



## **RESILIENCE:**

For Media Free of Hate and Disinformation

---

# **REGULATORY AND SELF-REGULATORY FRAMEWORKS AGAINST HATE SPEECH AND DISINFORMATION IN THE EUROPEAN UNION**

EXAMPLES AND CHALLENGES

**FACTSHEET**

## RESILIENCE : For Media Free of Hate and Disinformation

The regional project 'RESILIENCE: Civil society action to reaffirm media freedom and counter disinformation and hateful propaganda in the Western Balkans and Turkey' is implemented with the financial support of the European Union by partner organizations [SEENPM](#), the [Albanian Media Institute](#), [Mediacentar Sarajevo](#), [Kosovo 2.0](#), the [Montenegro Media Institute](#), the [Macedonian Institute for Media](#), the [Novi Sad School of Journalism](#), the [Peace Institute](#) and [Bianet](#).

Info: <https://seenpm.org/>

Contact: [admin@seenpm.org](mailto:admin@seenpm.org)

## REGULATORY AND SELF-REGULATORY FRAMEWORKS AGAINST HATE SPEECH AND DISINFORMATION IN THE EUROPEAN UNION – EXAMPLES AND CHALLENGES

### FACTSHEET

The factsheet compiled by: Brankica Petković and Barbara Novak

Editor, regional lead researcher: Brankica Petković

Language editor: Fiona Thompson

Design: Špela Kranjec for Filip Kranjec s.p., Ljubljana, Slovenia

Publishers: Peace Institute, Ljubljana and SEENPM, Tirana

Ljubljana, November 2021

© *Peace Institute and SEENPM*



South East European Network  
for Professionalization of Media



*This publication was produced with the financial support of the European Union.  
Its contents are the sole responsibility of the publishers  
and do not necessarily reflect the views of the European Union.*

# CONTENTS

<b>1. INTRODUCTION</b>	<b>7</b>
<b>2. EU: Initiative for a Digital Services Act (2020)</b>	<b>9</b>
2.1. Challenges and critical views	11
<b>3. EU: Audiovisual Media Services Directive (2018)</b>	<b>14</b>
3.1. Challenges and critical views	15
<b>4. EU: Code of Conduct on countering illegal hate speech online (2016)</b>	<b>16</b>
4.1. Challenges and critical views	17
<b>5. EU: The Code of Practice on Disinformation (2018)</b>	<b>18</b>
5.1. Challenges and critical views	20
<b>6. EU: European Digital Media Observatory, EDMO (2019)</b>	<b>21</b>
<b>7. GERMANY: Network Enforcement Act, NetzDG (2018)</b>	<b>22</b>
7.1. Challenges and critical views	24
<b>8. AUSTRIA: Communication Platforms Act (2021)</b>	<b>26</b>
8.1. Challenges and critical views	27
<b>9. SLOVENIA: Spletno oko hotline for reporting hate speech online (2006)</b>	<b>28</b>
9.1. Challenges and critical views	30
Literature and sources	31

## FACTSHEET

# REGULATORY AND SELF-REGULATORY FRAMEWORKS AGAINST HATE SPEECH AND DISINFORMATION IN THE EUROPEAN UNION

## EXAMPLES AND CHALLENGES

### 1. INTRODUCTION

The wide recognition of the urgent need to address the increasing threats of online hate speech and disinformation to democracy and human rights is reflected in the number of recent regulatory and policy initiatives by the European Union, and in individual member states.

Together with this urgent need, there is an enormous challenge to devise and adopt adequate and efficient regulatory and policy measures against online hate speech and disinformation, taking into account the complexity of the field – the internet content regulation. The risks of unduly excessive restrictions for freedom of expression and of privatization of law enforcement by IT companies are particular issues of concern.

Some of the instruments adopted in recent years take the form of laws and rely on the authorities to provide law enforcement, while others have introduced self-regulatory and accountability mechanisms for online platforms relying on their own enforcement of the rules. At the same time, various EU-funded projects have been developed to complement the effort with research, monitoring, education, reporting, EU-wide networking and collaboration to counter hate speech and disinformation with knowledge-sharing, innovative tools, and informing political debates and decisions on regulation.

On the level of the European Union, the expected legal basis for regulation in this field is Article 114 of the Treaty on the Functioning of the European Union which requires a sound demonstration that the regulatory initiative is needed for ensuring the functioning of the internal market.<sup>1</sup> Therefore, the

---

<sup>1</sup> Giuseppe B. Abbamonte, "Book Review: Giovanni Pitruzzella & Oreste Pollicino, *Disinformation and hate speech: a European constitutional perspective*", *EU Law Live*, 15 September 2020, <https://eulawlive.com/library/disinformation-and-hate-speech-a-european-constitutional-perspective/>.

subject of regulation are “services”, either audiovisual media services or digital services. There are, however, important non-economic objectives of the regulation of media and digital “services”, arising from their value for democracy and human rights and for fulfilling social and cultural functions of media and digital communication.

On the level of EU member states, there have recently been several regulatory and policy initiatives to regulate online hate speech and disinformation.

In this factsheet we present several examples of the EU level (initiatives for) regulation as well as self-regulatory mechanisms aimed at countering online hate speech and disinformation. Examples of regulation or self-regulation on the level of several member states will also be presented.

The brief presentation of examples of the regulatory and self-regulatory mechanisms in the European Union complements the factsheets with descriptions of national regulatory and self-regulatory frameworks against hate speech and disinformation in the countries of the Western Balkans and Turkey. The factsheets aim at informing the policy discussions in these countries within the regional project Resilience – Civil Society for Media Free of Hate and disinformation. They follow the series of the national reports and regional overviews on hate and propaganda models of media and communication, on hate narratives in online media and communication, and on media trust, produced within the Resilience project.<sup>2</sup>

The overview compiles information about the regulatory and self-regulatory documents and mechanisms in the European Union and their explanation by various legal experts. It includes also a brief presentation of critical views and challenges for each example of regulation or self-regulation, as elaborated by legal experts and advocacy groups.

---

<sup>2</sup> See the research reports at the Resilience project web page: <https://seenpm.org/resilience-research/>.

## 2. EU: Initiative for a Digital Services Act (2020)

The Digital Services Act (DSA)<sup>3</sup> is a legislative proposal by the European Commission<sup>4</sup> aiming to upgrade the current regulation of digital services in the EU,<sup>5</sup> particularly regarding illegal content (including illegal hate speech online), transparent advertising and disinformation.

Digital services range from simple websites to internet infrastructure services and online platforms. Illegal hate speech is defined in EU law under the Framework Decision on combating certain forms and expressions of racism and xenophobia by means of criminal law<sup>6</sup> as the public incitement to violence or hatred directed at groups or individuals on the basis of certain characteristics, including race, colour, religion, descent and national or ethnic origin.

The legislative proposal was submitted for approval to the European Parliament and the European Council in December 2020. The adoption procedure is expected to last until mid-2022.

The proposed DSA is intended to improve content moderation on social media platforms to address concerns about illegal content. The DSA proposal maintains the current rule according to which companies that host others' data are not liable for the content unless they actually know it is illegal, but adds the exception that once illegal content is flagged, companies are required to remove it. The DSA would introduce new obligations on platforms to disclose to regulators how their algorithms work, on how decisions to remove content are taken and on the way advertisers target users. Many of its provisions only apply to platforms which have more than 45 million users in the European Union. Platforms including Facebook, Google's subsidiary YouTube, Twitter and TikTok would meet that threshold and be subjected to the new obligations. Companies that do not comply with the new obligations risk fines of up to 6% on their annual turnover.<sup>7</sup>

---

3 The Digital Services Act is one of two proposals of the Digital Services Act package. The second proposal in the package is the Digital Markets Act (DMA).

4 European Commission, Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, 15 December 2020, <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN>.

5 Such as the Directive on Electronic Commerce (European Parliament and the Council of the European Union, Directive on Electronic Commerce, 8 June 2000, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0031&from=EN>).

6 The Council of the European Union, Council Framework Decision 2008/913/JHA on combating certain forms and expressions of racism and xenophobia by means of criminal law, 28 November 2008, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32008F0913&from=en>.

7 Wikipedia, Digital Services Act, [https://en.wikipedia.org/wiki/Digital\\_Services\\_Act](https://en.wikipedia.org/wiki/Digital_Services_Act).

