CIVIL
LIBERTIES
UNION FOR
EUROPE

# HOW THE EU CAN MITIGATE DISINFORMATION WITHOUT HARMING FUNDAMENTAL RIGHTS

APRIL, 2021

# *Executive Summary*

The aim of fighting against disinformation is to rebuild trust in democracies and ensure that people can participate in democratic debates and freely form their opinions. Liberties is of the opinion that policy-makers have effective means to defeat disinformation by (i) reinforcing the integrity of online services, (ii) limiting the monetization of disinformation, (iii) empowering users to exercise their rights to get access to information, and (iv) strengthening open and non-discriminatory cooperation between platforms, academic researchers, and fact-checkers. However, successful mitigation of the harms of disinformation depends on proper enforcement of the law. The General Data Protection Regulation (GDPR)[1] ensures proper safeguards against targeted and tailored deceptive messages. The upcoming Digital Services Act should introduce meaningful and robust transparency mechanisms for online advertising, content policing, and algorithmic system developments. These laws have to be enforced and applied to the problem of disinformation. We advocate for meaningful, transparent, and enforceable rules. Even the best laws and self- and co-regulatory mechanisms are pointless if they are not enforced and backed up by an appropriate oversight mechanism. This report was written to feed the Commission's approach to review the Code of Practice of Disinformation in March 2021.

[1] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) https://eur-lex.europa.eu/eli/reg/2016/679/oj

# The key findings of our report

- The EU is under an obligation to respect the Charter of Fundamental Rights in addressing disinformation. Therefore, any solution to mitigate the impact of disinformation should respect the users' fundamental rights, primarily freedom of expression and the protection of personal data.

- One of the EU's goals is to promote the values on which it is founded, including democracy. Healthy democracy furthers the well-being of citizens and provides an environment where commercial enterprises can flourish. Private commercial interests cannot be allowed to undermine the very democracy that allows them to prosper.

- The EU's response to disinformation should not be to further empower big tech companies. Authorising or mandating tech companies to engage in more data gathering, more tracking, more monitoring, and more fact-checking would give them more information about their users. This will provide them with even greater power and influence and allow them to collect the very kind of information that makes disseminating disinformation possible and profitable.

- A common definition is needed for disinformation in order to elaborate the scope of any regulation properly, let that be self- or co-regulation or legislative approach. The definition should be limited to avoid possible over-regulation that would pose unjustifiable limitations on free speech.

- To protect democracy and fundamental rights, policy-makers, researchers and regulators must understand the impact of tech companies on fundamental rights and democratic debate. Therefore, tech companies must be transparent about their activities. They should provide coherent reports, meaningful data sets, and both state- and language-level databases.

- Tech companies must also be transparent about their algorithms. Tech companies might have legitimate interests in selling their goods and services and protecting their intellectual property. However, this cannot be accepted as a justification to bar users, researchers, and regulators from understanding what goals and criteria companies have built into the algorithms they use in order to protect democracy and fundamental rights.

- Strengthening measures to protect the integrity of their services against the use of manipulative techniques will limit the amplification of disinformation campaigns. Therefore, risk assessment and risk mitigation, service design, including the recommender system, content curation and moderation, and the advertising system should be transparent and auditable.

- We warn against 'real account policy' or suspending the opportunity to communicate anonymously. At-risk groups, such as members of the LGBTQ community, people who live with mental illness, or victims of domestic violence, are either targeted by their governments or face societal discrimination. These groups rely on anonymity to protect themselves and should not be deprived of access to services, such as social media platforms.

- Tech companies employ micro-targeting/ surveillance advertising by using user data as the basis for decisions about what adverts to show them. This misuse of data to manipulate users reinforces the need for a strong ePrivacy Regulation to change the balance of incentives for companies away from a model that relies on data harvesting and data dissemination.
- Fact-checking is a somewhat limited solution to counter disinformation. The mere act of signalling to users that content they see has been fact-checked as accurate or not can play a role in mitigating the impact of disinformation. However, further action such as removing or blocking content should be based on a transparent mechanism with safeguards such as human review.

## 1. Mitigate disinformation while preserving free speech

Any steps taken to address the problem of disinformation will necessarily have an impact on freedom of expression, freedom to access information, the right to personal data protection, fair elections and the functioning of our democracies. Therefore, the EU should be cautious about what measures it introduces, how much power it offers to and what responsibilities it imposes on tech companies to police their platforms and their users.

