of 4 major social platforms (Twitter, Facebook, Youtube, TikTok). The chapter compares the approaches of these platforms and explores the similarities and differences between certain internal policies adopted to combat disinformation.

CHAPTER 1: THE SIGNIFICANCE AND RELEVANCE OF SOCIAL MEDIA PLATFORMS FOR TACKLING DISINFORMATION

The innovations brought about by digital media in the information ecosystem and communication are multidimensional, with social, cultural, economic and political aspects. Since it affects billions of people simultaneously in various aspects of their lives, these dimensions are intertwined and in a mutual interaction that is in constant change due to the digitalization of the media. Hence the analysis of approaches, policies and measures to combat disinformation in this field is too varied and complex to be analyzed through a single lens and interpreted globally due to the multidimensional nature of digital media. Therefore, in order to contextualize the issues to be examined in this paper for a comprehensive and meaningful analysis, this chapter will examine the relevance of social platforms, which are indispensable both for disinformation to reach the masses and for digital news media.

Social media platforms are of critical importance as they are public spaces where large masses of people spend a long time, are active and engage in social interactions. However, the importance of platforms does not end there. While platforms such as Twitter are a space for a variety of political and social debates and narratives, major online platforms also serve as important intermediaries for people to access news and news sources. As the recently published Reuters Institute Digital News Report notes, the steady rise in the global use of social platforms as a means of accessing online news suggests that this function will only grow in importance. The report shows that the share of social media in accessing news has increased from 22% in 2018 to 30%, while direct access to news websites and apps has decreased from 32% to 22%[1]

PROPORTION THAT SAY EACH IS THEIR MAIN WAY OF GETTING NEWS ONLINE (2018–2023) – ALL MARKETS

— Direct access to news websites/apps   — Social media access

Other gateways (and change from 2018):	
Search	25% (+1)
Mobile alerts	9% (+3)
Aggregators	8% (+2)
Email	5% (-1)

Q10a_new2017_rc. Which of these was the main way in which you came across news in the last week? *Base: All who used a news gateway in the last week in each market-year ≈ 2000. Note: Number of markets grew from 36 in 2018 to 46 from 2021 onwards. Markets listed in online methodology.*

Source: Reuters Institute Digital News Report 2023

---

1   Nic Newman et al., "Reuters Institute Digital News Report 2023," 2023. 11.

Based on this data, given that platforms are not uniform but structurally different in terms of shared and consumed content, it can be inferred that changes in the approaches adopted by social media platforms in the regulation of disinformation, or the decrease in the prevalence of certain social platforms and the increase in the use of certain platforms, may have wide-ranging effects and should therefore be examined. For instance while the use of social media platforms to access news is increasing globally, the share of each platform does not follow this trend. The percentage of people using Facebook for news fell from a peak of 42% in 2014 to 28% in 2023. Meanwhile, during this time period, the share of Instagram increased from 2% to 14%. Another dramatic change is that Tiktok's share increased from 1% in 2020 to 6% in 2023. According to the data, the fact that the percentage of users that use TikTok for news has increased up to 20% for users between the ages of 18-24 is an important indicator that should be taken into consideration regarding the change in the dynamics of this field in the future.[2]

In addition to the aspects and dimensions listed above, social media can also be considered as a game changer to increase the impact, speed and number of disinformation and influence operations. Along with the structural features of digital and social media that were analyzed and described in the first paper, social media platforms provide the environment and opportunities for the five functions of a campaign: reconnaissance, hosting, placement, propagation and saturation.[3]

Key trends in Social Media Platforms and the overall state of Social Platforms in 2023

Apart from the steady increase in the use of social media platforms for acquiring news, the proportional decline of Facebook's share in this domain, and the rise of Tiktok and Instagram, there are also important trends and features concerning disinformation and social media platforms' approaches to disinformation.

Trust in news and news media is declining around the world. An important point about this trend is that the prevalence and intensity of criticism of the media and the decline in trust in news go hand in hand. In countries such as the US, where political polarization has increased and the media has been targeted by various actors, levels of distrust has increased.[4] Turkiye, along with the US and Hungary, is among the countries most exposed to politicians' criticism of the media.[5]

There are also important distinctions between platforms in terms of news consumption and news sources. While news media platforms and journalists are still dominant on Facebook and Twitter, audiences on TikTok and Instagram prefer

2   "Overview and Key Findings of the 2023 Digital News Report," Reuters Institute for the Study of Journalism, accessed June 15, 2023, https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/dnr-executive-summary.
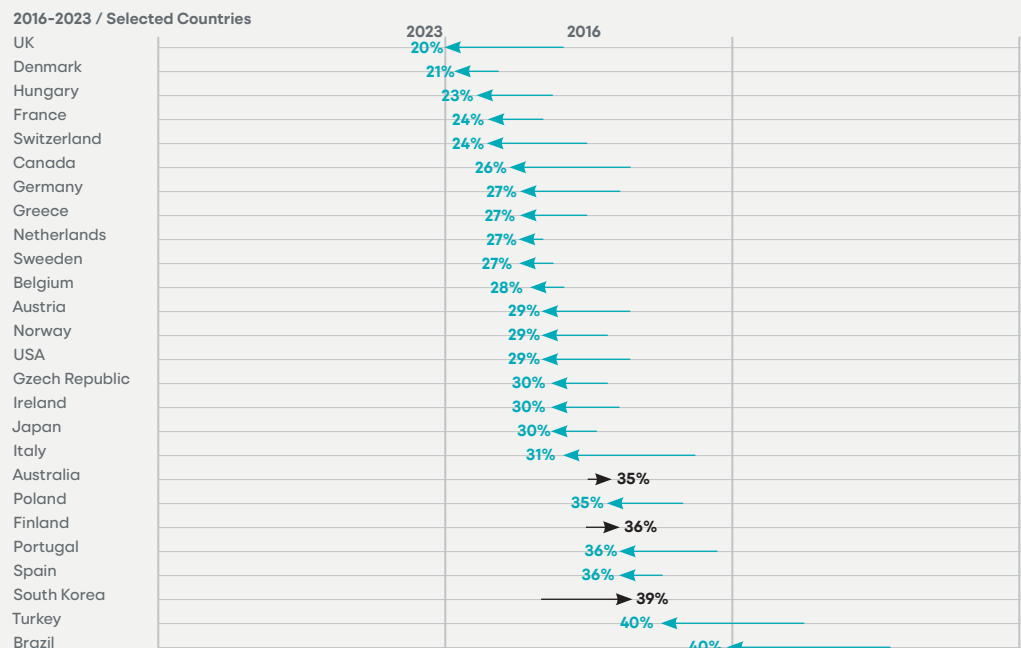3   Glenda Jakubowski, "What's Not to Like?: Social Media as Information Operations Force Multiplier," 2019.
4   "Overview and Key Findings of the 2023 Digital News Report," Reuters Institute for the Study of Journalism, accessed June 15, 2023, https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/dnr-executive-summary.
5   Nic Newman et al., "Reuters Institute Digital News Report 2023," 2023. 41.