The international law firm Allen & Overy summarized the main provisions of the DSA in more details as follows:

- Modernized liability regime for online intermediaries. The key principles from the e-Commerce Directive remain generally unchanged, but the DSA adds obligations to address notifications of content considered illegal. The DSA requires every hosting provider or online platform to put in place a user-friendly notice and takedown mechanisms that allow the notification of illegal content. Online platforms will need to establish internal complaint-handling systems, engage with out-of-court dispute settlement bodies to resolve disputes with their users, give priority to notifications of entities that have been qualified as so-called trusted flaggers by the authorities and suspend repeat infringers.
- New and far-reaching transparency obligations for online platforms relating to the measures taken to combat illegal information. If content is removed, an explanation needs to be provided to the person who uploaded that content. Online platforms must also publish detailed reports on their activities relating to the removal and the disabling of illegal content or content contrary to their terms and conditions.
- Obligation on online intermediaries to include in their terms and conditions information on any restrictions on the use of data provided by the users, with reference to the content moderation mechanisms applied, algorithmic decision-making and human review. This information must be in clear and unambiguous language and publicly available in an easily accessible format.
- Transparency obligations concerning online advertisements. For each advertisement and to each user, the online platforms must provide, in real time, clear and unambiguous information to users that a) they are seeing an advertisement, b) on whose behalf the ad is displayed, and c) provide meaningful information about the main parameters used to determine why a specific user is targeted by this ad.
- Steep fines for non-compliance of up to 6% of the annual income or turnover of the provider of intermediary services and periodic penalty payments for continuous infringements of up to 5% of the average daily turnover of the intermediary in the preceding financial year per day.
- Online intermediaries without establishment in the EU that provide services in the EU must designate a legal representative in the EU who will be required to cooperate with supervisory authorities, the European Commission and the European Board for Digital Services (a new pan-European group of coordinators that will assist with the harmonization of the DSA) and can be held liable for non-compliance with the DSA.

In addition to the rules set out above, very large platforms must also comply with the rules set out below. Very large online platforms are those platforms which have more than 45 million active monthly users in the EU, and they will have to:

- Analyze any systemic risk stemming from the use of their platforms and put in place effective content moderation mechanisms to address the identified risks (e.g. illegal content, privacy violations, etc).
- Provide transparency on the main parameters of the decision-making algorithms used to offer content on their platforms (the rankings mechanism) and the options for the user to modify those parameters. They must provide an option that is not based on profiling.
- Establish and maintain a public repository, available via application programming interfaces, with detailed information on the online advertisements they served on their platforms in the past year.
- An obligation to designate a dedicated compliance officer responsible for compliance with obligations under the DSA and undergo an annual independent audit.
- Upon request of the competent authority, very large online platforms must also give access to the data necessary to monitor their compliance with the DSA to the competent authority but also to vetted academic researchers that perform research into the systemic risks.
- In addition, the European Commission will have supervisory and enforcement powers in relation to very large platforms.<sup>8</sup>

## 2.1. Challenges and critical views

While acknowledging the efforts of the European Commission to create – with the new regulation – a safer digital space in which the fundamental rights of all users of digital services are protected, the advocacy groups specialized in defending rights and freedoms in the digital environment have expressed concerns about the implications of the proposed act for the right to freedom of expression.

The European Digital Rights network of civil and human rights organizations from across Europe (EDRi) criticizes the ‘delete first, think later’ principle introduced by the DSA proposal. They advocate additional safeguards in the DSA to prevent that a situation where “a system of privatized content control with arbitrary rules is established by online platforms and services beyond judicial and democratic scrutiny”.<sup>9</sup>

<sup>8</sup> “The Digital Services Act package is here,” Allen & Overy, 16 December 2020, <https://www.allenoverly.com/en-gb/global/news-and-insights/publications/the-digital-services-act-package-is-here>.

<sup>9</sup> Jan Penfrat, “The EU’s attempt to regulate Big Tech: What it brings and what is missing,” EDRi, 18 December 2020, <https://edri.org/our-work/eu-attempt-to-regulate-big-tech/>.



*As soon as anybody on the internet flags any content as potentially illegal, liability kicks in and would require the hosting company to “expeditiously” remove or disable access to the content in question. Removing or disabling content that has been flagged therefore becomes the most commercially reasonable action for companies in order to avoid the legal liability risk.<sup>10</sup>*

With the purpose of advocating the improved proposal of the Digital Services Act, recognizing the position of the European Union as a global trendsetter in internet legislation, civil society organizations from across the world formed the Digital Services Act Human Rights Alliance in May 2021. They describe their purpose as “to support the EU in upholding transparency and accountability, and establishing and promoting a world standard for platform governance”. The coalition includes Access Now, the Center for Democracy and Technology, Civil Liberties Union for Europe and other organizations.

In their Joint statement from October 2021<sup>11</sup> they highlighted the most important concerns. They urge: a) to avoid disproportionate demands on smaller providers that would put users’ access to information in serious jeopardy; b) to avoid legally mandated strict and short time frames for content removal due to their detrimental impact on the right to freedom of expression and opinion; c) not to impose legally mandated automated content moderation tools on online platforms, as this will lead to over-removals of legitimate speech; d) to avoid shifting states’ obligations to protect individuals’ rights to privately owned online platforms, thus allowing them to act as quasi-judicial bodies in the online ecosystem without any public scrutiny; e) not to impose mandatory reporting obligations on Law Enforcement Agencies (LEAs), especially without appropriate safeguards and transparency requirements; f) to prevent public authorities, including LEAs, from becoming trusted flaggers. Conditions for becoming trusted flaggers need to be subject to regular reviews and proper public oversight; g) to consider mandatory human rights impact assessment as the primary mechanism for examining and mitigating systemic risks stemming from platforms’ operations.<sup>12</sup>

Furthermore, the legal experts of Council of Europe<sup>13</sup> question what qualifies as “illegal content” in the DSA – whether it refers to content that has already been declared illegal by a relevant authority or at least has already been the object of specific measures under the provisions of the DSA, or rather points to the foreseeability that still-to-be-produced illegal information could end up being disseminated via the mentioned platforms. They also see the need for safeguards aimed at avoiding unnecessary and disproportionate impacts on

---

<sup>10</sup> *Ibid.*

<sup>11</sup> Human Rights Alliance, *Joint Statement of the Digital Services Act Human Rights Alliance*, 21 October 2021, [https://www.accessnow.org/cms/assets/uploads/2021/10/Digital\\_Services\\_Act\\_Human\\_Rights\\_Alliance\\_Statement.pdf](https://www.accessnow.org/cms/assets/uploads/2021/10/Digital_Services_Act_Human_Rights_Alliance_Statement.pdf).

<sup>12</sup> *Ibid.*

<sup>13</sup> European Audiovisual Observatory, *Unravelling the Digital Services Act package*, IRIS Special, 2021, <https://rm.coe.int/iris-special-2021-01en-dsa-package/1680a43e45>.

the exercise of the right to freedom of expression by users and third parties (neither by platforms themselves nor oversight bodies). The DSA provision on “intentional manipulation of ... service, including by means of inauthentic use or automated exploitation of the service, with an actual or foreseeable negative effect on the protection of public health, minors, civic discourse, or actual or foreseeable effects related to electoral processes and public security”, according to human rights law experts, clearly show that platforms may face the legal responsibility (overseen by public bodies) to restrict access to lawful content (and therefore protected under the freedom of expression clause) which can be considered “harmful” under the very vague criteria. These criteria are subjected to very open interpretations that are dependent on largely different political approaches and sensitivities within the European Union.<sup>14</sup>

In autumn 2021, a possibility to introduce an amendment on “media exemption” to the proposed Digital Service Act was made public. It would forbid platforms from “removing”, “disabling access to”, “suspending” or “interfering with” content by press publications or “editorial content providers”. In other words, any organization that publishes regularly on a topic with the intention to provide information to the public would be exempted from moderation or as otherwise designated exempting “the media”.<sup>15</sup> This prompted a call from journalists, fact-checkers and disinformation researchers to the members of the European Parliament to reject such an amendment as it “threatens to be a step backwards from the status quo, reversing years of progress in the fight against hate speech and disinformation online.”<sup>16</sup>

---

14 Joan Barata, “The Digital Services Act and social media power to regulate speech: obligations, liabilities and safeguards,” in *Unravelling the Digital Services Act package*, IRIS Special 2021, European Audiovisual Observatory, 2021, 5-19.

15 Diana Wallis, “Digital Services Act: a media exemption would open a loophole for disinformation”, Euroactiv, 20 October 2021, <https://www.euractiv.com/section/digital/opinion/digital-services-act-a-media-exemption-would-open-a-loophole-for-disinformation/>.

16 Disinfo.eu, Fact-checkers and experts call on MEPs to reject a media exemption in the DSA, 5 November 2021, <https://www.disinfo.eu/advocacy/fact-checkers-and-experts-call-on-meps-to-reject-a-media-exemption-in-the-dsa/>.