The EU is under an obligation to respect the Charter of Fundamental Rights in addressing disinformation. The EU's obligation to protect free speech implies that unwanted content such as disinformation and misinformation will always exist to a degree. Furthermore, disinformation on online platforms is not the cause but rather a symptom of broader societal problems, such as the dysfunction of politics, racism, sexism and inequality. It is not possible to eliminate disinformation without addressing these underlying factors, and trying to do so by crudely regulating speech over the internet would result in violations of freedom of expression. Therefore, Liberties is of the opinion that the most the EU can aspire to is to create an environment where disinformation is less likely to thrive and to mitigate the problems caused by disinformation, rather than to eradicate it completely.

As the Council of Europe Commissioner for Human Rights' statement warns, steps to combat disinformation that limit rights such as media freedom should be necessary, proportionate, and subject to regular oversight, including by Parliament and national human rights institutions.[2] This warning is echoed by the Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda (2017), which declares: "[S]tates may only impose restrictions on the right to freedom of expression in accordance with the test for such restrictions under international law, namely that they be provided for by law, serve one of the legitimate interests recognised under international law, and be necessary and proportionate to protect that interest."[3]

The discourse about disinformation focuses on social media and online platforms. However, evidence, such as the outcome of the 2016 US election or Brexit, shows that disinformation capable of influencing the democratic process is also carried over mainstream media.[4] Therefore, the Commission should include the traditional mainstream media within the scope of its efforts to mitigate disinformation, including public service media in EU countries where this is controlled by the government, such as Hungary or Poland.[5]

## 2. No extra power to the powerful

Although social media is not the root cause of disinformation, it does intensify the impact of false information. The business model of platforms such as Facebook and Twitter, and other tech giants such as Google and Amazon, is based on monetizing information of any kind, including disinformation. These companies

[2] Statement from the Council of Europe Commissioner for Human Rights: Press freedom must not be undermined by measures to counter disinformation about COVID-19, 2020. https://www.coe.int/en/web/commissioner/view/-/asset_publisher/ugj3i6qSEkhZ/content/press-freedom-must-not-be-undermined-by-measures-to-counter-disinformation-about-covid-19?_101_INSTANCE_ugj3i6qSEkhZ_languageId=en_GB

[3] The United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information, 2017. https://www.ohchr.org/_layouts/15/WopiFrame.aspx?sourcedoc=/Documents/HRBodies/SP/JointDeclaration3March2017.doc&action=default&DefaultItemOpen=1

[4] Benkler, Y., Faris, R, Roberts, H. Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics, Oxford University Press, 2018. p. 28.

[5] Public Media Alliance, Threats to media independence continue across Central Europe 2019. https://www.publicmediaalliance.org/threats-to-media-independence-continue-across-central-europe/

already have undue influence and power over culture, society, the economy and politics.[6]

The EU's response to disinformation should not end up further empowering these companies. Authorising, encouraging or mandating these companies to engage in more data gathering, more tracking, more monitoring, and more fact-checking would give them even more information about their users. This will not only give them even greater power and influence, but also allow them to collect the very kind of information that makes disseminating disinformation possible and profitable. (See also point 6)

We call on the Commission to change the paradigm and require more transparency and proper algorithm design from these companies, instead of entrusting them with more control over the users, their data, and their content.

### *3. The missing definition*

Creating any self- or co-regulatory mechanism is a challenge, especially if there is no common understanding of what is being regulated. The COVID-19 pandemic and related disinformation campaigns proved that it is impossible to regulate the field if stakeholders do not agree on the scope of the Code of Practice of Disinformation. As stated in the Joint Communication to the European Parliament, the European Council, the Council, the European Economic, and Social Committee and the Committee of the Regions of Tackling COVID-19 disinformation - Getting the facts right (2020), "one of the lessons learned from this crisis is the need to clearly differentiate between the various forms of false or misleading content revealed by the 'infodemic' and to calibrate appropriate responses."[7]

Liberties advocates[8] for using the definition, with slight edits, elaborated in the 'Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda' issued jointly by the United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and

[6] Bernal, P. The Internet, Warts and All, Free Speech, Privacy and Truth, Cambridge University Press, 2018. p. 257.

