Tackling Disinformation and Online Hate Speech:

EU and Member State approaches, so far

DEMOCRACY
REPORTING
INTERNATIONAL

Tackling Disinformation and Online Hate Speech:

EU and Member State approaches, so far

Funded by:

Federal Republic of Germany
Foreign Office

This publication is a result of a project carried out by Democracy Reporting International with the financial support from the Federal Republic of Germany Foreign Office. This publication was developed by the DRI Social Media & Democracy Team; Democracy Team; Michael Meyer-Resende, Rafael Goldzweig and Helena Schwertheim, and with valuable contributions from Nazrin Huseinzade.

December 2020

Contents

1. Executive Summary

The EU has addressed online disinformation and hate speech through a variety of policies and actions over the past three years. As far as regulation is concerned, it has adopted a soft regulatory approach, focusing on codes of conduct and practice, whereby online providers that shape public discourse have made voluntary commitments.

This approach will now change. In two major initiatives unveiled in December 2020, the European Commission has announced that it will propose binding EU legislation on issues including paid political advertising, algorithmic transparency and electoral integrity.

The question is: How will this regulatory drive relate to legislation and policies in the EU's 27 member states? What have they been doing on this front in recent years? In this paper and it associated database we provide an overview of the **EU 27 approaches**.

Main takeaways:

- Those EU governments most concerned about Russian foreign policy and disinformation campaigns have made the most systematic efforts to counter disinformation, both within government and society in general. In contrast, the German and Austrian governments have been most concerned with (mostly domestic) hate speech, and the protection of those targeted. The French government has legislated in both fields, but its law against hate speech was largely annulled by the Constitutional Council.

On online disinformation

- Those EU member states most concerned about Russian disinformation campaigns (The Baltic and Nordic countries, and the Czech Republic) have developed the most comprehensive approach to combating disinformation, often based on cross-governmental task forces and "all-society" approaches. In an assessment of the best digital literacy campaigns, they fare best.

- In 2020, Germany adopted legislation obliging companies to become more transparent about the algorithms that determine how content is ranked high or recommended, in order to develop a better understanding of whether algorithms boost disinformation. The European Commission has announced that it will require similar transparency.

- Some member states have legislation that predates the internet on false speech that can harm public order and peace (Croatia, Cyprus, France, Malta, Hungary). Following the onset of the Covid-19 pandemic, the Hungarian parliament added a controversial provision to the relevant article in the Criminal Code, expanding its scope.

- France is the only state in the EU that adopted a law specifically targeting disinformation during elections periods. The main innovation in the law is the possibility for a judge to rule within 48 hours upon referral, that the spread of disinformation on a platform should be halted. The law has been criticized for including only a vague definition of disinformation.

On online hate speech

- The most notable and controversial legal innovation in this field has been Germany's NetzDG (Network Enforcement Act), which obliges platforms to remove "manifestly illegal" content within 24 hours of receiving a complaint, or seven days, where the illegality of the content is not clear. The law has been criticized, in particular, for delegating authority for removing speech to private companies. The government has defended the approach, arguing that companies must take responsibility for the use that is made of their platforms and invest more in protecting potential victims from illegal speech online.

- In September 2020, Austrian government proposed a similar law, but more specifically articulating complaints and appeals mechanisms, both for those filing complaints and for those against whose content complaints are being filed.

- France adopted a similar law ("Loi Avia") in 2019, but this was ultimately annulled by France's Constitutional Council, which ruled that it represented a disproportionate limitation of freedom of speech. Despite the overturning of the law, in July 2020, French authorities launched an "Online Hate Observatory" to monitor and analyse online content.

- In the Czech Republic and Spain, some criminal provisions carry a higher potential penalty if they are committed online. The provisions have been criticized for assuming that all online speech reaches a wider audience, without requiring a determination of the actual reach of criminal content.

- The Hungarian government lost landmark cases in the European Court of Human Rights, when its Supreme Court tried to hold platforms liable for comments made in their comment sections by users.

2. Introduction

The EU is overhauling its regulation on internet platforms through the Digital Services Act (DSA) and is hoping to address specific threats to democratic discourse online through the European Democracy Action Plan (EDAP).

In recent years, many member states have developed their own approaches to tackling disinformation and online hate speech. This policy brief maps the national and EU-level regulatory frameworks[1] for addressing online hate speech and disinformation. In doing so, it provides insights into regulatory successes and shortcomings, where current strengths lie, and where the EU could further strengthen and complement national efforts.