### 3. EU: Audiovisual Media Services Directive (2018)

In addition to prohibition of incitement to hatred based on race, sex, religion or nationality in audiovisual media services (television broadcasting or on-demand), specified already in the 2010 Audiovisual Media Services Directive,<sup>17</sup> the revised Audiovisual Media Services Directive (2018)<sup>18</sup> extends the definition of the grounds for hatred to include “any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation” (referring to Article 21 of the Charter of Fundamental Rights of the European Union).

Furthermore, the 2018 AVMSD extends the scope of the regulation – in addition to television broadcasting and on-demand audiovisual media services – to one category of online platforms, the video-sharing platforms. According to the Directive, the video-sharing platforms should be subject to appropriate and proportionate measures by the EU member states, preferably through co-regulation, in order to protect the general public from illegal content such as hate speech (incitement to violence or hatred). Those measures must be appropriate in the light of the nature of the content, the category of persons to be protected and the rights and legitimate interests at stake and be proportionate taking into account the size of the platforms and the nature of the provided service.<sup>19</sup>

As emphasized by Vidgen, Burden and Margetts from the Alan Turing Institute, these provisions bring the video-sharing online platforms in line with existing requirements for traditional television broadcasting and video-on-demand platforms.<sup>20</sup> The 2018 AVMSD establishes minimal requirements for online platforms to remove illegal online hate, and creates space for them to tackle hate that is legal but violates their Terms of Service. As such, it is likely to mostly affect the smaller and less well-moderated platforms rather than bigger platforms which already address this content.<sup>21</sup>

In their report commissioned by the UK regulator Ofcom in the context of the requirements of the revised Audiovisual Media Services Directive for video sharing platforms, Vidgen, Burden and Margetts suggest four key steps for creating a moderation system to tackle online hate:

---

17 European Parliament and the Council of the European Union, *Audiovisual Media Services Directive*, 10 March 2010, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32010L0013&from=EN>.

18 European Parliament and the Council of the European Union, *Audiovisual Media Services Directive*, 14 November 2018, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32018L1808&from=SL>.

19 Joan Barata, *Regulating content moderation in Europe beyond the AVMSD*, 25 February 2020, <https://blogs.lse.ac.uk/mediase/2020/02/25/regulating-content-moderation-in-europe-beyond-the-avmsd/>.

20 Bertie Vidgen, Emily Burden and Helen Margetts, *Facing the challenge of online hate: the AVMSD and the role of regulation*, 30 March 2021, <https://blogs.lse.ac.uk/mediase/2021/03/30/facing-the-challenge-of-online-hate-the-avmsd-and-the-role-of-regulation/>.

21 *Ibid.*

a) Characterize online hate

First, platforms need to provide a clear account of online hate, clearly establishing where the line falls between hate and non-hate

b) Identify online hate

Second, platforms need to develop strategies to identify hate. Three planks form the basis of most content-moderation processes for identifying online hate: user reports, Artificial Intelligence (AI) and human review.

c) Handle online hate

Third, content identified as hateful needs to be handled with an intervention. Public discourse often focuses on the effects of bans but, in practice, platforms use many other interventions. Vidgen, Burden and Margetts identified 14 moderation strategies available to video-sharing platforms,<sup>22</sup> each of which imposes different levels of friction on users' activity. These range from hosting constraints, such as banning or suspending users and their content, through to engagement constraints, such as limiting the number of times that users can like or comment on content. Imposing any sort of friction risks impinging on users' freedom of expression and privacy, and it is important that the degree of friction is always proportionate to the harm that is likely to be inflicted.

d) Enable users to appeal decisions

Fourth, online platforms should create a robust and accessible review procedure so that users are able to challenge moderation decisions.<sup>23</sup>

### 3.1. Challenges and critical views

In the adoption procedure for the 2018 Audiovisual Media Services Directive, Civil Liberties Union for Europe (Liberties) expressed concerns that freedom of expression will be unduly restricted by the directive. They criticized the extended scope of the directive, suggesting that video-sharing platforms as well as on-demand audiovisual media services should be regulated rather by the E-commerce Directive (ECD). According to them, the ECD has a better-balanced regime and well-elaborated method for taking down problematic content. "The rules in place under the ECD make it more likely that limitations on free expression will be proportionate."<sup>24</sup> With regard to the co-regulatory and self-regulatory measures against hate speech, Liberties urged that "this would create an easy way to censor user content without the protection of legal safeguards and remedies. Creating the possibility to self-regulate and co-regulate without proper state regulation will result in unjustified restrictions."<sup>25</sup>

<sup>22</sup> Bertie Vidgen, Emily Burden, and Helen Margetts, *Understanding online hate: VSP Regulation and the broader context*, The Alan Turing Institute, February 2021, [https://www.ofcom.org.uk/\\_data/assets/pdf\\_file/0022/216490/alan-turing-institute-report-understanding-online-hate.pdf#page=84](https://www.ofcom.org.uk/_data/assets/pdf_file/0022/216490/alan-turing-institute-report-understanding-online-hate.pdf#page=84).

<sup>23</sup> *Ibid*

<sup>24</sup> Eva Simon, *Recommendations for the reform of the Audiovisual Media Services Directive*, Civil Liberties Union for Europe, September 2017, [https://drive.google.com/file/d/0B07ejuhu6mZ5Zk5aUGJyTmt5OWWM/view?resourcekey=0-1S2CLG5-68jk\\_2xCDkUkcQ](https://drive.google.com/file/d/0B07ejuhu6mZ5Zk5aUGJyTmt5OWWM/view?resourcekey=0-1S2CLG5-68jk_2xCDkUkcQ).

<sup>25</sup> *Ibid*.

## 4. EU: Code of Conduct on countering illegal hate speech online (2016)

The Code of Conduct on countering illegal hate speech online<sup>26</sup> is a self-regulatory mechanism established by the European Union to respond to the proliferation of racist and xenophobic hate speech online. The EU Code of Conduct was agreed on 31 May 2016 to ensure that requests to remove racist and xenophobic internet content are dealt with quickly by the major IT companies. The European Commission launched the Code of Conduct together with four major IT companies (Facebook, Microsoft, Twitter and YouTube). To date, ten companies have adhered to the Code: in addition to the first group, Instagram, Dailymotion and Snapchat also joined the mechanism in 2018, Jeuxvideo.com in 2019, TikTok in 2020 and LinkedIn in 2021.<sup>27</sup>

The implementation of the Code of Conduct is based on cooperation involving the European Commission, IT platforms, civil society organizations and national authorities.

How does it work? When the IT companies receive a request to remove from their online platform content deemed to be illegal, they assess this request against their rules and community guidelines and, where necessary, national laws transposing EU law on combating racism and xenophobia. The companies commit to reviewing the majority of these requests in less than 24 hours and to removing the content if necessary, while respecting the fundamental principle of freedom of speech.

Implementation of the Code is evaluated through a monitoring exercise by a network of civil society organizations located in different EU countries. Using a commonly agreed methodology, these organizations test how the IT companies apply the Code in practice. They do this by regularly sending the IT companies requests to remove content and recording how long it takes the IT companies to assess the request, what the outcomes are, and the feedback they receive from the companies.

Between 2016 and 2020, the Code of Conduct had delivered continuous progress: in 2020, on average the companies assessed 90% of flagged content within 24 hours and 71% of the content deemed illegal hate speech was removed.

However, the sixth evaluation of the Code of Conduct on Countering Illegal Hate Speech Online, conducted in 2021, shows that while the average of notifications reviewed within 24 hours remains high (81%), it has decreased compared to 2020 (90.4%). At 62.5% the average removal rate was also lower

<sup>26</sup> European Commission, *Countering illegal hate speech online*, <https://ec.europa.eu/newsroom/just/items/54300>.