[7] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020JC0008

[8] We  have been arguing for this understanding since 2018. See Access Now, Civil Liberties Union for Europe, EDRi, Shadow Report, Informing the Disinformation Debate, 2018. https://dq4n3btxmr8c9.cloudfront.net/files/2r7-0S/online_disinformation.pdf

Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information (the "special mandate holders").[9] Liberties bases its suggestion for slight adjustments to this definition on the work of Yochai Benkler, Robert Faris and Hal Roberts, which includes two components:[10]

*disinformation:* statements that are known or reasonably should be known to be false. It manipulates and/or misleads the population intentionally to achieve political ends, and, as a side effect, it interferes with the public's right to know and the right of individuals to seek, receive, and impart information.

*propaganda:* statements that demonstrate a reckless disregard for verifiable information.

Here Liberties would add *misinformation:* it is false information, but the person who is disseminating it believes it to be true and is publishing it without meaning to be wrong or having a political purpose in communicating the false information.

## 4. Transparency

### 4.1 Data disclosure

The Code of Practice from 2018[11] set out self-regulatory standards on behalf of tech companies to fight disinformation. However, it did not achieve its goal of enabling access to data held by platforms to allow for monitoring, fact-checking, and research activities. Platforms are still secretive about their practices and shield their actions by invoking trade secrets, copyright, or even the GDPR to avoid the mandatory transparency requirements, even though the GDPR clearly states that the "Regulation does not, therefore, concern the processing of such anonymous information, including for statistical or research purposes".

The annual self-assessment reports also demonstrate the shortcomings of the Code.[12] They contain very minimal information about disinformation-related activities, the information shared by signatories is difficult to compare, and they do not help evaluate the mechanisms introduced by tech companies to

[9] https://www.ohchr.org/Documents/Issues/Expression/JointDeclaration3March2017.doc (Microsoft Word document)

[10] Benkler, Y., Faris, R, Roberts, H. ibid. p. 24.

[11] https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation

[12] Annual self-assessments are available here: https://ec.europa.eu/digital-single-market/en/news/annual-self-assessment-reports-signatories-code-practice-disinformation-2019

fight against disinformation and manipulation campaigns.

The Commission should at a minimum require platforms and tech companies to provide coherent reports, meaningful data sets and state- and language-level databases. If not voluntarily, then by mandatory legal obligation.

Liberties is of the opinion that transparency should be a multi-layer obligation. Multi-layer obligation means that different stakeholders need different data sets to fulfill their roles. Regulators, researchers, fact-checkers, civil society, and the general public need different data sets. Data required by individuals are regulated by the GDPR, ensuring the right of access by the data subject to their personal data under Article 14 (2 )(c) and Article 15 (1) of the GDPR.

We also call on the Commission to oblige platforms and tech companies to adhere to the proper understanding of the GDPR. The latter does not in any way hinder their ability to work transparently and share relevant data about their algorithms (see point 4.2). The GDPR protects users from micro-targeting techniques, profiling, and content curation based on their data – practices that many of the signatories do on a daily basis.

The European Democracy Action Plan (EDAP)[13] flags the need to ensure effective data disclosure for research on disinformation by developing a framework in line with applicable regulatory requirements and based on the involvement of all relevant stakeholders. David Kaye, a former Special Rapporteur of the United Nations, states in his report, "transparency includes knowing what rules States and companies use to moderate content, the rules regarding content, how those rules are applied, what kind of appeals process exists and what kind of accountability there is for wrongful take down of content."[14]

### *4.2 Algorithmic transparency and the right to receive an explanation*

Algorithms are not neutral servants of freedom of expression or freedom of information. There is a common myth that technology is neutral and that therefore tech companies ensure platform neutrality and algorithmic neutrality to overcome biased individual decision-making. These systems and services are not neutral because they are designed by people, and people have assumptions and

[13] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2020%3A790%3AFIN&qid=1607079662423https://ec.europa.eu/info/strategy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan_en

[14] UN Human RIghts Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 2018, https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/096/72/PDF/G1809672.pdf?OpenElement

a certain logic that informs their actions.[15] The neutrality of technology was a leading concept 20 years ago, and it is tangible in the eCommerce Directive,[16] considering merely 'technical' and 'automatic' processing is 'passive' and neutral, creating limited liability for platforms.[17] The eCommerce Directive treated intermediaries, hosting user content services, as neutral. But content curation, AI, and other technical developments proved that services have a direct influence on content offered to users. The upcoming regulation, including both DSA and the Code of Practice, should change this assumption. Platforms should be held liable, not for user-generated content, but for the algorithms they create.