influencers, celebrities and social media personalities over journalists as news sources.[6] On platforms such as Instagram and Tiktok, where fast-paced visual content is predominant, lacking nuance and easily decontextualized videos and content, all of these features, coupled with the fact that users prefer social media personalities over professional journalists as news sources, further complicates existing challenges and increases the risks associated with disinformation. In addition, the public's fears and concerns about misinformation and disinformation, which have been observed in recent years, have not changed. Globally, the proportion of people who worry about distinguishing between what is real and what is fake online has increased by 2 percentage points to 56% in 2022. Among those who use social media as their main source of news, this figure reaches 64%.[7] The steadily increasing concern about distinguishing between real and fake news suggests that there may be a correlation with the previously mentioned growing distrust of media.

Another important and growing global concern is about the algorithms of social media platforms and their impact on users. People's opinions about platforms' algorithmically tailored news selection are increasingly negative. A large proportion of people are skeptical about the automatic selection of news stories by social media platforms based on what they or their friends have consumed in the past. More interestingly, this skepticism tends to persist even when news selection is done by editors and journalists.[8] One reason for this steady decline could be that the risks and consequences of social media platforms' algorithms and over personalization have become more and more evident and increasingly present in the public debate. Turkiye is along with Brazil one of the countries demonstrating the sharpest decline.

Proportion that agree that having stories selected for them based on what they have consumed in the past is a good way to get news

2016-2023 / Selected Countries

Country	2023	2016
UK	20%	
Denmark	21%	
Hungary	23%	
France	24%	
Switzerland	24%	
Canada	26%	
Germany	27%	
Greece	27%	
Netherlands	27%	
Sweeden	27%	
Belgium	28%	
Austria	29%	
Norway	29%	
USA	29%	
Gzech Republic	30%	
Ireland	30%	
Japan	30%	
Italy	31%	
Australia	35%	
Poland	35%	
Finland	36%	
Portugal	36%	
Spain	36%	
South Korea	39%	
Turkey	40%	
Brazil		40%

6   "Overview and Key Findings of the 2023 Digital News Report," Reuters Institute for the Study of Journalism, accessed June 15, 2023, https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/dnr-executive-summary.
7   Ibid.
8   "Attitudes towards Algorithms and Their Impact on News," Reuters Institute for the Study of Journalism, accessed June 17, 2023, https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/attitudes-towards-algorithms-impact-news.

Turkiye's Digital Media Landscape: Trends, Challenges and Characteristics

In order to understand the main characteristics of anti-disinformation policies in Turkiye, their impact on social media platforms and the main characteristics of anti-disinformation policies in Turkiye and their impact on society, it is necessary to analyze Turkiye's digital media and in particular social media landscape. To this end, this section will first provide an overview of the impact of digital media in Turkiye by examining its prevalence and its evolution over time. It will then examine the structural features and challenges of digital media and news consumption in Turkiye, such as high distrust, partisanship, polarization and censorship.

Overview of the trends

According to datareportal's Digital 2023 report, the number of internet users in Turkiye at the beginning of 2023 is 71.38 million, and 83.4% of the total population has access to the internet. 73.1% of the population uses social media, and 59.6% of social media users are male and 40.4% female. Moreover, 51% of social media users in Turkiye use social media to read news.[9] According to the Reuters Institute, in 2023 the main social platforms used for news are Instagram and Youtube.[10]

There have been significant specific changes in social media platforms regarding their prevalence and reach in more recent years. Meta's social media platforms, Facebook and Instagram, saw their online ad reach decline according to the report. Although Facebook has 32.8 million users, its ad reach decreased by 1.6 million people between 2022 and 2023, down 4.7 percent. Instagram saw a 6.7 percent decline in the same time period. Facebook's worldwide decline can also be observed in Turkiye.[11]

Youtube, with 57.9 million total users, increased its ad reach by 0.9 percent between 2022-2023. Moreover, it should be emphasized that Youtube attracts attention in terms of gender balance with 49.5 percent female and 50.5 percent male users in Turkiye compared to other platforms. Like the rest of the world, TikTok increased its reach in Turkiye by 12.4 percent between 2022 and 2023, which is quite high. A similar increase of 15.2% is observed on Twitter.[12] Although the number of TikTok users over the age of 18 is 29.8 million, considering the high prevalence of TikTok among the young population, it can be inferred that this figure is even higher and has the potential to increase significantly in the future.

Various studies have examined the role of social media platforms in terms of access to news. For example, a study conducted in 2018 found that online news, including social media, is used as a source by 89% of the population in Turkiye. And yet in the same survey, 48 percent of the respondents stated that they use television as their

---

9   "Digital 2023: Turkiye," DataReportal – Global Digital Insights, February 13, 2023, https://datareportal.com/reports/digital-2023-Turkiye.
10   Nic Newman et al., "Reuters Institute Digital News Report 2023," 2023. 105.
11   "Digital 2023: Turkiye," DataReportal – Global Digital Insights, February 13, 2023, https://datareportal.com/reports/digital-2023-Turkiye.
12   Ibid.

main source, while 39 percent stated that they prefer online news (including social media) as their main source.[13] The online news exposure of 89%, is an important indicator that demonstrates the potential and importance of digital disinformation.

Characteristics and challenges

A characteristic feature of digital media in Turkiye is a high level of polarization. Polarization facilitates the spread of disinformation and simultaneously complicates the fight against disinformation, as it can lead different groups in society to prefer different news sources, isolate themselves from opposing views and discourses, and even lead them to live in alternative realities with alternative facts.

The discourses and narratives on digital media and social media platforms emerge from, feed on and are shaped by Turkiye's political and social environment.