3. The European Union's Liability Regime

3.1 Waiting on a new liability regime: The DSA is to replace the ECD

Until the Digital Services Act comes into force, the primary legal act governing online platforms[2] in the EU will be the **E-Commerce Directive (ECD)**.[3] The Directive exempts intermediaries (online platforms) from liability for content they manage, if they fulfil certain conditions. The ECD's assumption that intermediaries merely "pass on" content is out of date. While these platforms do not create content, they shape public discourse actively by ranking, recommending or selecting content for each user. With the proliferation of illegal content online, such as hate speech, and concerns about disinformation, some EU member states have adopted legislation stipulating notice and takedown procedures,[4] and this has led to legal frag mentation throughout the EU and, often, the practical non-enforceability of these measures in many states.

Following the logic of the single market, the DSA is intended to overcome this fragmentation by providing clarity and a horizontal framework for regulatory oversight, accountability and transparency of the online space. The Directive and its sister package, the European Democracy Action Plan (EDAP) are aimed at updating and harmonising EU approaches to countering disinformation and online hate speech.

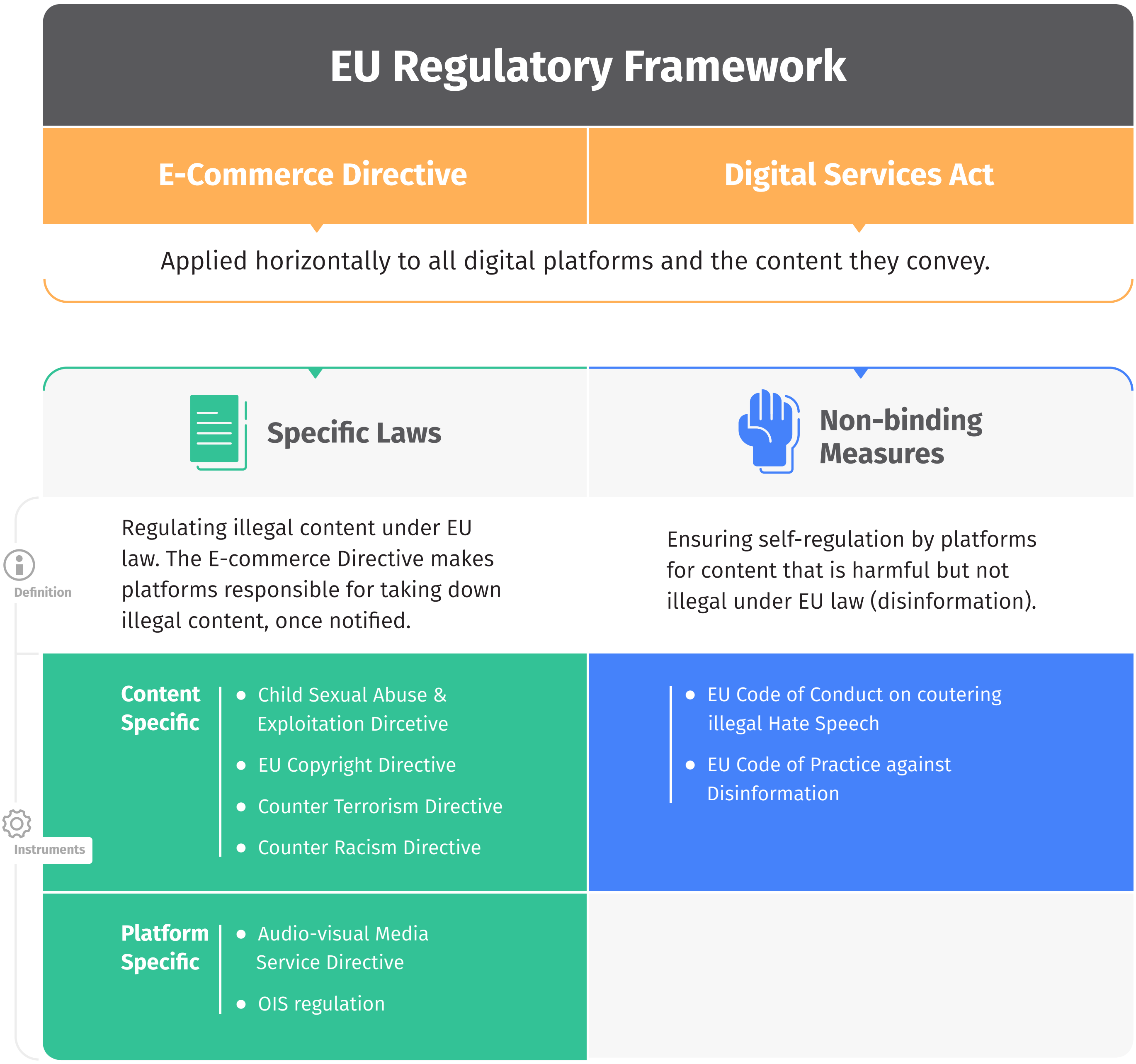
[1] As of December 2020

[2] For the purpose of this policy brief, "online platform" will be used to denote a digital service provider that facilitates interactions on the Internet between two or more natural/legal persons by offering and performing a variety of functions and services. The most typical examples include search engines (e.g. Google, Bing), online marketplaces (e.g. Amazon, eBay), audio-visual and music platforms (e.g.Deezer, Spotify, Netflix), video sharing platforms (e.g. YouTube, Vimeo), payment systems (e.g. PayPal, Apple Pay), social networks (e.g. Facebook, Linkedin), app stores (e.g. Apple App Store)

[3] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, see https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031 (15 December 2020)

[4] Notice and takedown refers to a procedure implied in article 14 (2) and explicitly mentioned in article 21 (2) of the ECD whereby the hosting provider is given notice about the illegal/potentially illegal content, and has to take it down within a given time-frame.

The DSA also proposes new rules to ensure greater accountability on how platforms moderate content, on advertising and on algorithmic processes. While not providing binding measures, it strengthens policies on disinformation and signals new efforts to address online political advertising, as well as to safeguard elections in the online sphere.



EU Regulatory Framework

E-Commerce Directive | **Digital Services Act**

Applied horizontally to all digital platforms and the content they convey.

Specific Laws

Definition: Regulating illegal content under EU law. The E-commerce Directive makes platforms responsible for taking down illegal content, once notified.

Instruments:

Content Specific