<sup>27</sup> European Commission, *The Code of conduct on countering illegal hate speech online*, 22 June 2020, [https://ec.europa.eu/commission/presscorner/detail/en/qanda\\_20\\_1135](https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1135).

than in 2019 and 2020. However, broken down by IT company the progress of Instagram (66.2% removals in 2021, 42% in 2020) and Twitter (49.8% versus 35.9%) is noteworthy. TikTok was included in the evaluation for the first time and performed well (80.1% removals).<sup>28</sup>

The results of the Code of Conduct monitoring feed into the wider work of the European Commission on how online platforms should be more proactive in the prevention, detection and removal of illegal content, including the recent proposal of the Digital Services Act Package. The Commission also announced that it will continue to facilitate the dialogue between IT companies and civil society organizations working on the ground to tackle illegal hate speech.<sup>29</sup>

## 4.1. Challenges and critical views

Human rights law expert Barbora Bukovská, Senior Director for Law and Policy at the international non-governmental organization Article 19, believes that the Code of Conduct has a problematic legal basis and unclear process of implementation, and is thus a misguided policy on the part of the European Commission. She argues that “for companies, it is likely to amount to no more than a public relations exercise. Despite its nonbinding character, the Code can lead to more censorship by private companies – and thus undermine the rule of law and create a chilling effect on freedom of expression on the platforms they run”.<sup>30</sup>

Bukovská sees the key problem with the Code of Conduct in its normative basis for defining “illegal online content.” According to her assessment, the Code of Conduct puts companies – rather than the courts – in the position of having to decide the legality of content. It allows law enforcement to pressure companies to remove content in circumstances where the authorities do not have the power to order its removal because the content itself is legal. Importantly, the Code does not require the adoption of any safeguards against misuse of the notice procedure and is silent on remedies to challenge wrongful removals, Bukovská emphasizes.<sup>31</sup>

The 2021 evaluation report on the EU Code of Conduct on countering online hate speech revealing a significant decline of number of notifications reviewed and in the removal rate of online hate speech by the IT companies, is seen by the Centre for Democracy and Technology as confirmation of the “persistent

---

28 Didier Reynders, *Countering illegal hate speech online: 6th evaluation of the Code of Conduct*, European Commission, 7 October 2021, [https://ec.europa.eu/info/sites/default/files/factsheet-6th-monitoring-round-of-the-code-of-conduct-october2021\\_en\\_1.pdf](https://ec.europa.eu/info/sites/default/files/factsheet-6th-monitoring-round-of-the-code-of-conduct-october2021_en_1.pdf).

29 European Commission, *Commission publishes EU Code of Conduct on countering illegal hate speech online continues to deliver results*, 22 June 2020, [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_20\\_1134](https://ec.europa.eu/commission/presscorner/detail/en/IP_20_1134).

30 Barbora Bukovská, *Article 19: The European Commission’s Code of Conduct for Countering Illegal Hate Speech Online: An analysis of freedom of expression implications*, Transatlantic Working Group, 7 May 2019, <https://www.ivir.nl/publicaties/download/Bukovska.pdf>.

31 *Ibid*, pp. 6.



concerns regarding the European Commission's approach to tackling issues such as hate speech, disinformation, and online gender-based violence."<sup>32</sup>

At the same time, Euroactiv reported about the concerns of the European Commission that larger legislative efforts, such as the proposed Digital Services Act, could distract online platforms and overshadow their compliance with the voluntary commitments they have made. "[A] gentleman's agreement alone will not suffice here", said Věra Jourová, the Commission's vice-president for values and transparency, as reported by Euroactiv. "The Digital Services Act will provide strong regulatory tools to fight against illegal hate speech online."<sup>33</sup>

## 5. EU: The Code of Practice on Disinformation (2018)

The Code of Practice on Disinformation,<sup>34</sup> initiated by the European Commission, was published on 26 September 2018. It sets out principles and commitments for online platforms, leading social networks, advertisers and the advertising industry to counter the spread of disinformation online in the European Union, which its signatories agreed to implement.

Signing up to the Code is a voluntary decision of the platform. Within the self-regulatory framework, platforms are not obliged to implement any specific practice but agree to report their activities to the Commission. It became the world's first self-regulatory instrument to fight disinformation.

Current signatories involve major online platforms active in the EU, as well as trade associations and relevant players in the online and advertising ecosystems. They are: Google, Facebook, Twitter, Microsoft, TikTok, Mozilla, DOT Europe (Former EDiMA), the World Federation of Advertisers (WFA) and its Belgian counterpart, the Union of Belgian Advertisers (UBA); the European Association of Communications Agencies (EACA), and its national members from France, Poland and the Czech Republic – respectively, Association des Agences Conseils en Communication (AACC), Stowarzyszenie Komunikacji Marketingowej/Ad Artis Art Foundation (SAR), and Asociace Komunikacnich Agentur (AKA); the Interactive Advertising Bureau (IAB Europe), Kreativitet & Kommunikation, and Goldbach Audience (Switzerland) AG.<sup>35</sup>

---

32 Molly Killeen, "Progress stalls on Commission's online hate speech efforts", *Euractiv*, 7 October 2021, <https://www.euractiv.com/section/digital/news/progress-stalls-on-commissions-online-hate-speech-efforts/>.

33 *Ibid.*

34 European Commission, *Code of Practice on Disinformation*, 1 October 2021, <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>.

35 Current signatories involve major online platforms active in the EU, as well as trade associations and relevant players in the online and advertising ecosystems. They are: Google, Facebook, Twitter, Microsoft, TikTok, Mozilla, DOT Europe (Former EDiMA), the World Federation of Advertisers (WFA) and its Belgian counterpart, the Union of Belgian Advertisers (UBA); ...

In October 2021, eight new prospective signatories joined the revision process. Amongst them are online video platforms like Vimeo, new types of social networks like Clubhouse, and advertising technology providers like DoubleVerify, as well as organizations that provide specific expertise and technical solutions to fight disinformation, such as Avaaz, Globsec, Logically, NewsGuard, and WhoTargetsMe.<sup>36</sup>

The current Code has been considered a good first step, but the Commission's Assessment in 2020 revealed significant shortcomings.<sup>37</sup> These include inconsistent and incomplete application of the Code across platforms and member states, gaps in the coverage of the Code's commitments, a lack of appropriate monitoring mechanism, including key performance indicators, a lack of commitments on access to platforms' data for research on disinformation and limited participation from stakeholders, in particular from the advertising sector.<sup>38</sup>

In May 2021, the Commission issued Guidance to strengthen the Code of Practice on Disinformation in order to address the shortcomings identified in the Commission's 2020 Assessment of the Code<sup>39</sup> and draw from the lessons learned in the COVID-19 disinformation monitoring programme.<sup>40</sup>

The 2021 Guidance calls for reinforcing all chapters of the Code: the monetization of disinformation must be reduced, measures against manipulative techniques must be stepped up and users should have access to tools to understand and flag disinformation and safely navigate in the online environment. According to the Guidance, the Code should also increase the coverage of fact checking across all member states and languages, and drastically improve access to platforms' data for research. The signatories should also set up a publicly accessible Transparency Centre, and a Permanent Task Force will be in charge of adapting the Code in view of technological, societal, market and legislative developments. The reinforced Code needs to be complemented by an effective monitoring framework,

---

...the European Association of Communications Agencies (EACA), and its national members from France, Poland and the Czech Republic – respectively, ... Association des Agences Conseils en Communication (AACC), Stowarzyszenie Komunikacji Marketingowej/Ad Artis Art Foundation (SAR), and Asociace Komunikacnich Agentur (AKA); the Interactive Advertising Bureau (IAB Europe), Kreativitet & Kommunikation, and Goldbach Audience (Switzerland) AG.

36 European Commission, Code of Practice on disinformation: Commission welcomes new prospective signatories and calls for strong and timely revision, 1 October 2021, [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_21\\_4945](https://ec.europa.eu/commission/presscorner/detail/en/IP_21_4945).

37 European Commission, Disinformation: EU assesses the Code of Practice and publishes platform reports on coronavirus related disinformation, 10 September 2020, [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_20\\_1568](https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1568).

38 European Commission, Code of Practice on disinformation: Commission welcomes new prospective signatories and calls for strong and timely revision, 1 October 2021, [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_21\\_4945](https://ec.europa.eu/commission/presscorner/detail/en/IP_21_4945).

39 European Commission, Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement, 10 September 2020, <https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>.

40 European Commission, First baseline reports – Fighting COVID-19 disinformation Monitoring Programme, 10 September 2020, <https://digital-strategy.ec.europa.eu/en/library/first-baseline-reports-fighting-covid-19-disinformation-monitoring-programme>.



based on clear key performance indicators for measuring the efficacy of the actions implemented under the 2021 Code. Signatories are expected to deliver the revised Code by the end of 2021.<sup>41</sup>

## 5.1. Challenges and critical views

The COVID-19 disinformation monitoring programme has provided an in-depth overview of the actions taken by platforms to fight false and misleading information around coronavirus and vaccines.<sup>42</sup> However, the platforms' monthly reports are very diverse. They broadly follow platforms' commitments to the EU Code of Practice on Disinformation, but each platform has its own reporting style and fills in the metrics according to its preferences, and none provides much granular country-level data. The reports thus only become comparable with significant effort.<sup>43</sup>

As reported by Euroactiv, there is a concern that the expected revision of the Code of Practice on Disinformation by the end of 2021 will not be possible since the attention of online platforms is currently focused on negotiations on the Digital Services Act.<sup>44</sup> The Code is an instrument of soft law, hence it is voluntary and non-binding. However, some of its provisions, especially those anticipated in the Commission's 2021 Guidance, may become mandatory once the Digital Service Act is adopted. The Code's signatories might use the shorter time frame in which it was agreed to resist complying with the measures included in the Commission's guidance.<sup>45</sup>

---

41 European Commission, *Code of Practice on disinformation: Commission welcomes new prospective signatories and calls for strong and timely revision*, 1 October 2021, [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_21\\_4945](https://ec.europa.eu/commission/presscorner/detail/en/IP_21_4945).