The core reasons for the high level of political polarization in Turkiye are the existing social and political cleavages namely the center-periphery, secular-religious cleavages and political tensions over the Kurdish question.[14] The effects of polarization on social media is generally more severe. While people prefer to discuss politically polarizing or polarized issues within their own families, close circles of friends or neighborhoods, they prefer to express their opinions or engage with opposing views at relatively much lower rates on social media, even though it is considered as an alternative public space.[15] The article entitled Dimensions of Political Polarization in Turkiye demonstrates that the media preferences of the public are in line with political preferences in Turkiye (opposition party supporters prefer different newspapers and channels, while the ruling party and its partners prefer different newspapers and channels) and they do not differentiate on social media, and that users on Facebook and Twitter overwhelmingly interact with like-minded people and posts.[16]

Two trends, reinforced by this high polarization, make Turkiye's digital and media ecosystem more vulnerable to disinformation. These are high distrust of the media and high partisanship, reinforced by the sharp fragmentation of the media. A significant number of government supporters, as well as opposition voters, think that the media is not able to act freely and report every issue and news deemed important.[17]

13   "Digital News Report 2018 - Turkiye Supplementary Report," Reuters Institute for the Study of Journalism, accessed June 17, 2023, https://reutersinstitute.politics.ox.ac.uk/our-research/digital-news-report-2018-Turkiye-supplementary-report.
14   "Dimensions of Polarization in Turkiye | Strengthening Transatlantic Cooperation," accessed June 17, 2023, https://www.gmfus.org/news/dimensions-polarization-Turkiye.
15   "Dimensions of Polarization in Turkiye 2017 - TurkuazLab - Strategies and Tools for Mitigating Polarization in Turkiye Project," November 12, 2020, https://www.turkuazlab.org/en/dimensions-of-polarization-in-Turkiye-2017/.
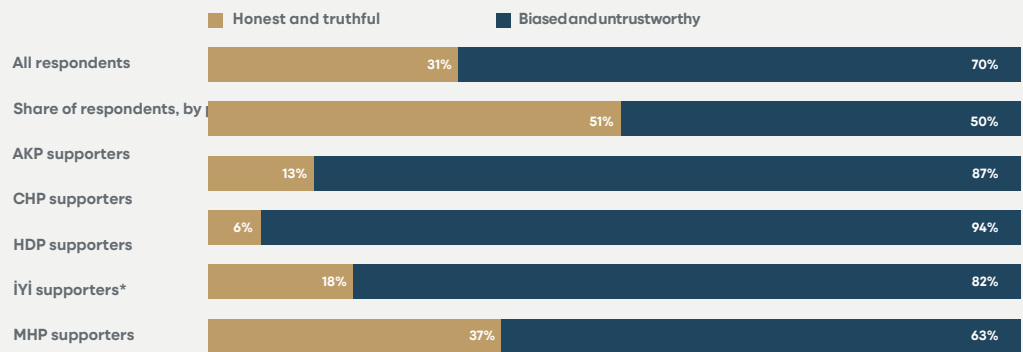16   Ibid.
17   "Turkiye's Changing Media Landscape," Center for American Progress (blog), June 10, 2020, https://www.americanprogress.org/article/Turkiyes-changing-media-landscape/.

This vulnerability should be seen as a challenge that facilitates disinformation campaigns by both domestic and foreign actors, with the potential for great harm. Masses that are highly polarized on certain key political and social issues and due to the fragmentation of the media, are not exposed to pluralist views are more susceptible to manipulation. This risk is especially heightened during periods of high tension and social engagement, such as election periods. The reactivation of nearly 12,000 Russian and Hungarian accounts in Turkish on Twitter before the 2023 elections in Turkiye points to the potential problems that such disinformation operations can create with this polarization in the Turkish digital and social media ecosystem.[18]

Reflecting these attitudes and sentiments towards the media, in a similar vein, a very high level of distrust of the media is observed, predominantly among supporters of opposition parties. Rising distrust, coupled with polarization, leads to a clear and sharp divide between opposition and pro-government media outlets in Turkiye. As a result, the opposition audience is more likely to use social media and online sources to access news than pro-government voters.[19] Based on survey data, a CAP study argues that this trend is motivated by the need to access non-censored information.

FIGURE 1
Distrust of the media is very high in Turkey
? | In general, do you think that the media in Turkey presents honest and truthful information or that it presents biased and untrustworthy information?

Legend: ■ Honest and truthful   ■ Biased and untrustworthy

	Honest and truthful	Biased and untrustworthy
All respondents	31%	70%
Share of respondents, by	51%	50%
AKP supporters	13%	87%
CHP supporters	6%	94%
HDP supporters	18%	82%
İYİ supporters*	37%	63%
MHP supporters		

* When the survey was conducted—from May to June 2018—the İYİ Party had not yet competed in an election. Therefore, to identify İYİ Party supporters, the authors looked at respondents who voted for the MHP in November 2015 but stated in CAP's survey that they approved of the İYİ Party and did not plan to vote for the AKP-MHP alliance in the 2018 elections.
Note: Percentages may not total 100 due to rounding.
Source: CAP national survey of 2,534 respondents in Turkey, May 24 to June 4, 2018.

CAP

Source: "Turkiye's Changing Media Landscape," Center for American Progress (blog), June 10, 2020, https://www.americanprogress.org/article/Turkiyes-changing-media-landscape/.

---

18   https://www.facebook.com/middleeasteye, "Thousands of Russian Twitter Accounts Reactivated in Turkish," Middle East Eye, accessed June 18, 2023, http://www.middleeasteye.net/news/Turkiye-elections-thousands-russian-speaking-accounts-activated-twitter.
19   "Turkiye's Changing Media Landscape," Center for American Progress (blog), June 10, 2020, https://www.americanprogress.org/article/Turkiyes-changing-media-landscape/.

The trends and dynamics elaborated so far in this section lay the foundation for a more vulnerable Turkish media ecosystem to disinformation. High social and political polarization leads to people being trapped in their own bubbles on social media, the prevalence of biased media outlets combined with deep media fragmentation increases the speed and reach of disinformation that is deliberately spread to cause harm, and the lack of a common ground for social engagement and debate makes it substantially difficult to combat disinformation.