- Child Sexual Abuse & Exploitation Dirctive
- EU Copyright Directive
- Counter Terrorism Directive
- Counter Racism Directive

Platform Specific

- Audio-visual Media Service Directive
- OIS regulation

Non-binding Measures

Definition: Ensuring self-regulation by platforms for content that is harmful but not illegal under EU law (disinformation).

Instruments:

- EU Code of Conduct on coutering illegal Hate Speech
- EU Code of Practice against Disinformation

4. Online Platforms and Disinformation

4.1 Disinformation measures at the EU level

Disinformation is an area where rigid legislation by the EU or member states is not desirable, mainly because doing so would constitute a disproportionate interference with freedom of speech. In the absence of such legislation, the soft law instrument in play is the EU Code of Practice on Disinformation – a voluntary arrangement between the European Commission and big tech companies such as Facebook, Google, Twitter and Mozilla, who signed on in 2018, and, more recently, Microsoft (2019) and TikTok (2020).[5] EDAP signals that, in 2021, the Commission will introduce legislation providing for greater transparency in online political advertising.[6]

4.2 Different approaches to tackling disinformation at the national level

Public debate around disinformation made the headlines for its role in the 2016 United States Elections, when manipulation of the online environment by Russian actors was uncovered. This echoed differently among the EU member states, who adopted varying measures to counter the threat ahead of the 2019 elections for the European Parliament.

It is worth noting that most of the countries who took the lead in responding to the threat were those most concerned about Russian disinformation, which they see as a matter of national security. They were also the most active supporters of initiatives at the EU level (such as EU StratCom), along with taking responsibility for effectively protecting their citizens in their own hands in the long term, adopting a more horizontal approach when it came to the policy instruments available to them.

4.2.1 The “Task-Force” approach

A group of countries approached the issue with a more comprehensive action plan. Former Socialist countries, the Nordic countries and the Czech Republic took an approach that involved various government ministries and departments, as well as society in general.

For **Denmark**, **Estonia**, **Finland**, **Lithuania** and **Sweden**, disinformation has a cybersecurity connotation, with activities to counter this led by ministries of defense, or with laws and centers dedicated to countering disinformation attacks/campaigns. Examples include the Czech Centre against Terrorism and Hybrid Threats (under the Ministry of the Interior), Estonia’s Defense League (under the Ministry of Defense) and Latvia’s Baltic Media Centre of Excellence. These will likely be joined by Sweden’s[7] Psychological Defence Authority and Defence Research Agency, as well as the new European Centre of Excellence for Countering Hybrid Threats, to be hosted in Finland. Denmark has adopted the Cyber and Information Security Strategy 2018-2021.