42 European Commission, *Code of Practice on Disinformation*, 1 October 2021, <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>.

43 Trisha Meyer, Alexandre Alaphilippe, and Claire Pershan, *The good, the bad and the ugly: how platforms are prioritising some EU member states in their COVID-19 disinformation responses*, 28 April 2021, <https://www.disinfo.eu/publications/the-good-the-bad-and-the-ugly-how-platforms-are-prioritising-some-eu-member-states-in-their-covid-19-disinformation-responses/>.

44 Luca Bertuzzi and Molly Killeen, *"Commission pushes for 'timely' update of disinformation code of practice"*, Euroactiv, 1 October 2021, <https://www.euractiv.com/section/digital/news/commission-pushes-for-timely-update-of-disinformation-code-of-practice/>.

45 *Ibid.*

## 6. EU: European Digital Media Observatory, EDMO (2019)

The European Digital Media Observatory (EDMO) is an EU-funded project that supports the independent community working to combat disinformation. The creation of the European Digital Media Observatory (EDMO) is one of the elements in the European Commission's detailed action plan against disinformation,<sup>46</sup> published in December 2018. It has been financed by the European Union since 2019.

EDMO is managed by a consortium led by the European University Institute in Florence, Italy. The consortium includes the company Athens Technology Center from Greece, Aarhus University from Denmark, and the fact-checking organization Pagella Politica from Italy. EDMO has a governance structure completely independent from public authorities, including the European Commission.<sup>47</sup>

The European Digital Media Observatory (EDMO)<sup>48</sup> serves as a hub for fact-checkers, academics and other relevant stakeholders to collaborate with each other. It encourages them to actively link with media organizations and media literacy experts, and provide support to policymakers. This helps to coordinate actions in the fight against disinformation.

The activities of EDMO are based on five strands:

- a) mapping fact-checking organizations in Europe and supporting them by fostering joint and cross-border activities and dedicated training modules.
- b) mapping, supporting and coordinating research activities on disinformation at the European level, including the creation and regular update of a global repository of peer-reviewed scientific articles on disinformation.
- c) building a public portal providing media practitioners, teachers and citizens with information and materials aimed at increasing awareness, building resilience to online disinformation and supporting media literacy campaigns.
- d) design of a framework to ensure secure and privacy-protected access to platforms' data for academic researchers working to better understand disinformation.
- e) support to public authorities in the monitoring of the policies put in place by online platforms to limit the spread and the impact of disinformation; this includes support to the European Regulators Groups for Audiovisual Media Services (ERGA), reporting to the European Commission, etc.<sup>49</sup>

---

46 European Commission, *Action Plan against Disinformation*, 5 December 2018, <https://digital-strategy.ec.europa.eu/en/library/action-plan-against-disinformation>.

47 European Commission, *European Digital Media Observatory (EDMO)*, 22 September 2021, <https://digital-strategy.ec.europa.eu/en/policies/european-digital-media-observatory>.

48 EDMO, *United against disinformation*, 2020, <https://edmo.eu/>.

49 *Ibid.*

## 7. GERMANY: Network Enforcement Act, NetzDG (2018)

In 2017, Germany passed the Network Enforcement Act (NetzDG), also called the “Facebook Act”.<sup>50</sup> The law came into effect on 1 January 2018.

Under the NetzDG law, social media companies with more than 2 million registered users in Germany are required to establish an effective and transparent procedure to receive and review complaints of allegedly “unlawful” content. The “unlawful content” is defined in 22 provisions of the Criminal Code, ranging widely from insult of public office to actual threats of violence. The social media companies must block or remove “manifestly unlawful” content within 24 hours of receiving a complaint but have up to one week or potentially more if further investigation is required. In especially complex cases, companies can refer the case to an industry-funded but government-authorized body that is required to make determinations within a seven-day window.<sup>51</sup>

Human Rights Watch summarized the main aspects of NetzDG as follows:

- a) Under the law, the Federal Ministry of Justice and Consumer Protection can fine a responsible individual up to 5 million euros and the company up to 50 million euros for failing to establish a compliance system or for failing to issue a public report on their actions related to the law every six months. The amount of the fine depends on the gravity of the offence and the number of users on the platform, but the Ministry has not yet produced the fine structure.
- b) Social media companies must inform users of all decisions made in response to complaints and provide justification, but the law does not provide for meaningful judicial oversight or a process of judicial appeal when users want to contest a corporate or industry body decision to block or remove a post.
- c) To comply with the law, social media companies have created new mechanisms to report allegedly unlawful content and hired reviewers to analyze those reports. These reviewers have joined the teams these companies already had in place to monitor compliance with their user agreements. According to the social media companies Google (which owns YouTube) and Facebook, they each employ 10,000 of content reviewers globally, either directly or via contractors, to monitor violations of their community standards and violations of NetzDG.<sup>52</sup>

---

<sup>50</sup> Bundesministerium der Justiz und für Verbraucherschutz, FAQ: Act to Improve Enforcement of the Law in Social Networks, 2017, <https://www.bmjv.de/SharedDocs/FAQ/EN/NetzDG/NetzDG.html>.

<sup>51</sup> Human Rights Watch, Germany: Flawed Social Media Law: NetzDG is Wrong Response to Online Abuse, 14 February 2018, <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law>.

<sup>52</sup> *Ibid.*

The NetzDG has been amended several times. Following the 2019 murder of a politician, Walter Lübcke, by neo-Nazis – which was preceded by targeted threats and hate speech online – the reform of NetzDG, in 2020, extended the scope of the law by placing an additional reporting obligation on social media platforms which requires them to report certain types of “unlawful” content to the Federal Criminal Police Office.<sup>53</sup>

A wide-ranging reform of the NetzDG law was introduced in June 2021,<sup>54</sup> aiming to improve the user-friendliness of the reporting channels for complaints about unlawful content and to increase the content and comparability of social media providers’ transparency reports. Furthermore, the amendment introduces an appeals procedure for measures taken by the social network provider. The powers of the Federal Office of Justice are expanded to include supervisory powers. Due to new requirements under the EU Audiovisual Media Services Directive, video-sharing platform services are included in the scope of the Network Enforcement Act.<sup>55</sup>

As elaborated in more detail by the Global Legal Monitor, these four main reforms include:

a) User-Friendliness of Complaint Procedures: the amended Network Enforcement Act provides that the procedure for submitting complaints about unlawful content to social network providers must be “easily recognizable, directly accessible, and permanently available.” However, this requirement has not been uniformly implemented by providers, and some forms to submit complaints are hidden or available only after several clicks. To improve the user-friendliness of the complaint procedure, the amendment also adds language requiring that the complaint procedure be “easy to use.”

b) Appeals Procedure and Arbitration: The amendment obligates social media providers to establish a mechanism whereby the complainant and the affected user may initiate a review of the decision to remove or not remove flagged content (appeals procedure). The appeal of the original decision must be submitted within two weeks. Furthermore, the amendment authorizes the Federal Office of Justice to approve arbitration bodies organized under private law for out-of-court settlements of disputes between complainants/users and social media providers. Participation in arbitration is voluntary.

c) Transparency Reports: Social media networks that receive more than 100 complaints about illegal content in a calendar year are required to publish biannual reports in German on how they deal with these complaints. The amendment requires that more information be included in the transparency

---

53 Natasha Lomas, “Germany tightens online hate speech rules to make platforms send reports straight to the feds,” *Techcrunch*, 19 June 2020, <https://techcrunch.com/2020/06/19/germany-tightens-online-hate-speech-rules-to-make-platforms-send-reports-straight-to-the-feds/>.