In this context, difficulties related to online freedoms often perpetuate the problems stated above. The decline in Reporters Without Borders' Press Freedom Index from an already poor ranking of 149 worldwide in 2022 to 165th in 2023[20] is to be highlighted in this respect. Reporters Without Borders' analysis of the state of Turkish media in 2023 emphasized the prevalence of a systematic policy of censorship on the internet, discriminatory legal sanctions against certain journalists and media outlets. The analysis also states that 90 percent of the media is under government control, either directly or through pro-government private companies.[21]

Considering all these factors, it is clear that the government's approach to the media, regulations and legislation have a significant impact on the media landscape. Therefore, it is important to examine the policies developed to combat disinformation on social media through this lens, to compare the Turkish approach with other contemporary practices and to analyze the differences, commonalities and its potential impact on society.

20   "Türkiye | RSF," May 31, 2023, https://rsf.org/en/country-t%C3%BCrkiye.
21   "Türkiye | RSF."

CHAPTER 2: COMBATING DISINFORMATION IN THE EU AND TURKIYE

In this chapter, the first section examines the evolution of practices aimed at combating disinformation in the EU. The second section will cover government regulations and initiatives aimed at combating disinformation in Turkiye with a special focus on social media.

Social Media Platforms and the EU: EU Code of Practice, Digital Services Act and the Digital Markets Act

The EU's policies and approach to disinformation include both institutional initiatives and legislative frameworks. This section discusses the EU's approach towards combating disinformation on social media platforms since the approach of the EU has gradually evolved and changed over time. Starting with the first of its kind EU Code of Practice in 2018, which was voluntary to join and sign, the process continues with the Digital Services Acts (DSA) and the Digital Markets Act (DMA), a more comprehensive and legally binding framework. Nevertheless, the EU's approach to social media platforms aims to find the most effective practices and frameworks to combat disinformation, mostly by involving various stakeholders from social media platforms and civil society organizations, and to set standards based on long-term assessments.

The EU Code of Practice on Disinformation

The EU Code of Practice on Disinformation is the world's first self-regulatory framework for combating disinformation, adopted and signed with the voluntary participation of online platforms, tech companies and other actors in the sector.[22] "The 2018 Code of Practice aimed at achieving the objectives set out by the Commission's Communication presented in April 2018 by setting 21 commitments in different domains, from transparency in political advertising to demonetization of purveyors of disinformation."[23]. In 2018, Facebook, Twitter and Google signed the EU Code of Practice. TikTok joined in 2020. All these platforms have developed separate roadmaps to fulfill the commitments in the Code of Practice and have shared them with the Commission.[24]

Another important aspect of the Eu Code of Practice on Disinformation is that signatories of the Code regularly produce annual self-assessment reports and share them with the Commission, providing an opportunity for an overall assessment of the status of the fight against disinformation on social media platforms, the

22 "2018 Code of Practice on Disinformation | Shaping Europe's Digital Future," June 16, 2022, https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation.

23 "2018 Code of Practice on Disinformation | Shaping Europe's Digital Future," June 16, 2022, https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation.

24 "Roadmaps to Implement the Code of Practice on Disinformation | Shaping Europe's Digital Future," October 16, 2018, https://digital-strategy.ec.europa.eu/en/news/roadmaps-implement-code-practice-disinformation.

results and effectiveness of the policies implemented. This practice has become an important framework for assessing and monitoring the disinformation policies and the level and status of the partnerships of social media platforms with other stakeholders such as researchers and factcheckers.[25]

As a result of the revision process of the EU Code of Practice, the initial code in 2018 became the Strengthened Code of Practice on Disinformation in 2022. The Strengthened Code includes 44 commitments and 128 specific measures. The revision process of this Code involved 34 signatories, which were parties to the 2018 Code.[26]

The responsibilities and measures that social media platforms should take to combat disinformation are divided into the following thematic areas: Scrutiny of ad placements, political advertising, integrity of services, empowering users, empowering the research community, empowering the fact-checking community. In addition, the Code of Practice also includes the setting up of a transparency center and a permanent task force and establishes a more comprehensive monitoring framework to measure the impact of the Code and the implementation of the Code's commitments and measurements.[27] One of the major focus of the Code is on the relationship between disinformation and advertising, which is the source of income for social media platforms. Specifically, certain articles in the Code aim to prevent ads from containing disinformation and to prevent disinformation from becoming a lucrative business by advertising disinformation content. In addition, demonetization is adopted as a method to combat disinformation, aiming to combat the financial dynamics that sustain disinformation on social media platforms. The Code of Practice on Disinformation also acknowledges the importance of political advertising on social media and emphasizes that social media platforms should increase transparency in this area.[28] Another important dimension is the empowerment of users and other important stakeholders such as researchers and fact-checkers. For the empowerment of the users, the strategy consists of making it easier for users to identify disinformation, developing recommender systems and designing services in a safer way so that users can avoid disinformation. For the empowerment of the researchers and fact-checkers the EU Code of Practice contains commitments on increasing the cooperation between the fact-checkers/ researchers and social media platforms. These commitments also include enabling and facilitating factcheckers' access to information and researchers' access to relevant data.[29]

25   "Annual Self-Assessment Reports of Signatories to the Code of Practice on Disinformation 2019 | Shaping Europe's Digital Future," October 29, 2019, https://digital-strategy.ec.europa.eu/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019.

26   "The 2022 Code of Practice on Disinformation | Shaping Europe's Digital Future," June 2, 2023, https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation.

27   "2022 Strengthened Code of Practice on Disinformation | Shaping Europe's Digital Future," June 16, 2022, https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation.

28   Ibid.

29   Ibid.

This revised EU Code of Practice is intended to be part of a more comprehensive regulatory framework (as stated by the Commission). This strengthened Code of practice will function as a code of conduct and a set of measures against disinformation within the co-regulatory framework established with the DSA.[30]

Digital Services Act Package

The Digital Services Act Package is composed of two separate acts. The Digital Services Act and the Digital Markets Act. According to the Commission the two main goals of these Acts are: creating "a safer digital space in which the fundamental rights of all users of digital services are protected"[31] and establishing "a level playing field to foster innovation, growth, and competitiveness, both in the European Single Market and globally."[32].