[5] European Commission, Code of Practice on Disinformation, 26 September 2018.
https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation (15 December 0202)

[6] Specifically, the EDAP states the EC will “Propose, in 2021, legislation to ensure greater transparency in the area of sponsored content in a political context (‘political advertising’)”.
See: European Commission, European Democracy Action Plan, 3 December 2020.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A790%3AFIN&amp;qid=1607079662423 (15 December 2020)

[7] As of December 2020 proposals for this agency have not been implemented

Such agencies and task forces have been established in the last five years and have been working as centres of knowledge and action to counter new hybrid threats. Estonia's Defense League runs an anti-propaganda blog,[8] with a focus not only on countering harmful narratives, but also on highlighting corporate practices related to social media – "outing" individuals and posts designed to further disinformation ("naming and shaming") – and advocating for media literacy. The Swedish government mandated a special investigator to analyse and submit proposals for the development of a Psychological Defence Authority, aimed at preserving the free exchange of knowledge and information in society to protect democracy and the rule of law. The mandate ranges from counteracting deception and disinformation (something the Swedish Defence Research Agency already does),[9] to ensuring that government authorities can effectively communicate with the public during crises.

Media literacy is part of the agenda for most countries, but mainstreaming actions throughout different ministries appears to have been more successful than more limited attempts; According to the 2019 Media Literacy Index[10] produced by the Open Society Institute – Finland leads all countries included, followed by Denmark, the Netherlands, Sweden and Estonia. Some examples might help explain why these countries have been successful in creating societal resilience against disinformation.

Finland comes first in the ranking because it put the issue at the top of the government's agenda, involving several ministries and departments and working with a different set of stakeholders, from journalists to political parties and citizens, to mainstream the issue.[11] Media literacy is a cross-departmental priority, and a key strategic aim of the Finnish Ministry of Education and Culture. An inter-ministerial approach is also part of the Danish Cyber and Information Security Strategy 2018-2021, which incorporates digital competencies throughout the educational system. The Swedish strategy includes five organizations acting at the national level, with 15 measures to mainstream media literacy and anti-disinformation measures across different groups in society, from students to public servants, to the general public and journalists.[12]

4.2.2 Legislative approaches of the EU-27

Some of these countries have also adopted legislation on the matter. A 2019 amendment[13] in **Denmark** criminalises the dissemination of disinformation that "aids or enables" a foreign state actor influence public opinion in the country. **Lithuania's** Law on Cyber Security was passed in 2018, with its Article 8(11) empowering the National Cyber Security Centre to order electronic communications providers, such as servers, to temporarily shut down (for up to 48 hours) without a court order if they are used to mount a "cyber incident", such as a disinformation attack. These countries focused on laws targeting mainly foreign interference, while not applying legal measures to potential disinformation challenges coming from within their own borders.

In 2020, **Germany** legislated on aspects of the online infrastructure behind disinformation campaigns. While the government cannot legislate on aspects of false or misleading content, this legislation acknowledges the role which ranking and recommendation functions (algorithms) on intermediary platforms

[8] PropaStop https://www.propastop.org (15 December 2020)

[9] Swedish Defence Research Agency https://www.foi.se/en/foi.html (15 December 2020)

[10] The Media Literacy Index 2019 https://osis.bg/?p=3356&amp;lang=en (15 December 2020)

[11] Rosalind Kenny Birch, "How Finland shuts down fake news" Apolitical, 3 December 2019. https://apolitical.co/en/solution_article/how-finland-shuts-down-fake-news (15 December 2020)

[12] B Hanh et al., "Digital Media Literacy: Sweden", Wonder, 22 December 2019. https://askwonder.com/research/digital-media-literacy-sweden-pq23nd3qg (15 December 2020)

[13] The amendment imposes a maximum penalty of 12-years' imprisonment for offenses carried out in connection with Danish or EU parliamentary elections (since social media platforms are not considered foreign powers, comments and posts published on Facebook or other social platforms are not covered by the amendment). Denmark, Law No. 269 of 26/03/2019 on Penal Code, https://perma.cc/Y4US-BTH3. (15 December 2020)

may have in privileging such false or misleading content, even unintentionally. To understand more about this danger, there needs to be more transparency in how these algorithms work. Germany's overhauled Federal Inter-State Media Law, which transposes the EU's revised Audiovisual Media Services Directive, requires media intermediaries (i.e., social media platform providers) to keep the following information easily and permanently available:

> 1. The criteria that determine the accessibility of content on a platform;
>
> 2. The central criteria for the aggregation, selection and presentation of content and their recommendation system, including information on the functioning of the algorithms used;
>
> 3. Identification and labelling of "social bots", without prohibiting their use.