Search engines, news curation and profiling all depend on algorithms. Tech companies use algorithms to make choices that can produce outcomes harmful to individuals and to society. The legitimate interests that tech companies might have in selling their goods and services and protecting their intellectual property cannot bar users, researchers and regulators from understanding what goals and criteria companies have built into the algorithms they use for the reasons of protecting democracy and fundamental rights. An obligation for transparency of algorithms is essential. However, it is not enough in itself  to understand algorithms and prevent users and society from harm. It is unreasonable to expect the average user to dig deeply into data to understand algorithms. Rather, tech companies should be obliged to provide users with a clear explanation of the criteria applied by algorithms. Article 13(2)(f)[18] of the GDPR requires tech companies to provide an explanation of automated decision-making. This provision applies to targeted disinformation campaigns because profiling and content

[15] Turner N, Resinck P., Barton G. Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms, Center for Technology and Innovation, 2019. https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/

[16] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce')

[17] Bernal, P. The Internet, Warts and All, Free Speech, Privacy and Truth, Cambridge University Press, 2018. p. 71.

[18] In addition to the information referred to in paragraph 1, the controller shall, at the time when personal data are obtained, provide the data subject with the following further information necessary to ensure fair and transparent processing: (f) | the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

curation is personal data processing. We call this the right to explanation, which is connected to the integrity of services (see point 5).

## 5. Integrity of services

As set out in EDAP, the new Code of Practice needs to require online platforms to strengthen the integrity of their services against the use of manipulative techniques to limit the amplification of disinformation campaigns. Therefore, we believe that risk assessment and risk mitigation, service design, including the recommender system, content curation and moderation, and the advertising system should be transparent and auditable.

### 5.1 Profile takedowns and the real-name policy

The Oxford Internet Institute's report states that platforms combat disinformation campaigns and cyber troops disrupting elections mostly through account takedowns.[19] Account takedown, however, is the most blurry and crude response employed by platforms. Platforms remove accounts without proper justification, offering non-transparent appeal mechanisms without the possibility for human interaction. Account takedowns often happen to politicians, artists, and political activists. Sometimes the removal process is triggered by malevolent requests from rival political actors who wish to silence their opposition.

Signatories of the Code should commit themselves to change the account takedown measures and introduce safeguards, such as mandatory explanation to the users, human interventions, proper step-by-step mechanisms with the possibility of meaningful counter-argument, and redress mechanisms.

We warn against 'real account policy' or suspending the opportunity to communicate anonymously. At-risk groups, such as members of the LGBTQ community, people who live with mental illness, or victims of domestic violence, are either targeted by their governments or face societal discrimination. These groups rely on anonymity to protect themselves and should not be deprived of access to services, such as social media platforms.

[19] Public announcements by Facebook and Twitter reveal that between January 2019 and November 2020 more than 10,893 Facebook accounts, 12,588 Facebook pages, 603 Facebook groups, 1,556 Instagram accounts, and 294,096 Twitter accounts were taken down by the platforms (see Figure 2). In this timeframe, Facebook also reported that almost US $10 million was spent on political ads. https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/127/2021/02/CyberTroop-Report20-Draft9.pdf
Industrialized Disinformation 2020 Global Inventory of Organized Social Media Manipulation, Oxford Internet Institute, Oxford University.

### *5.2 The illusion of choice*

Social media is designed primarily to share and get access to information. The business model behind it is based on data collection about users and monetization of user profiles. Social media is not designed in a way that allows users to decide what they can see on the platforms they use or to determine what information is collected about them. This situation has to change. Users should be empowered to make an informed decision about their personal data and who has access to it. If platforms use content curation, they should inform users how and why they see the content that platforms present to them.

However, the opportunity to choose does not fit with the reality we have now. Our life is organized around big platforms, and the key functions are so complex that people do not have the time or resources to invest in researching their choices. We have learned this from the failure of cookie banners, which give the appearance of allowing users to choose what cookies apply, but through a design that only creates consent fatigue. The illusion of choice only makes people more vulnerable, because they are likely to behave according to a mistaken belief that their data is safe.

### *5.3 Algorithmic accountability and audit*

Companies need to be held accountable for the algorithms they use. An algorithmic audit means testing and analyzing specific harms caused by algorithms. Algorithmic accountability and audit have been advocated by academia, such as AI Now Institute and others.[20] The algorithmic audit should not be left in the hands of tech companies themselves. Besides tech companies' internal audits, researchers and regulators should have access to information that allows them to conduct independent audits.