With these two acts, online services are classified as intermediary services, hosting services, online platforms and very large online platforms (VLOPs) according to their reach and market share. It imposes new obligations and responsibilities depending on the size of the platforms and their impact on the market.[33] Platforms such as Alibaba AliExpress, Amazon Store, Facebook, Google Play, Google Maps, Instagram, LinkedIn, Snapchat, TikTok, Twitter and Youtube are all defined as Very Large Online Platforms. The main clauses affecting the regulation of disinformation on social media platforms are related to user empowerment and content moderation.   With the Digital Services Act Package, platforms will be obliged to label advertisements and inform users who is providing and promoting these advertisements, and to evaluate the content reported by users in a detailed and systematic manner based on rules and without arbitrariness. Moreover, platforms are prevented from targeting users and displaying ads based on the user's ethnicity, political opinions and sexual orientation. In terms of content moderation, platforms will be obliged to create mechanisms for users to flag content and be more active in this area in line with these regulations, analyze specific potential problems and risks on their platforms and take measures to address these problems.[34]

In conjunction with these obligations, to ensure transparency and accountability within this coregulatory framework, platforms must allow external audits, provide access to data to researchers and publish transparency reports on their decisions regarding content moderation.[35]

30   "The 2022 Code of Practice on Disinformation | Shaping Europe's Digital Future," June 2, 2023, https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation.
31   "The Digital Services Act Package | Shaping Europe's Digital Future," June 13, 2023, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.
32   Ibid.
33   "The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications," accessed June 19, 2023, https://www.tandfonline.com/doi/epdf/10.1080/17579961.2023.2184136?needAccess=true&role=button.
34   "The Digital Services Act (DSA)," accessed June 19, 2023, https://www.eu-digital-services-act.com/.
35   Ibid.

Table 1. New obligations for gatekeepers in the DSA.

		Intermediary services	Hosting services	Online platforms	Very large online platforms
Transparency measures for online platforms	Transparency reporting	✓	✓	✓	✓
	Requirements on terms of services due account of fundamental rights	✓	✓	✓	✓
	Notice-and-action and obligation to provide information to users		✓	✓	✓
	User-facing transparency of online advertising			✓	✓
	Transparency of recommender systems and user choice for access to information				✓
Oversight structure to address the complexity of the online space	Cooperation with national authorities following orders	✓	✓	✓	✓
	Points of contact and, where necessary, legal representative	✓	✓	✓	✓
	Complainant and redress mechanism and out of court dispute settlement			✓	✓
	External risk auditing and public accountability				✓
	Crisis response cooperation				✓
Measures to counter illegal goods, services, or content online	Trusted flaggers			✓	✓
	Measures against abusive notices and counter-notices			✓	✓
	Vetting credentials of third-party suppliers ("KYBC")			✓	✓
	Reporting criminal offences			✓	✓
	Risk management obligations and compliance officer				✓
	Codes of conduct				✓
Access for researchers to key data	Data sharing with authorities and researchers				✓

Source: authors' elaboration adapted from the literature.

Source: "The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications," accessed June 19, 2023, https://www.tandfonline.com/doi/epdf/10.1080/17579961.2023.2184136?needAccess=true&role=button.

With the DSA and DMA, the EU's approach to combating disinformation on social media platforms has shifted from a voluntary approach prioritizing self-regulation to a binding coregulatory framework prioritizing transparency, and accountability which involves users, researchers and factcheckers in the process. The Commission is emerging as an important co-regulator in this domain. With the Digital Services Act Package, the Commission has acquired significant enforcement powers as a core regulator, with direct supervision and enforcement powers over VLOPs and can impose fines of up to 6% of global turnover on VLOPs, and has the right to temporarily suspend access to "rogue" platforms with a court order if they refuse to comply with the obligations and requirements of the Digital Services Act Package.[36]

The withdrawal of Elon Musk's Twitter from the EU Code of Practice on Disinformation is case in point.[37] Because even if Twitter has withdrawn from the voluntary Code of Practice, it will be obliged to comply with the regulations under the DSA and DMA on August 25, 2023. As the EU's internal market commissioner, Thierry Breton states : Twitter leaves EU voluntary Code of Practice against disinformation. But obligations remain. You can run but you can't hide. Beyond voluntary commitments, fighting disinformation will be legal obligation under #DSA as of August 25. Our teams will be ready for enforcement."[38].

36   "Questions and Answers: Digital Services Act," Text, European Commission - European Commission, accessed May 12, 2023, https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348.
37   "Twitter Pulls out of Voluntary EU Disinformation Code," BBC News, May 27, 2023, sec. Europe, https://www.bbc.com/news/world-europe-65733969.
38   "Thierry Breton on Twitter; 'Twitter Leaves EU Voluntary Code of Practice against Disinformation. But Obligations Remain. You Can Run but You Can't Hide. Beyond Voluntary Commitments, Fighting Disinformation Will Be Legal Obligation under #DSA as of August 25. Our Teams Will Be Ready for Enforcement.' / Twitter," accessed June 19, 2023, https://twitter.com/ThierryBreton/status/1662194595755704321?s=20.

Turkiye: The 2020 Bill on Social Media Platforms

In 2020 the Turkish Parliament amended the law on Regulating Internet Publications and Combating Crimes Committed by Means of Such Publications. These changes entered into force on 1 October 2020. These amendments have significantly increased the control of government authorities over social media platforms. The amendments have introduced a licensing requirement have required major platforms above a certain threshold of users to have a representative in Turkiye. The new clauses have also setout new obligations on these platforms to to remove content and to share information with state authorities. Failure to comply can result in sanctions ranging from fines to complete blocking of the platform.[39]

The bill makes references to the German NetzDG law in this social platform law, and the government has argued that these amendments are quite similar to the laws passed in the EU.[40] Although provisions such as the removal of offensive content by court order and fines for non-compliance are similar to the German NetzDG, the NetzDG has much more lenient measures and does not include provisions such as blocking access to platforms, accessing the data of posters and filing lawsuits as in the Turkish law.[41] There are also specific challenges to Turkiye in view of shortcomings related to the independence of courts as highlighted by the most recent European Commission Report on Turkiye.[42]

The Disinformation Law of 2022

The Turkish Parliament adopted a specific Disinformation Law in October 2022. Accordingly posting false information with the intention of spreading fear and panic or endangering the security of the country, public order or public health has been criminalized. Irrespective of the original producers of the content, the law also criminalizes the dissemination of misinformation. The broad and vague nature of the definition of disinformation is a major problem.[43] In a polarized country like Turkiye, where different media outlets can make completely opposite interpretations of the same event at the same time, determining what is true and what is false information is a challenge. Moreover, providing legal certainty for how courts are to be interpret the concepts of endangering national security or public order will be particularly challenging. As a result, it has been claimed that the law will have a chilling effect on online freedoms, increasing self-censorship on social media.[44]

39   "Turkish Law Tightening Rules on Social Media Comes into Effect," euronews, October 1, 2020, https://www.euronews.com/my-europe/2020/10/01/turkish-law-tightening-rules-on-social-media-comes-into-effect.