As far as the question of paid ads are concerned, in 2019, the **Irish** government proposed legislation to regulate the transparency of online paid political advertising within election periods. The main objective is to "respect the role of the internet in the public sphere of political discourse and ensure that the public have access to legitimate information required in order to make autonomous voting decisions".[14]

In November 2020, the **Spanish** government approved the Procedure for Intervention against Disinformation. This involves the detection and analysis of disinformation campaigns, and of their possible impact on national security, as well as a role for the Secretary of State for Communication in conducting public campaigns to respond to disinformation campaigns.

One potential harmful consequence of such regulations is that they can be misused against journalists and civil society activists. In Europe and around the world, disinformation laws can have good intentions, but be misused or be cast so generally (in the case, for example, of defamation laws) that they curtail freedom of expression, by going after those who openly criticize public figures or the government.

As far as legislation directed at certain types of speech is concerned, **Croatia, Cyprus, France and Malta** all have a criminal provision on disinformation that may disturb public peace or alarm public opinion. These provisions date back to the pre-internet age. In the case of Cyprus the fines are small, and there has never been a punishment imposed on these grounds. Nevertheless, while the intention behind the measure is understandable, its vagueness raises concerns over its effect on freedom of speech.

Hungary has drawn criticism over recent additions to the penal code and their potentially negative effects on free speech. Hungary already had a provision in its penal code to penalise scaremongering (Article 337: "claiming or spreading a falsehood or claiming or spreading a distorted fact before a large public that could potentially alarm or agitate a large group of people during the site of a public emergency.") In the context of the COVID-19 pandemic, the parliament added a paragraph to the article, introducing penalties for "claiming or spreading a falsehood or a distorted truth before a large public during an emergency legal regime in a way

[14] Merrion Street, "Proposal to Regulate Transparency of Online Political Advertising", 5 November 2019, https://merrionstreet.ie/en/news-room/news/proposal_to_regulate_transparency_of_online_political_advertising.html (15 December 2015)

that is suitable for obstructing or preventing the successful defence.[15]" The provision was much criticized for its vague nature, use of terms such "site of public emergency", and its potential to stifle democratic debate and have a chilling effect on opinions other than those of the government opinion,[16] especially in a context where the government has repeatedly had extremely aggressive reactions to such opinions, even where they have been expressed by official expert bodies.[17]

With the beginning of the COVID-19 pandemic, **Bulgaria** made two attempts in the spring of 2020 alone to pass legislation on disinformation, but both failed.[18] A third attempt was under consideration as of December 2020.[19]

In **Romania**, President Klaus Iohannis issued a decree on 16 March 2020 declaring a state of emergency (initially for 30 days, then prolonged for another 30 days) in response to the COVID-19 pandemic. Article 54 of the decree limited the right to freedom of information, and has been used by the Ministry of Interior to suspend access, without a court order or clear means for judicial redress, to 15 online resources "spreading false news". On 15 May 2020, when the emergency period came to an end, access to the 15 websites was restored by the decision of the National Authority for Administration and Regulation in Communications (ANCOM).

France is the only state that specifically adopted a law against disinformation during electoral periods. Its main innovation is allowing a judge to rule within 48 hours upon referral that the spreadof the identified disinformation on a platform should be halted. The concern was that a targeted disinformation campaign could derail an electoral process, particularly close to election day (as was the attempt in the case of the "Macron leaks" during the French presidential elections in 2017[20]). The law has been criticized for its vague definition of disinformation. Between elections, the text establishes a duty of cooperation for these platforms, obliging them to introduce measures to address disinformation threats.

In conclusion, EU member states have been addressing the issue of disinformation mainly through task forces and policies like those to provide public education. They focus on a collaborative effort between society and government. In a novel development, there is has been legislation aimed at increasing the transparency of the ranking and recommendation systems and algorithms used by platforms. While narrow aspects of "false content" are penalised in many EU member states (namely slander and libel), criminal law has not been part of the response toolbox, except in Hungary, where it has followed the strong-man tendencies of the ruling party.