54 The Act to Amend the Network Enforcement Act (2021) is available at: <https://perma.cc/9W8E-GSWM>.

55 Jenny Gesley, “Germany: Network Enforcement Act Amended to Better Fight Online Hate Speech,” *Library of Congress*, 2021, <https://www.loc.gov/item/global-legal-monitor/2021-07-06/germany-network-enforcement-act-amended-to-better-fight-online-hate-speech/>.

reports. Among other things, providers must report whether procedures for the automated detection of unlawful content are used and, if yes, how they work. In addition, reports must further subdivide the numbers of complaints according to the amount of time it took to remove the flagged content (within 24 hours, within 48 hours, within a week, or at a later date).

d) Expansion of Powers of the Federal Office of Justice: Under the old version of the Network Enforcement Act, the Federal Office of Justice was able to issue fines only for noncompliance with the provisions of the Network Enforcement Act. The amendment gives the Federal Office of Justice powers to supervise compliance with the act. Once it has determined that an infringement has occurred, the office can require the social media provider to remedy the infringement. It can also request information about implementation measures, the number of registered users in Germany, and the number of complaints received in the past calendar year. Witnesses are obligated to testify in the administrative procedure.<sup>56</sup>

As explained by Judit Bayer, simply deleting or blocking the content did not help to stop the tide of increasing extremist expression in Germany.

Therefore, the amended act aimed at channelling the information about online hate speech and threatening with hate crime into the criminal authorities, extending the scope of cases when access to user data can be granted to the Federal Criminal Police Authority and national security services. Beyond that, the law also added new actions to the Criminal Code to include behaviours which became prevalent in the online context. For example, the crime of ‘threatening’ meant threatening with a crime that entailed at least one-year imprisonment – this did not include threats against sexual self-determination, bodily integrity, or the threat of causing severe material damage – typical online threats.<sup>57</sup>

## 7.1. Challenges and critical views

The Network Enforcement Act has been largely criticized since it was adopted in 2017.

Human rights organizations, such as Human Rights Watch, said the German law is fundamentally flawed. They recognized the valid concerns of the governments and the public about the proliferation of illegal or abusive content online. However, as Germany director at Human Rights Watch explained their views in early 2018, the law is “vague, overbroad, and turns private companies into overzealous censors to avoid steep fines, leaving users with no judicial oversight or right to appeal.”<sup>58</sup>

---

<sup>56</sup> *Ibid.*

<sup>57</sup> Judit Bayer, *Germany: New law against right-wing extremism and hate crime*, 24 April 2021, <https://inform.org/2021/04/24/germany-new-law-against-right-wing-extremism-and-hate-crime-judit-bayer/>.

<sup>58</sup> Human Rights Watch, *Germany: Flawed Social Media Law: NetzDG is Wrong Response to Online Abuse*, 14 February 2018, <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law>.

The criticism increased with the implementation of the law, particularly when some high-profile users of social media were blocked or their accounts were temporarily suspended. Also, four large political parties opposed the law: The Left, which voted against the law; the Free Democrats and the Alternative for Germany, which were not in Parliament when the law was passed; and the Green Party, which abstained from the parliamentary vote.<sup>59</sup>

Another concern was raised because, according to the Human Rights Watch, the NetzDG has had a domino effect since governments around the world used the example of the German law to restrict online speech “by forcing social media companies to act as their censors”.<sup>60</sup>

The legislative process for adoption of the amended NetzDG (2021) was also challenging. In addition to the contesting by opposition parties for different reasons, the President of the Parliament refused to sign the bill into the law, expressing constitutional concerns and calling for specific improvements regarding the reporting obligation of social media platforms which gave access to users’ data, including their password and other information effectively giving access to the user profile. Taking into account the Constitutional Court’s decision on similar issue, in 2020, and in order to overcome the constitutional problem, a new “repair act” was passed to narrow the scope of cases in which authorities may demand access to users’ data to specified criminal actions.<sup>61</sup>

In the 2021 amended law the reporting obligation foresees a judicial order to receive passwords and other data which allow access to users’ profiles or even devices. As noticed by Judit Bayer, the authorities expect a large wave of reporting, therefore several hundred new positions have been created in the Ministry. To be able to deal with the increased workload, a new department was created within the Federal Criminal Police Authority.<sup>62</sup>

---

59 *Ibid.*

60 *Ibid.*

61 Judit Bayer, Germany: New law against right-wing extremism and hate crime, 24 April 2021, <https://inform.org/2021/04/24/germany-new-law-against-right-wing-extremism-and-hate-crime-judit-bayer/>.

62 *Ibid.*



## 8. AUSTRIA: Communication Platforms Act (2021)

As part of a regulatory package against hate speech on the internet, Austria introduced new obligations for online communication platforms. The new Communication Platforms Act (“Kommunikationsplattformen-Gesetz”, KoPIG)<sup>63</sup> applies to domestic and foreign providers of for-profit communication platforms that have more than 100,000 users in Austria or revenues in Austria of more than EUR 500,000. The new law came into force on 1 January 2021.

Gabriela Staber and Angelika Stütz, legal experts of the law firm CMS-Law Now in Vienna, have elaborated the main elements of the new law as follows:<sup>64</sup>

a) Excluded are online marketplaces, non-profit online encyclopaedias (such as Wikipedia), and learning platforms, newspaper and television company platforms hosting their journalistic offerings, and apps used for individual communication.

b) The video content on video-sharing platforms is governed by the rules of the Audio-visual Media Services Act, while the Communication Platforms Act applies to the rest of their content.

c) The regulatory authority is required to keep a (declarative) list of platforms governed by the new rules. This list must be compiled on the authority’s own initiative since there is no obligation for a platform to notify the authority when it commences or terminates its services.

d) The new rules require platforms to:

- Appoint and notify the authority of a) an authorized representative who is responsible for compliance with the new laws, and b) a representative who is able to accept official communications on behalf of the platform (it may be the same person, but they must speak German).

- Set up an “effective and transparent procedure” for reporting and deleting illegal content: a) Deletion must take place within 24 hours if the illegality is “obvious to a legal layman (...)”, or within seven days if a detailed examination is necessary; b) There must be a complaints procedure for users affected by deletion or blocking to avoid “overblocking”; c) Illegal content comprises, for example, defamation, harassment, pornography involving minors, racist, discriminatory, or national-socialist content, unauthorized photographs, stalking by means of telecommunication.

---

63 Final text of the law is available on the platform of the European Commission at: [https://ec.europa.eu/growth/tools-databases/tris/index.cfm/en/search/?trisaction=search\\_detail&year=2020&num=544&mLang=EN](https://ec.europa.eu/growth/tools-databases/tris/index.cfm/en/search/?trisaction=search_detail&year=2020&num=544&mLang=EN).

64 Gabriela Staber and Angelika Stutz, *Communication platforms face new obligations and high fines in Austria*, 22 March 2021, [https://www.cms-lawnow.com/ealerts/2021/03/au-test?cc\\_lang=en](https://www.cms-lawnow.com/ealerts/2021/03/au-test?cc_lang=en).

- Report the number and results of notifications annually (or semi-annually if the number of users exceeds one million).

- Store deleted postings for at least ten weeks for any possible prosecution. Pay a financing contribution to the budget of the regulatory authority.

e) For failure to comply with these obligations, fines of up to EUR 10 million can be imposed on the platform (depending on several factors, such as revenue, number of users, prior misconduct, severity and length of violation). Moreover, it is possible to impose a fine of up to EUR 1 million on the members of the managing board if upon request from the authority they fail to appoint an authorized representative or a representative to accept official documents served by the authority.<sup>65</sup>

## 8.1. Challenges and critical views

As pointed out by Staber and Stütz, the conformity of Austria's new rules with EU law is questionable.<sup>66</sup>

As they highlight, the EU Commission outlined in a letter to the Austrian Government that the Communication Platforms Act is likely to impede the freedom to provide services. In particular, the EU Commission is concerned about the additional costs and administrative burdens associated with implementing a notification and take-down mechanism and the need to have the necessary language skills and local cultural and legal knowledge to be able to assess content and interact with the authorities. By imposing stricter requirements on operators than those of the country in which they are based, the Communication Platforms Act does not comply with the e-Commerce Directive and does not meet its requirements for exceptions to the country-of-origin principle.<sup>67</sup>

In addition, the EU Commission questions why Austria is passing its own set of rules now while the Digital Services Act is in the making, which is aimed at harmonizing at the EU level many items governed by the Austrian Communication Platforms Act. Since the Digital Services Act will take the form of a regulation and hence be directly applicable in each member state, the EU Commission points out that Austria will have to repeal part of its new laws when it comes into force. In particular, some of the provisions of the Austrian law seem to incentivize the fast take-down of seemingly infringing content – something the Digital Services Act aims to avoid in order to safeguard free speech.<sup>68</sup>