40   "What's Behind Turkiye's New Internet Law? | Heinrich Böll Stiftung," Heinrich-Böll-Stiftung, accessed June 18, 2023, https://tr.boell.org/en/node/21327.

41   "Blog: Regulating Disinformation and Social Media Platforms 'Alla Turca,'" ARTICLE 19, May 19, 2022, https://www.article19.org/resources/regulation-of-social-media-platforms-in-Turkiye-internet-law/.

42   European Commission. "Key findings of the 2022 Report on Türkiye". 12 October 2022   https://ec.europa.eu/commission/presscorner/detail/en/country_22_6088

43   Berk Esen. "Turkiye's New Disinformation Law: An Alarming Trend Towards Cyber-Authoritarianism," Stiftung Wissenschaft und Politik (SWP), accessed June 18, 2023, https://www.swp-berlin.org/publikation/Turkiyes-new-disinformation-law-an-alarming-trend-towards-cyber-authoritarianism.

44   Ibid.

These laws and regulations described and detailed above show that the Turkish government prefers to take a direct interventionist approach when it comes to combating disinformation on social media platforms. In that respect, Turkish authorities have diverged from European practices which tend to prioritize the establishment of standards for content moderation and anti-disinformation methods on social media platforms. In addition to the stated laws, there is also ongoing work to adopt a DMA like legislation. There are as of yet no drafts shared with the public, but it will be important for this potentially new piece of legislation to restrain from creating a more unpredictable regulatory framework regarding disinformation given that existing laws have already allocated responsibilities to different government agencies but also to courts.

CHAPTER 3: SELF-REGULATION ON MAJOR SOCIAL PLATFORMS: APPROACHES, METHODS AND CHALLENGES

A general overview of self-regulation models and the methods adopted is necessary for a comprehensive understanding and contextualization of the fight against disinformation on social media platforms. This section will examine some of the self-regulation methods and practices adopted by Google, TikTok, Meta and Twitter on their platforms. It should be noted that self-regulations are in a constant state of flux, and in the case of Twitter in particular, have experienced many different changes over the course of 2023.

Platforms have adopted various models and methods of self-regulation over a long period of time due to the different structures of platforms, the different purposes users use them for, and the different content they share on these platforms. However it should be noted that, even though the existence and implementation of these regulations have been announced and reported by the platforms, the extent to which they are implemented in practice has been questioned. Various studies and analyses have shown that up to 59% of content rated as false by factcheckers were not removed or labeled by Twitter, and this rate is between 20%-30% on platforms such as YouTube and Facebook. Studies have found that on Facebook only 16 percent of health-related information had a warning label, and that there were large differences between the languages used on the platform in terms of labeling problematic content.[45]

The set of self-regulations that the social media platforms adopt can take various forms. One common method is to perform content moderation with fact-checking, applying sanctions such as adding or removing a label to the content after a period of evaluation. This method can be inefficient in preventing disinformation from spreading to too many people, as it requires a lot of time to assess whether the content is disinformation or not. Therefore, instead of dealing with each piece of content individually, another approach is to reduce the algorithmic amplification of the topic related to misinformation or disinformation, rather than the specific shared content in order to limit the dissemination of disinformation to a wider audience.[46] However, a weakness of this approach is that it can only be applied if the information is completely true or false. Since most disinformation posts are misleading and partially inaccurate, it is not a comprehensive approach to categorize information as either black or white. Therefore, a different approach to such content is to add labels that advise the users to be careful and warn them that they are potentially encountering false or misleading information.[47]

45   Nandita Krishnan et al., "Research Note: Examining How Various Social Media Platforms Have Responded to COVID-19 Misinformation," Harvard Kennedy School Misinformation Review, December 15, 2021, https://doi.org/10.37016/mr-2020-85.

46   Joseph B. Bak-Coleman et al., "Combining Interventions to Reduce the Spread of Viral Misinformation," Nature Human Behaviour 6, no. 10 (October 2022): 1372–80, https://doi.org/10.1038/s41562-022-01388-6.

47   Ibid.

Different platforms have implemented such methods to combat covid-related disinformation, especially during the Covid era. For example, Meta platforms Instagram and Facebook used credibility labels based on the ratings of fact-checkers, indicating whether the content shared was true or false. Similar to other platforms, TikTok uses "contextual labels" and adds labels with contextual information that users can access to posts that may be misleading if not mentioned in the content itself.

For content that is not blatantly false and based on truth but is potentially misleading, approaches such as Facebook's downranking are used to reduce the likelihood of other users seeing it.[48] Most of the regulations discussed so far have opted to interfere with, label or algorithmically reduce the visibility of content that is already shared and posted on the platform. However, the regulations imposed by social media platforms are not limited to these. While Google acknowledges that disinformation is a significant challenge, it also argues that it is not in a position to broadly confirm what is true and what is false. Google uses ranking algorithms to promote official or respected sources and high-quality information. Additionally, Google is tightening its regulations on monetization to prevent financial gain, an important aspect of disinformation.[49] According to Google's statements and reports, algorithms determine which content is truly authoritative and reputable without human interference (which could be subjective and ideologically motivated). Instead these algorithms "focus on measurable signals that correlate with how users and other websites value the expertise, trustworthiness, or authoritativeness of a webpage on the topics it covers."[50].