[15] Csaba Győry, "Fighting Fake News or Fighting Inconvenient Truths?", 11 April 2020. https://verfassungsblog.de/fighting-fake-news-or-fighting-inconvenient-truths/ (15 December 2020)

[16] Csaba Győry, "Fighting Fake News or Fighting Inconvenient Truths?", 11 April 2020. https://verfassungsblog.de/fighting-fake-news-or-fighting-inconvenient-truths/ (15 December 2020)

[17] Fanni Kaszás, "Coronavirus: Hungarian Medical Chamber's Public Proposals Not Welcomed by Gov't", 16 March 2020. https://hungarytoday.hu/coronavirus-hungarian-medical-chambers-public-proposals-not-welcomed-by-govt/ (15 December 2020)

[18] The first – the Emergency Bill – would have made the transmission of false information on the spread of infectious disease punishable by up to three years in prison (in case of serious damage five years). The draft bill was vetoed by the President. The second attempt was to make the Radio and Television Act apply to internet platforms. The media regulator, the Council for Electronic Media, would have new powers over disinformation in the internet environment, including powers to announce that a website spreads disinformation online, and to ask for a court order to discontinue access to the website. The attempt was rejected by the parliamentary Culture and Media Commission.

[19] The third attempt suggests that "disinformation in the internet environment" should be a theme to be included in the Personal Data Protection Act. If passed, the owners of websites, online blogs and, in certain cases, social networks might be responsible for the dissemination of disinformation online.

[20] This leak was promoted on Twitter by an army of trolls and fake accounts (bots), with the hashtag #MacronLeaks appearing in almost half a million tweets in twenty-four hours, and so the attack is now remembered as "the Macron Leaks." However, the leak itself was only the pinnacle of a coordinated operation that started months before, with a disinformation campaign and a hack.

5. Online Platforms and Online Hate Speech

5.1 Online hate speech at the EU level: Minimal involvement

The only binding instrument at the EU level addressing hate speech is the **Counter-Racism Framework Decision**,[21] which contains generic obligations to make racist and xenophobic speech punishable under criminal law, but does not mention online platforms. The EU has no significant competence on matters of substantive criminal law[22] – it only has competence to legislate on online speech when it constitutes terrorist or child sexual abuse content.

A voluntary, self/co-regulatory measure does exist – the Code of Conduct countering hate speech online, which online platforms are encouraged to sign. The Code of Conduct suffers from the absence of any obligations to produce verifiable data on the effect of measures taken on hate speech, such as what proportion of posts with identified hate speech have successfully been taken down.[23] Also, this is self-reported data, which lacks independent scrutiny.

5.2 Online hate speech at the EU-27 national level

Online hate speech is regulated more evenly across member states, as compared to approaches to tackling disinformation. Specific forms of hate speech are themselves criminalised in most states,[24] reflecting obligations under the EU's Counter-Racism Framework Decision. In addition, six countries have special laws on online hate speech, or differentiations in their general hate speech laws law depending on whether this occurs online or otherwise (Austria, the Czech Republic, France, Germany, Italy and Spain).

In **Italy**, in 2013, the Supreme Court extended the application of Article 416 of the Penal Code (criminal conspiracy) to hate speech perpetrated within virtual communities, blogs, chats and on social networks. In 2017, the application of "hate speech" was extended to cyberbullying – the instigation of hatred online towards individuals on various protected grounds.

To date, there are only three member states that have adopted (with varying success) or are planning to adopt legislation setting out specific responsibilities for online platforms (Germany, France and Austria). None of these laws create new categories of illegal content but, instead, impose obligations on online platforms to take down content that is illegal under the respective national laws. Despite being routinely referred to as "online hate speech laws", the scope of these acts is much wider, and covers, inter alia, defamation, terrorist content, public incitement to violence, endangering state security, child sexual abuse materials, insult and blasphemy.[25]

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

[22] According to article 83 (1) TFEU, the crime areas in question are: terrorism, trafficking in human beings and sexual exploitation of women and children, illicit drug trafficking, illicit arms trafficking, money laundering, corruption, counterfeiting of means of payment, computer crime and organised crime.

[23] European Commission, "Code of Practice on Disinformation One Year On" 29 October 2019. https://ec.europa.eu/commission/presscorner/detail/en/statement_19_6166 (15 December 2020)

[24] Estonia and Romania are two EU member-states critiqued for too narrow a definition of hate speech in their laws, as such, since 30 October 2020, they are subject to EU Commission's infringement proceedings.

[25] NetzDG captures 22 accounts of criminal law violations, its Austrian analogue Ko-Pl-G - 15.

Germany was the first country in the world to adopt a law mandating online platforms to respond to justified complaints by users by removing illegal content, on pain of fines of up to 50 million euros. The NetzDG came into force on 1 October 2017, after an intense political debate.

If a user complains against content that is "manifestly illegal", the platform has to remove it within 24 hours. In other cases, it has seven days to respond and to liaise with law enforcement agencies. The critics of the law point out that NetzDG does not require a court order prior to content removals. The task of law enforcement is "outsourced" to the platforms. In view of the strict time frames, the threat of regulatory fines might encourage platforms to remove more content than necessary, resulting in the excessive removal of content. The Ministry of Justice and Consumer Protection, in its report published in September 2020, has assessed the first three years of the act, and claimed that it did not lead to excessive removal.