## *6. Addressing the online manipulation business model*

Tech companies' policies about what content to show and promote to users and what to sideline, and the way they police content is driven almost purely by their economic interests. The business model of online platforms – the monetization of disinformation – is the core problem.[21]

As the European Data Protection Supervisor (EDPS) stated, measures to foster online accountability have "focused on transparency

[20] https://ainowinstitute.org/

[21] In 2018 Liberties with Access Now and EDRi submitted a Shadow Report to the Report of the High Level Expert Group on Fake News and Online Disinformation. Even though almost three years have passed, our statements regarding the business model are still valid. https://dq4n3btxmr8c9.cloudfront.net/files/2r7-0S/online_disinformation.pdf

measures, exposing the source of information while neglecting the accountability of players in the ecosystem who profit from harmful behaviour."[22] When discussing harmful behaviour that promotes disinformation, it is of paramount importance to separate two issues. First, the role of online platforms and the economic interests behind the spreading of dis-/misinformation. Second, state-led "hybrid threats" such as cyber attacks and disinformation campaigns. When it comes to the economic aspect associated with online platforms, the EDPS rightly points out that "fake news is a symptom of concentrated, unaccountable digital markets, constant tracking and reckless handling of personal data".

Certain contemporary political campaigns have been successful in spite of an easily demonstrable lack of respect for basic facts. This phenomenon has been aided in part by the use of social media; specifically, platforms that profit from the collection and analysis of user data. Such data processing operations are based on promoting media that is more likely to spread while disregarding the veracity of the content. The more sensational the "news", the more a user's attention is grabbed and the more profiling data is generated – and it is such profiling data that generates profits for the platform.

Companies such as Facebook employ micro-targeting/surveillance advertising by using user data as the basis for decisions about the advertisements that users see in their news feeds, based on what will likely appeal to them and what they will subsequently engage with and click on. This type of data manipulation reinforces the need for the ePrivacy Regulation to enter into force as a means of changing the balance of incentives for companies away from a model that relies on sensationalism and shock. The ePrivacy Regulation could explicitly stop surveillance advertising techniques, such as the use of third-party cookies and other tracking methods. This needs to change to ensure that the right to privacy in the electronic communications sector is prioritised ahead of current unsustainable approaches. The New York Times investigated one of the widely known disinformation stories of the 2016 US presidential election and found it to be motivated by advertising revenue that was successfully generated by Google.[23] Simply encouraging platforms to adopt ineffective mechanisms of removal or verification (such as flagging and 'disputed tags') cannot solve the problem while the fundamental business model of the platform itself facilitates or propagates disinformation.

[22] https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf

[23] http://europa.eu/rapid/press-release_IP-18-5681_en.htm

## *7. Fact-checking is a limited solution*

According to a recent study published by the Center for Media, Data and Society at the Central European University, there are several challenges fact-checkers face. One of the main challenges for fact-checkers seems to be difficulty in reaching their audiences. The study finds that the "impact of fact-checking remains a research gap as there is no solid evidence to understand how effective fact-checking is."[24]

We have to add that fact-checkers are not neutral decision-making bodies about speech, and we should not allow them to police speech. The verdicts of independent fact-checkers often come under scrutiny for ideological reasons, reliability of data, or inherent bias. Nonetheless, fact-checking may not be sufficient to combat skepticism towards the media system or the lack of trust in democratic institutions. The mere act of alerting users that the content they are seeing has been fact-checked as true or false can play a role in mitigating the impact of disinformation. But a further decision to remove or block content should be based on a transparent mechanism with safeguards such as human review.

[24] What keeps fact-checking organizations up at night?, Report, Central European University, Center for Media, Data and Society, 2021. https://cmds.ceu.edu/sites/cmcs.ceu.hu/files/attachment/article/2006/whatkeeps-fact-checkingorganizationsupatnight.pdf

The Civil Liberties Union for Europe (Liberties) is a non-governmental organisation promoting and protecting the civil liberties of everyone in the European Union. We are headquartered in Berlin and have a presence in Brussels. Liberties is built on a network of national civil liberties NGOs from across the EU. Unless otherwise indicated, the opinions expressed by Liberties do not necessarily constitute the views of our member organisations.

**Website:** liberties.eu

**Contact info:** info@liberties.eu

Written by Eva Simon

**Address:**
The Civil Liberties Union for Europe e. V.
Ringbahnstr. 16-20
12099 Berlin Germany