Despite the Commission's concerns, the Austrian government has not postponed its new regulations, which were triggered by several highly

---

<sup>65</sup> *Ibid.*

<sup>66</sup> *Ibid.*

<sup>67</sup> *Ibid.*

<sup>68</sup> *Ibid.*



publicized incidents of hate speech against well-known Austrian politicians.<sup>69</sup> In the analysis of the Austrian law in the proposal stage, European Digital Rights (EDRi), an association of civil and human rights organizations from across Europe, acknowledged “the big improvement compared to the German NetzDG” regarding complaint and redress mechanisms to supervise the platform’s decisions and increase the quality of their content moderation processes.<sup>70</sup> However, EDRi raised concern regarding the penalty regulations. Particularly, the power of the regulatory authority KommAustria “to shape online discourse” is seen as “enormous”, taking into account the power to issue administrative and penalty decisions regarding the content moderation procedures and decisions of the platforms. At the same time, ECRi considers the transparency obligations for online platforms in the Austrian law very positive.<sup>71</sup>

## 9. SLOVENIA: Spletno oko hotline for reporting hate speech online (2006)

The Spletno oko (Web Eye) hotline allows internet users in Slovenia to anonymously report hate speech and child sexual abuse images if they come across them online. The hotline was established in September 2006. It is a part of the Safer Internet Center, which is coordinated by the University of Ljubljana, Faculty of Social Sciences, in cooperation with partners: Arnes, Slovenian Association of Friends of Youth and Centre MISSS (Youth Information and Counselling Center of Slovenia). Among members of the Advisory Board of the Safer Internet Center are representatives of the Office of the State Prosecutor General and of the General Directorate of Police, representatives of the media and representatives of organizations working in the field of child protection.

The main task of the Spletno oko hotline is to reduce the amount of child sexual abuse images and hate speech online, in cooperation with the police, internet service providers, and other governmental and non-governmental organizations.

The main activities of the Spletno oko hotline include the reporting mechanism which enables anonymous reports of illegal content on the internet as well as fast and effective analysis and processing of the received reports.

The reporting point is located on the Spletno oko website and is available to all internet users. The form guides the user through the offered categories of potentially illegal content until the submission of a duly completed report.

---

69 *Ibid.*

70 EDRi, *First Analysis of the Austrian Anti-Hate Speech Law (NetDG/KoPIG)*, 10 September 2020, <https://edri.org/our-work/first-analysis-of-the-austrian-anti-hate-speech-law-netdg-koplg/>.

71 *Ibid.*

At the same time, the website gives users an explanation of whether the content is potentially illegal and advice on options for further actions. A specially trained reviewer of Spletno oko reviews the submitted reports and forwards those identified as potentially illegal for consideration to the police, in accordance with pre-agreed procedures.

In 2020, Spletno oko processed 2,268 reports of hate speech online (compared to 773 in 2019). In 2020, they forwarded to the police 67 reports (compared to 90 in 2019). Among 67 reports of hate speech online evaluated by the Spletno oko reviewer as “potentially unlawful” and forwarded to the police, in 2020, 60 per cent contained incitement to murder or killing and 33 per cent contained denial or glorification of the Holocaust or war crimes.

Among important achievements of Spletno oko is the Code of Hate Speech Regulation on Web Portals in Slovenia.<sup>72</sup> It was developed and agreed jointly by Spletno oko and the main web portals in Slovenia, in 2010. Six online media outlets joined the initiative in 2010, and later three other online media outlets also signed the Code. By joining the Code, the online media adopted uniformed guidelines for the regulation of hate speech in comment sections, which provided for consistent moderation of hateful, intolerant and offensive comments, mandatory registration of users and unification of the form for submitting comments in the comment sections of the online media outlets. As a part of implementation of the Code, the comment sections of web portals should contain a warning about the legal provisions on hate speech in Criminal Code, and on own community standards, and a special button for reporting hate speech in the users’ comments. The report submitted in that way informs both Spletno oko and the administrators of the online media outlets. The latter should remove comments that violate the rules of the web portal as soon as possible, and Spletno oko reviews them separately and forwards them for consideration to the police in case of potentially unlawful content.

Since 2016, the Spletno oko hotline has also carried out monitoring of the self-regulatory mechanisms of global social media companies and their compliance with the EU Code of Conduct on countering illegal hate speech online. As a part of the monitoring, the Spletno oko hotline has recorded the extent to which social media companies remove reported cases in Slovenian language, at what time and how they inform complainants about it.

---

72 See: [https://www.spletno-ok.si/sites/default/files/kodeks\\_oblikovan\\_0.pdf](https://www.spletno-ok.si/sites/default/files/kodeks_oblikovan_0.pdf).

## 9.1. Challenges and critical views

Some signatories of the Code of Hate Speech Regulation among online media outlets in Slovenia no longer have a reporting button installed on the portals, especially those who have installed social network plug-ins for the purpose of commenting by users. There is lack of motivation of some online media portals for more active cooperation and for appropriate regulation of user comments on their websites.

A huge disparity between the large number of reports on hate speech online submitted to the Spletno oko hotline in 2020 (as well as in many previous years) and much smaller number of reports forwarded to the police by Spletno oko as potentially unlawful according to the Criminal Code, and the even smaller number of cases actually submitted by the police to the prosecutors or convicted in the court procedures, illustrate the challenges of regulation and self-regulation of hate speech online. Although the large amount of the hate speech content perceived as such by the internet users did not qualify for criminal prosecution, it clearly indicates the need to address this subject with other legal and non-legal instruments.

## Literature and sources

Abbamonte, Giuseppe B. “Book Review: Giovanni Pitruzzella & Oreste Pollicino, Disinformation and hate speech: a European constitutional perspective.” *EU Law Live*, 15 September 2020. <https://eulawlive.com/library/disinformation-and-hate-speech-a-european-constitutional-perspective/> (accessed November, 2021).

“The Digital Services Act package is here.” *Allen & Overy*. 16 December 2020. <https://www.allenoverly.com/en-gb/global/news-and-insights/publications/the-digital-services-act-package-is-here> (accessed November, 2021).

Barata, Joan. “The Digital Services Act and social media power to regulate speech: obligations, liabilities and safeguards.” In *Unravelling the Digital Services Act package, IRIS Special 2021*, European Audiovisual Observatory, 5-19. 2021.

Barata, Joan. *Regulating content moderation in Europe beyond the AVMSD*. 25 February 2020. <https://blogs.lse.ac.uk/medialse/2020/02/25/regulating-content-moderation-in-europe-beyond-the-avmsd/> (accessed November, 2021).

Bayer, Judit. *Germany: New law against right-wing extremism and hate crime*. 24 April 2021. <https://inform.org/2021/04/24/germany-new-law-against-right-wing-extremism-and-hate-crime-judit-bayer/> (accessed November, 2021).

Bertuzzi, Luca and Killeen, Molly. “Commission pushes for ‘timely’ update of disinformation code of practice.” *Euractiv*, 1 October 2021. <https://www.euractiv.com/section/digital/news/commission-pushes-for-timely-update-of-disinformation-code-of-practice/> (accessed November, 2021).

Bukovská, Barbora. *Article 19: The European Commission’s Code of Conduct for Countering Illegal Hate Speech Online: An analysis of freedom of expression implications*. Transatlantic Working Group, May 7, 2019. <https://www.ivir.nl/publicaties/download/Bukovska.pdf> (accessed November, 2021).

Bundesministerium der Justiz und für Verbraucherschutz. *FAQ: Act to Improve Enforcement of the Law in Social Networks*. 2017. <https://www.bmjv.de/SharedDocs/FAQ/EN/NetzDG/NetzDG.html> (accessed November, 2021).

Disinfo.eu. *Fact-checkers and experts call on MEPs to reject a media exemption in the DSA*. 5 November 2021. <https://www.disinfo.eu/advocacy/fact-checkers-and-experts-call-on-meps-to-reject-a-media-exemption-in-the-dsa/> (accessed November, 2021).

EDMO. *United against disinformation*. 2020. <https://edmo.eu/> (accessed November, 2021).

EDRI. *First Analysis of the Austrian Anti-Hate Speech Law (NetDG/KoPiG)*. 10 September 2020. <https://edri.org/our-work/first-analysis-of-the-austrian-anti-hate-speech-law-netdg-koplg/> (accessed November, 2021).

European Audiovisual Observatory. *Unravelling the Digital Services Act package, IRIS Special*. 2021. <https://rm.coe.int/iris-special-2021-01en-dsa-package/1680a43e45> (accessed November, 2021).