Despite these downsides, the NetzDG has served as the inspiration for a similar **French** Law (the Law on Fighting Hateful Content on the Internet, or Loi Avia), as well as an **Austrian** analogue (the Communications Platform Law, or KoPl-G).[26]

Loi Avia was adopted in July 2019, but the French Constitutional Council struck down most of its provisions, finding that it infringed upon the freedom of expression beyond a degree proportionate to the aim pursued[27] and, hence, was unconstitutional. Despite the overturning of the Loi Avia, France did ultimately establish the "Online Hate Observatory",[28] officially launched in July 2020. Set up by the Conseil supérieur de l'audiovisuel (Superior Audio-visual Council), the role of the observatory is to monitor and analyse hateful online content, in collaboration with online platforms, associations and researchers.[29]

Like the NetzDG, the Austrian KoPl-G applies to a number of existing criminal offenses. Platforms must provide an accessible complaint mechanism if users allege that such offenses have been committed, and react either within 24 hours, if the content is manifestly illegal to laypersons, or within 7 days, if otherwise. The Austrian draft law includes stronger complaint and redress mechanisms than NetzDG. Under the KoPl-G, both parties (the complainant and the author of the content) can appeal the decision taken by the platform – first, to the platform itself, and after 14 days, to the arbitration body of the telecoms and media regulator. As to transparency reports, under the Austrian proposal, all decisions regarding content moderation are to be displayed in statistics, irrespective of whether the decision was taken under the law or under the internal community guidelines of the platform. This will provide more data, thus making it easier to assess the effectiveness of the law in the future.

Regulations such as the NetzDG have been criticized for providing a legal blueprint for authoritarian regimes seeking to repress freedom of speech. Similar laws have been adopted in Russia, Turkey, and Venezuela.[30]

[26] Communications Platform Law (in German – Kommunikationsplattformen-Gesetz). See: Republic of Austria, Bundesgesetz über Maßnahmen zum Schutz der Nutzer auf Kommunikationsplattformen, https://www.parlament.gv.at/PAKT/VHG/XXVII/ME/ME_00049/index.shtml (15 December 2020)

[27] See Para 8 of the decision of the Constitutional Council n ° 2020-801 DC: Republic of France, Decision of the Constitutional Council n ° 2020-801 DC, 18 June 2020. https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm&gt; (15 December 2020)

[28] Counseil Superieur de l'audiovisuel, "Lutte contre la haine sur internet: le CSA met en place un observatoire de la haine en ligne", 23 July 2020. https://www.csa.fr/Informer/Espace-presse/Communiques-de-presse/Lutte-contre-la-haine-sur-internet-le-CSA-met-en-place-un-observatoire-de-la-haine-en-ligne (15 December 2020)

[29] Alessandra Venier, "Platforms against hate speech", Inline Policy, 14 September 2020. https://www.inlinepolicy.com/blog/platforms-actions-against-hate-speech (15 December 2020)

[30] Jacob Mchangama and Joelle Fiss, The Digital Berlin Wall: How Germany (Accidentally) Created a Prototype for Global Online Censorship, Justicia, November 2019. http://justitia-int.org/wp-content/uploads/2019/11/Analyse_The-Digital-Berlin-Wall-How-Germany-Accidentally-Created-a-Prototype-for-Global-Online-Censorship.pdf (15 December 2020)

Both the **Czech Republic** and **Spain** differentiate between hate speech occurring online and offline, with both having harsher penalties for online instances. In Spain, certain types of hate speech carry aggravated penalties if expressed online, such as engaging in "public praise or justification" of terrorism (Article 578.2 of its penal code[31]). Presumably, the reason behind this is that online speech reaches further than that offline. The rights organisation Article 19 has critiqued this, saying that "the government has not made the for an increase in sentencing for online speech".[32]

In both **Hungary** and **Italy**, there is no legislation mandating that platforms restrict access to hate speech material without a prior court order or notice. However, case law has revealed a different picture, as courts in both countries have held platforms accountable for online hate speech.