	Facebook	Instagram	YouTube	Twitter	Linked-In	Snapchat	Reddit	TikTok	Twitch	Tumblr	Messenger	WhatsApp	Total (N = 12)
Labels, warnings, notifications, & links	✓	✓	×	✓	×	×	✓	✓	×	×	✓¹	✓¹	7
Decreasing visibility & spread	✓	✓	✓	✓	×	×	✓	✓	×	×	✓²	✓²	8
Content removal	✓³	✓	✓	✓	✓	✓*	✓³	✓	×	×	×	×	8
Account suspension/ban	✓	✓	✓	✓	×	×	✓*	✓⁴	×	×	×⁵	✓	7
Ad restrictions	✓	✓	✓	✓	✓	✓*	×	✓	×	×	N/A	N/A	7

Notes: ✓ indicates presence and × indicates absence of remedy; N/A indicates remedy not applicable. ¹Messenger and WhatsApp apply forwarding labels to indicate to recipients that the message was not written by the sender; WhatsApp also allows users to fact check the content of viral messages through Google search. ²Messenger and WhatsApp apply forwarding limits, which limit the number of people or groups a message can be forwarded to at one time. ³Misinformation in Facebook groups and Reddit communities can also be removed by moderators. ⁴Applies to users who violate any of TikTok's Community Guidelines and Terms of Service, including their misinformation policies. ⁵Because of how Messenger is connected to Facebook, Facebook account bans may apply to Messenger; however, this was not explicitly mentioned. *This remedy was identified in the updated rapid review conducted on November 23, 2021.

Table 2. Remedies to Address COVID-19 Misinformation Across Leading Social Media and Messaging Platforms

Source: Nandita Krishnan et al., "Research Note: Examining How Various Social Media Platforms Have Responded to COVID-19 Misinformation," Harvard Kennedy School Misinformation Review, December 15, 2021, https://doi.org/10.37016/mr-2020-85.

Compared to Google, which aims to promote accurate and reliable information and emphasizes the use of algorithms, apart from utilizing algorithms, TikTok adopts an approach that also involves users more in the fight against disinformation.

48   Emily Saltz et al., "Misinformation Interventions Are Common, Divisive, and Poorly Understood," Harvard Kennedy School Misinformation Review, October 27, 2021, https://doi.org/10.37016/mr-2020-81.
49   "How Google Fights Disinformation". Google. February 2019.
50   "How Google Fights Disinformation". Google. February 2019. 11.

TikTok encourages citizen-led debunking to combat disinformation on its platform. With two editing features on the platform, users can debunk disinformation videos by splitting the video screen in two or adding another video to the beginning of the video. With this method, they have introduced a method where platform users can respond to disinformation that random actors can spread by taking videos.[51] The highly decentralized nature of this citizen-led debunking could raise concerns since it could produce chaotic results. However, this method of combating such user-generated and viral disinformation videos with viral debunking videos can be stemming from TikTok's structural differences. While fact-checking and monitoring are quite difficult on a platform that is entirely video-based, and it is relatively easier for the posts of accounts with low follower numbers to go viral on TikTok than on other platforms,[52] the adoption of such methods can be due to the main characteristics of the platform. Therefore, the specific structure of platforms should be taken into account as an important factor in analyzing self-regulation models.

CONCLUSION

As this study has demonstrated, combating disinformation is an ever changing and arduous task that will require a close degree of public private partnership. Fundamentally, the assessment of which specific content can actually be categorized as disinformation will always be challenging. This challenge is set to become even more complex with the advent of AI backed technologies that will for instance allow malign actors to create deep fakes. Given that the establishment of a "Ministry of Truth" cannot be envisaged in democracies, other methods will have to be devised. At first, the large social media platforms were expected to develop their own internal rules and regulations. From a public policy standpoint, this lighter approach is warranted. The EU Code of Practice is a reflection of this reality.

And yet, the methods devised by the platforms have not, in the absence of a more binding regulation, been able to address the concerns of policy makers. Efforts undertaken by these large platforms should not be underestimated but clearly they remained insufficient given the scale of the challenge. As a result, policy makers in democratic societies have had to re-consider their stance. As a jurisprudence in the vanguard of digital legislation, the EU has moved ahead with a co-regulatory framework with the Strengthened EU Code of Practice and the adoption of the DSA and DMA. While this framework includes legally binding requirements for social media platforms, these requirements are generally aimed at improving the standards of social media platforms' internal policies, increasing their transparency and ensuring that stakeholders who are part of the community (users, researchers, fact-checkers) participate more effectively in the process,

51   Puneet Bhargava et al., "How Effective Are TikTok Misinformation Debunking Videos?," Harvard Kennedy School Misinformation Review, March 29, 2023, https://doi.org/10.37016/mr-2020-114.
52   geert, "TikTok: The New Frontier in Disinformation," Tilt (blog), August 31, 2022, https://www.tiltstudio.co/tiktok-the-new-frontier-in-disinformation/.

and introducing more stringent monitoring measures to monitor and evaluate whether these obligations are being met. It remains to be seen how effective this new approach shall be.

From the standpoint of public policy, combating disinformation is likely to involve an era of regulatory trial and error. The risk however is for countries to misalign their public policy objectives in terms of securing the right balance between the freedom of expression and containing disinformation. This risk is especially palpable for countries where the rule of law is weaker and the independence of regulatory authorities as well as the judiciary is weaker. An over reliance on judicial procedures would not only hinder online freedoms but also create an environment propitious for self censorship. The aim for lawmakers intent on fostering a more effective environment to combat disinformation should always be to create an inclusive resilient ecosystem, both in terms of empowering researchers, fact-checkers, users, and in terms of identifying best practices and measuring the effectiveness of the impact of the internal policies adopted by social media platforms. From that perspective, the approach adopted by Turkish authorities has been overly skewed to the direction of strict law enforcement. The ongoing work on the adoption of a DMA like legislation could however provide an opportunity to engage in a public deliberation on how Turkiye should revise its strategy to combat online disinformation.

BIBLIOGRAPHY

1) "2018 Code of Practice on Disinformation | Shaping Europe's Digital Future," June 16, 2022. https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation.

2) "2022 Strengthened Code of Practice on Disinformation | Shaping Europe's Digital Future," June 16, 2022. https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation.

3) Akyavas, Renan. "Turkiye: First Journalist Sentenced under New 'Disinformation' Law." Free Turkiye Journalists (blog). Accessed June 18, 2023. https://freeTurkiyejournalists.ipi.media/Turkiye-first-journalist-sentenced-under-new-disinformation-law/.

4) "Annual Self-Assessment Reports of Signatories to the Code of Practice on Disinformation 2019 | Shaping Europe's Digital Future," October 29, 2019. https://digital-strategy.ec.europa.eu/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019.

5) ARTICLE 19. "Blog: Regulating Disinformation and Social Media Platforms 'Alla Turca,'" May 19, 2022. https://www.article19.org/resources/regulation-of-social-media-platforms-in-Turkiye-internet-law/.