European Commission. *Action Plan against Disinformation*. 5 December 2018. <https://digital-strategy.ec.europa.eu/en/library/action-plan-against-disinformation> (accessed November, 2021).

European Commission. *Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement*. 10 September 2020. <https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement> (accessed November, 2021).

European Commission. *Code of Practice on Disinformation*. 1 October 2021. <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation> (accessed November, 2021).

European Commission. *Code of Practice on disinformation: Commission welcomes new prospective signatories and calls for strong and timely revision*. 1 October 2021. [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_21\\_4945](https://ec.europa.eu/commission/presscorner/detail/en/IP_21_4945) (accessed November, 2021).

European Commission. *Commission publishes EU Code of Conduct on countering illegal hate speech online continues to deliver results*. 22 June 2020. [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_20\\_1134](https://ec.europa.eu/commission/presscorner/detail/en/IP_20_1134) (accessed November, 2021).

European Commission. *Countering illegal hate speech online*. <https://ec.europa.eu/newsroom/just/items/54300> (accessed November, 2021).

European Commission. *Disinformation: EU assesses the Code of Practice and publishes platform reports on coronavirus related disinformation*. 10 September 2020. [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_20\\_1568](https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1568) (accessed November, 2021).

European Commission. *European Digital Media Observatory (EDMO)*. 23 September 2021. <https://digital-strategy.ec.europa.eu/en/policies/european-digital-media-observatory> (accessed November, 2021).

European Commission. *First baseline reports – Fighting COVID-19 disinformation Monitoring Programme*. 10 September 2020. <https://digital-strategy.ec.europa.eu/en/library/first-baseline-reports-fighting-covid-19-disinformation-monitoring-programme> (accessed November, 2021).

European Commission. *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*. 15 December 2020. <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN> (accessed November, 2021).

European Commission. *The Code of conduct on countering illegal hate speech online*. 22 June 2020. [https://ec.europa.eu/commission/presscorner/detail/en/qanda\\_20\\_1135](https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1135) (accessed November, 2021).

European Parliament and the Council of the European Union. *Audiovisual Media Services Directive*. 10 March 2010. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32010L0013&from=EN> (accessed November, 2021).

European Parliament and the Council of the European Union. *Audiovisual Media Services Directive*. 14 November 2018. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32018L1808&from=SL> (accessed November, 2021).

European Parliament and the Council of the European Union. *Directive on Electronic Commerce*. 8 June 2000. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0031&from=EN> (accessed November, 2021).

Gesley, Jenny. "Germany: Network Enforcement Act Amended to Better Fight Online Hate Speech." Library of Congress, 2021. <https://www.loc.gov/item/global-legal-monitor/2021-07-06/germany-network-enforcement-act-amended-to-better-fight-online-hate-speech/> (accessed November, 2021).

Human Rights Alliance. *Joint Statement of the Digital Services Act Human Rights Alliance*. 21 October 2021. [https://www.accessnow.org/cms/assets/uploads/2021/10/Digital\\_Services\\_Act\\_Human\\_Rights\\_Alliance\\_Statement.pdf](https://www.accessnow.org/cms/assets/uploads/2021/10/Digital_Services_Act_Human_Rights_Alliance_Statement.pdf) (accessed November, 2021).

Human Rights Watch. *Germany: Flawed Social Media Law: NetzDG is Wrong Response to Online Abuse*. 14 February 2018. <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law> (accessed November, 2021).

Killeen, Molly. "Progress stalls on Commission's online hate speech efforts." Euractiv, 7 October 2021. <https://www.euractiv.com/section/digital/news/progress-stalls-on-commissions-online-hate-speech-efforts/> (accessed November, 2021).

Lomas, Natasha. "Germany tightens online hate speech rules to make platforms send reports straight to the feds." Techcrunch, 19 June 2020. <https://techcrunch.com/2020/06/19/germany-tightens-online-hate-speech-rules-to-make-platforms-send-reports-straight-to-the-feds/> (accessed November, 2021).



Meyer, Trisha, Alaphilippe, Alexandre, and Pershan, Claire. *The good, the bad and the ugly: how platforms are prioritising some EU member states in their COVID-19 disinformation responses*. 28 April 2021. <https://www.disinfo.eu/publications/the-good-the-bad-and-the-ugly-how-platforms-are-prioritising-some-eu-member-states-in-their-covid-19-disinformation-responses/> (accessed November, 2021).

Penfrat, Jan. "The EU's attempt to regulate Big Tech: What it brings and what is missing." EDRi, 18 December 2020. <https://edri.org/our-work/eu-attempt-to-regulate-big-tech/> (accessed November, 2021).

Reynders, Didier. *Countering illegal hate speech online: 6th evaluation of the Code of Conduct*. European Commission, 7 October 2021. [https://ec.europa.eu/info/sites/default/files/factsheet-6th-monitoring-round-of-the-code-of-conduct-october2021\\_en\\_1.pdf](https://ec.europa.eu/info/sites/default/files/factsheet-6th-monitoring-round-of-the-code-of-conduct-october2021_en_1.pdf) (accessed November, 2021).

Simon, Eva. *Recommendations for the reform of the Audiovisual Media Services Directive*. Civil Liberties Union for Europe, September, 2017. [https://drive.google.com/file/d/0B07ejuhu6mZ5Zk5aUGJyTmt50WM/view?resourcekey=0-1S2CLG5-68jk\\_2xCDkUkcQ](https://drive.google.com/file/d/0B07ejuhu6mZ5Zk5aUGJyTmt50WM/view?resourcekey=0-1S2CLG5-68jk_2xCDkUkcQ) (accessed November, 2021).

Staber, Gabriela and Stutz, Angelika. *Communication platforms face new obligations and high fines in Austria*. March 22, 2021. [https://www.cms-lawnow.com/ealerts/2021/03/au-test?cc\\_lang=en](https://www.cms-lawnow.com/ealerts/2021/03/au-test?cc_lang=en) (accessed November, 2021).

The Council of the European Union. *Council Framework Decision 2008/913/JHA on combating certain forms and expressions of racism and xenophobia by means of criminal law*. 28 November 2008. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32008F0913&from=en> (accessed November, 2021).

Vidgen, Bertie, Burden, Emily, and Margetts, Helen. *Facing the challenge of online hate: the AVMSD and the role of regulation*. 30 March 2021. <https://blogs.lse.ac.uk/medialse/2021/03/30/facing-the-challenge-of-online-hate-the-avmsd-and-the-role-of-regulation/> (accessed November, 2021).

Vidgen, Bertie, Burden, Emily, and Margetts, Helen. *Understanding online hate: VSP Regulation and the broader context*. The Alan Turing Institute, February, 2021. [https://www.ofcom.org.uk/\\_data/assets/pdf\\_file/0022/216490/alan-turing-institute-report-understanding-online-hate.pdf#page=84](https://www.ofcom.org.uk/_data/assets/pdf_file/0022/216490/alan-turing-institute-report-understanding-online-hate.pdf#page=84) (accessed November, 2021).

Wallis, Diana. "Digital Services Act: a media exemption would open a loophole for disinformation." Euroactiv, 20 October 2021. <https://www.euractiv.com/section/digital/opinion/digital-services-act-a-media-exemption-would-open-a-loophole-for-disinformation/> (accessed November, 2021).

Wikipedia. *Digital Services Act*. [https://en.wikipedia.org/wiki/Digital\\_Services\\_Act](https://en.wikipedia.org/wiki/Digital_Services_Act) (accessed November, 2021).

This publication is a part of the RESILIENCE project research and advocacy component. It includes a series of factsheets on NATIONAL REGULATORY AND SELF-REGULATORY FRAMEWORKS AGAINST HATE SPEECH AND DISINFORMATION in Albania, Bosnia and Hercegovina, Kosovo, Montenegro, North Macedonia, Serbia and Turkey. The series also includes a factsheet with examples of regulatory and self-regulatory mechanisms on the EU level and in the EU member states.

Nine media development organizations in the Western Balkans and Turkey have joined forces under an EU-funded project 'RESILIENCE: Civil society action to reaffirm media freedom and counter disinformation and hateful propaganda in the Western Balkans and Turkey'. The three-year project is coordinated by the South East European Network for Professionalization of Media (SEENPM), a network of media development organizations in Central and South East Europe, and implemented in partnership with: the Albanian Media Institute in Tirana, the Foundation Mediacentar Sarajevo, Kosovo 2.0 in Pristina, the Montenegro Media Institute in Podgorica, the Macedonian Institute for Media in Skopje, the Novi Sad School of Journalism in Novi Sad, the Peace Institute in Ljubljana, and bianet in Istanbul.