In Hungary, the state lost a case before the European Court of Human Rights in 2016, after Hungarian courts had held online news portals liable for not deleting comments made by users in their comment sections.[33] In 2018, the state lost another case before the Court, which found that Hungarian courts had overreached in holding a news site liable for content to which a hyperlink in an article had referred.[34]

In conclusion, online hate speech is regulated relatively evenly across EU member states, through their penal codes. Only six countries have differentiation or specialised laws for hate speech occuring online. The most avant-garde approaches can be seen the "network enforcement acts" of Germany, Austria and France, which reflect the view that addressing hate speech online requires legal innovation. The current lack of transparent, verifiable and meaningful data from online platforms makes it difficult to measure the efficacy of these laws, much as is the case with the EU's Code of Conduct. One innovation in these laws is their recognition of online platforms' responsibilities in content spread – something other hate speech laws specific to the online sphere lack. Other approaches in the EU simply tend to create new crimes specific to the online sphere (e.g., cyberbullying), or result in different penalties, depending on whether the crime occurred online or offline.

[31] Spain, Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal, 23 November 1995. https://www.boe.es/buscar/act.php?id=BOE-A-1995-25444 (15 December 2020)

[32] Article 19, "Spain: Speech related offences of the Penal Code", March 2020, p. 13. https://www.article19.org/wp-content/uploads/2020/03/Spain-Penal-Code-analysis-March-2020-Final.pdf (15 December 2020)

[33] European Digital Rights, "MTE v. Hungary: the ECtHR rules again on intermediary liability", 10 February 2016. https://edri.org/our-work/mte-v-hungary-the-ecthr-rules-again-on-intermediary-liability/ (15 December 2020)

[34] Columbia Global Freedom of Expression, "Magyar Jeti Zrt v. Hungary", https://globalfreedomofexpression.columbia.edu/cases/magyar-jeti-zr-v-hungary/ (15 December 2020)

6. Concluding Remarks

Legislating on disinformation and online hate speech is a complex task. There is no silver bullet or a widely accepted best practice, and the interconnectivity of contemporary digital communications in spaces designed and moderated by a handful of private companies makes this environment a difficult one to regulate. So far, the EU has responded with voluntary Codes of Conduct and Practice. The European Democracy Action Plan and the Digital Services Act will change this, as the two initiatives introduce binding EU legislation on issues like paid political advertising, algorithmic transparency and electoral integrity. In the meantime, a few member states have defined their own approaches to tackling disinformation and online hate speech.

Online disinformation

- Several EU countries have developed cross-departmental task forces within their governments to address disinformation. These countries (the Baltic and Nordic countries) are also those whose populations score the best in digital literacy assessments, often seen as a long-term basis for tackling disinformation at the societal level.

- Another promising approach is one which shifts the focus from the obliging platforms to moderate content to obliging them to be more transparent about the algorithms they use to prioritize and recommend content. Germany is at the forefront of this approach, which promises to empower users to better understand how their content is curated, through its Inter-State Media Treaty. The European Commission announced that it will ultimately require similar transparency.

- The COVID-19 pandemic and resulting states of emergency have also been taken advantage by several EU governments, with worrying consequences on the freedom of speech and of the media, and in the shrinking of the civic space. This should serve as a reminder of the potential consequences of legislation – disinformation laws (even those with good intention) can be misused or be so general as to curtail freedom of expression, by going after journalists or activists who openly criticize public figures (defamation laws) or the government. If misused, legislation has the potential to curtail fundamental rights, and broad language and sanctions should be avoided. This is as true in the EU as it is in the rest of the world.

On online hate speech

- There is less discrepancy among EU member states on regulating online hate speech, reflecting obligations under the EU's Counter-Racism Framework Decision. A small group of countries (six) differentiate between online and offline hate in their laws.

- The most controversial, but also innovative approach is the "network enforcement act", with different variations in Germany, France and Austria. Based on notice-and-takedown systems, these laws have been criticized for "privatising the judiciary", as they delegate the responsibility for takingmdown content to the online platforms themselves. At the same time, this is an acknowledgement of the power and role that platforms have in the spread of illegal content, and deviates from the EU's self-regulatory Code of Conduct approach. This approach must be balanced carefully against freedom of expression concerns, and requires a strong rule of law culture to ensure that it is not misused – a criticism that has been raised by human rights advocates.

Legislation has the power to increase accountability, but as a rule of thumb, it should be guided towards increasing transparency and ensuring that educational actors, civil society, academia and the media are well funded, in order to properly shape and deliver anti-disinformation efforts, ranging from ensuring a more plural media environment to promoting greater media literacy. Sufficient funding is also needed to support research to shed light on how disinformation campaigns work and on the role of tech companies in enabling them. If misused, legislation has the potential to curtail fundamental rights, so broad language and sanctions, which can increase the opportunity for such misuse, should be avoided.

DEMOCRACY
REPORTING
INTERNATIONAL