6) Bak-Coleman, Joseph B., Ian Kennedy, Morgan Wack, Andrew Beers, Joseph S. Schafer, Emma S. Spiro, Kate Starbird, and Jevin D. West. "Combining Interventions to Reduce the Spread of Viral Misinformation." Nature Human Behaviour 6, no. 10 (October 2022): 1372–80. https://doi.org/10.1038/s41562-022-01388-6.

7) BBC News. "Twitter Pulls out of Voluntary EU Disinformation Code." May 27, 2023, sec. Europe. https://www.bbc.com/news/world-europe-65733969.

8) Bhargava, Puneet, Katie MacDonald, Christie Newton, Hause Lin, and Gordon Pennycook. "How Effective Are TikTok Misinformation Debunking Videos?" Harvard Kennedy School Misinformation Review, March 29, 2023. https://doi.org/10.37016/mr-2020-114.

9) Center for American Progress. "Turkiye's Changing Media Landscape," June 10, 2020. https://www.americanprogress.org/article/Turkiyes-changing-media-landscape/.

10) DataReportal – Global Digital Insights. "Digital 2023: Turkiye," February 13, 2023. https://datareportal.com/reports/digital-2023-Turkiye.

11) "Dimensions of Polarization in Turkiye | Strengthening Transatlantic Cooperation." Accessed June 17, 2023. https://www.gmfus.org/news/dimensions-polarization-Turkiye.

12) "Dimensions of Polarization in Turkiye 2017 - TurkuazLab - Strategies and Tools for Mitigating Polarization in Turkiye Project," November 12, 2020. https://www.turkuazlab.org/en/dimensions-of-polarization-in-Turkiye-2017/.

13) euronews. "Turkish Law Tightening Rules on Social Media Comes into Effect," October 1, 2020. https://www.euronews.com/my-europe/2020/10/01/turkish-law-tightening-rules-on-social-media-comes-into-effect.

14) European Commission - European Commission. "Questions and Answers: Digital Services Act." Text. Accessed May 12, 2023. https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348.

15) geert. "TikTok: The New Frontier in Disinformation." Tilt (blog), August 31, 2022. https://www.tiltstudio.co/tiktok-the-new-frontier-in-disinformation/.

16) Heinrich-Böll-Stiftung. "What's Behind Turkiye's New Internet Law? | Heinrich Böll Stiftung | Niwêneriya Tirkiyeyê." Accessed June 18, 2023. https://tr.boell.org/en/node/21327.

17) "Thousands of Russian Twitter Accounts Reactivated in Turkish." Middle East Eye. Accessed June 18, 2023. http://www.middleeasteye.net/news/Turkiye-elections-thousands-russian-speaking-accounts-activated-twitter.

18) Jakubowski, Glenda. "What's Not to Like?: Social Media as Information Operations Force Multiplier," 2019.

19) Krishnan, Nandita, Jiayan Gu, Rebekah Tromble, and Lorien C. Abroms. "Research Note: Examining How Various Social Media Platforms Have Responded to COVID-19 Misinformation." Harvard Kennedy School Misinformation Review, December 15, 2021. https://doi.org/10.37016/mr-2020-85.

20) Newman, Nic, Richard Fletcher, Kirsten Eddy, Craig T Robertson, and Rasmus Kleis Nielsen. "Reuters Institute Digital News Report 2023," 2023.

21) POLITICO. "Turkiye's Social Media Law: A Cautionary Tale," March 29, 2021. https://www.politico.eu/article/Turkiyes-social-media-law-a-cautionary-tale/.

22) Reuters Institute for the Study of Journalism. "Attitudes towards Algorithms and Their Impact on News." Accessed June 17, 2023. https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/attitudes-towards-algorithms-impact-news.

23) Reuters Institute for the Study of Journalism. "Digital News Report 2018 - Turkiye Supplementary Report." Accessed June 17, 2023. https://reutersinstitute.politics.ox.ac.uk/our-research/digital-news-report-2018-Turkiye-supplementary-report.

24) Reuters Institute for the Study of Journalism. "Overview and Key Findings of the 2023 Digital News Report." Accessed June 15, 2023. https://reutersinstitute.politics. ox.ac.uk/digital-news-report/2023/dnr-executive-summary.

25) "Roadmaps to Implement the Code of Practice on Disinformation | Shaping Europe's Digital Future," October 16, 2018. https://digital-strategy.ec.europa.eu/ en/news/roadmaps-implement-code-practice-disinformation.

26) Saltz, Emily, Soubhik Barari, Claire Leibowicz, and Claire Wardle. "Misinformation Interventions Are Common, Divisive, and Poorly Understood." Harvard Kennedy School Misinformation Review, October 27, 2021. https://doi.org/10.37016/mr-2020-81.

27) Stiftung Wissenschaft und Politik (SWP). "Turkiye's New Disinformation Law: An Alarming Trend Towards Cyber-Authoritarianism." Accessed June 18, 2023. https:// www.swp-berlin.org/publikation/Turkiyes-new-disinformation-law-an-alarming-trend-towards-cyber-authoritarianism.

28) "The 2022 Code of Practice on Disinformation | Shaping Europe's Digital Future," June 2, 2023. https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation.

29) "The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications." Accessed June 19, 2023. https://www.tandfonline.com/doi/ epdf/10.1080/17579961.2023.2184136?needAccess=true&role=button.

30) "The Digital Services Act (DSA)." Accessed June 19, 2023. https://www. eu-digital-services-act.com/.

31) "The Digital Services Act Package | Shaping Europe's Digital Future," June 13, 2023. https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.

32) "Thierry Breton on Twitter: 'Twitter Leaves EU Voluntary Code of Practice against Disinformation. But Obligations Remain. You Can Run but You Can't Hide. Beyond Voluntary Commitments, Fighting Disinformation Will Be Legal Obligation under #DSA as of August 25. Our Teams Will Be Ready for Enforcement.' / Twitter." Accessed June 19, 2023. https://twitter.com/ThierryBreton/status/166219459575 5704321?s=20.

33) "Türkiye | RSF," May 31, 2023. https://rsf.org/en/country-t%C3%BCrkiye.

34) York, Jillian C. "Turkiye's New Disinformation Law Spells Trouble For Free Expression." Electronic Frontier Foundation, November 7, 2022. https://www.eff. org/tr/deeplinks/2022/11/Turkiyes-new-disinformation-law-spells-trouble-free-expression.