



Doc. 15537
23 May 2022

The control of online communication: a threat to media pluralism, freedom of information and human dignity

Report¹

Committee on Culture, Science, Education and Media

Rapporteur: Mr Frédéric REISS, France, Group of the European People's Party

Summary

Online communication has become an essential part of people's daily lives. Therefore, it is worrying that a handful of large tech companies are *de facto* controlling online information flows. To address the dominance of a few internet intermediaries in the digital marketplace, member States should use anti-trust legislation. Crucial issues for internet intermediaries and for the public are the quality and variety of information, and plurality of sources available online.

The use of artificial intelligence and automated filters for content moderation is neither reliable nor effective. The role and necessary presence of human decision makers, as well as the participation of users in the establishment and assessment of content moderation policies are crucial. Legislation should deal with "illegal content" and avoid using broader notions such as "harmful content".

Member States must strike a balance between the freedom of private economic undertakings, with the right to develop their own commercial strategies, including the use of algorithmic systems, and the right of the public to communicate freely online, with access to a wide range of information sources. Internet intermediaries must assume their responsibility to ensure a free and pluralistic flow of information online which is respectful of human rights.

1. Reference to committee: [Doc. 15041](#), Reference 4496 of 6 March 2020.



Contents	Page
A. Draft resolution	3
B. Explanatory memorandum by Mr Frédéric Reiss, rapporteur	6
1. Introduction	6
2. Hosting services as an economic activity with gatekeeping powers	7
2.1. Platforms and market power	7
2.2. Platforms and informational power	8
2.3. Disinformation and elections as particular examples	10
3. Considerations regarding possible solutions	11
4. Conclusions	14

A. Draft resolution²

1. The Parliamentary Assembly holds that communication policies must be open, transparent and pluralistic, and it must build on unhindered access to information of public interest and the responsibility of those disseminating information to society. It notes that online communication has become an essential part of people's daily lives and is concerned that a handful of internet intermediaries are *de facto* controlling online information flows. This concentration in the hands of a few private corporations gives them huge economic and technological power, as well as the possibility to influence almost every aspect of people's private and social lives.
2. There are questions on the capacity and willingness of an economic-technological oligopoly to ensure diversity of information sources and pluralism of ideas and opinions online; on the expediency of entrusting artificial intelligence with the task of monitoring online pluralism; and on the real capacity of legal frameworks and democratic institutions in place to prevent the concentration of economic-technological-informational power from being converted into non-democratic political power. Indeed, as electoral communication shifts to the digital sphere, whoever controls online communication during election campaigns may become a formidable political force. Voters may be seriously encumbered in their decisions by misleading, manipulative and false information.
3. Main risk factors in this context are: the lack of transparency of new forms of online advertising, which can too easily escape the restrictions applicable to advertising on traditional media, such as those intended to protect children, public morals or other social values; the fact that journalists, whose behaviour is guided by sound editorial practices and ethical obligations, are no longer the ones playing the gatekeeper role; and the growing amount of disinformation available online, in particular when it is strategically disseminated with the intent to influence election results.
4. From an economic point of view, network effects and economies of scale create a strong tendency towards market concentration. In the context of oligopolistic competition driven by technology, inefficiencies and market failures may come from the use of market power to discourage the entrance of new competitors, from the creation of barriers to switching services or from information asymmetries. Therefore, to address the dominance of a few internet intermediaries in the digital marketplace, member States should use anti-trust legislation. This may enable citizens to have greater choice when it comes to choosing, to the extent possible, platforms that are likely to better protect their privacy and dignity.
5. A few innovative remedies to mitigate the power of internet intermediaries include giving users, where possible, the option of accessing, consulting and receiving services from third-party providers of their choice, which would rank and/or deliver content following a classification previously made by the user him/herself and could alert him/her in case of violent, shocking or other dangerous content.
6. Beyond the business model, crucial issues for internet intermediaries and for the public are the quality and variety of information, and plurality of sources available online. Internet intermediaries are increasingly using algorithmic systems, which are helpful to search the internet, automatically create and distribute content, identify potentially illegal content, verify information published online and moderate the communication online. However, algorithmic systems can be abused or used dishonestly to shape information, knowledge, the formation of individual and collective opinions and even emotions and actions. Coupled with economic and technological power of big platforms, this risk becomes particularly serious.
7. With the emergence of internet intermediaries, harmful content is spreading at a very high speed on the web. Internet intermediaries should be particularly mindful of their duty of care where they produce or manage the content available on their platforms, or where they play a curatorial or editorial role, while avoiding taking down third-party content, except for clearly illegal content.
8. The use of artificial intelligence and automated filters for content moderation is neither reliable nor effective. Big platforms already have a long record of mistaken or harmful content moderation decisions in areas such as terrorist or extremist content. Solutions to policy challenges such as hate speech, terrorist propaganda and disinformation are often multifactorial; therefore, mandating automated moderation by law is an inappropriate and incomplete solution. It is important to acknowledge and properly articulate the role and necessary presence of human decision makers, as well as the participation of users in the establishment and assessment of content moderation policies.

2. Draft resolution adopted unanimously by the committee on 10 May 2022.

9. Today there is a trend towards the regulation of social media platforms. Whilst increased democratic oversight is necessary, regulation enacted in practice often entails overbroad power and the discretion of government authorities over information flows, which endanger freedom of expression. Lawmakers should aim at reinforcing transparency and focus on companies' due processes and operations, rather than on the content itself. Moreover, legislation should deal with "illegal content" and avoid using broader notions such as "harmful content".

10. If lawmakers choose to impose very heavy regulations on all internet intermediaries, including new smaller companies, this might consolidate the position of big actors which are already in the market. In such a case, new actors would have little chance of entering the market. Therefore, there is a need for a gradual approach, to accommodate different types of regulations on different types of platforms.

11. The Parliamentary Assembly recalls that, in its [Recommendation CM/Rec\(2018\)2](#) on the roles and responsibilities of internet intermediaries, the Committee of Ministers of the Council of Europe indicates that any legislation should clearly define the powers granted to public authorities as they relate to internet intermediaries; and [Recommendation CM/Rec\(2020\)1](#) on the human rights impacts of algorithmic systems confirms that the rule of law standards must be maintained in the context of algorithmic systems.

12. Internet intermediaries must ensure a certain degree of transparency of the algorithmic systems they use, because this may have an impact on our freedom of expression. At the same time, in their capacity as private companies, they should enjoy, without prejudice for effective transparency and for human rights, their legitimate right to commercial secrecy. Member States must strike a balance between the freedom of private economic undertakings, with the right to develop their own commercial strategies, including the use of algorithmic systems, and the right of the public to communicate freely online, with access to a wide range of information sources. They should also recognise that content removal is not in itself a solution for societal harms, as more rigorous content moderation may displace the problem of online hate speech to less popular platforms rather than address its causes.

13. Internet intermediaries have the responsibility to ensure the protection of users' rights, including freedom of expression. Therefore, member States should ensure internet intermediaries' accountability for the algorithmic systems they develop and use in the automated production and distribution of information, as well as for their lines of funding and policies they implement for creating information flows and dealing with illegal content.

14. In particular, internet intermediaries should assume specific responsibilities based on international standards and national legislation regarding users' protection against manipulation, disinformation, harassment, hate speech and any expression which infringes privacy and human dignity. The functioning of internet intermediaries and technological developments behind their operation must be guided by high ethical principles. It is from both a legal and ethical perspective that internet intermediaries must assume their responsibility to ensure a free and pluralistic flow of information online which is respectful of human rights.

15. Consequently, the Assembly calls on Council of Europe member States to:

15.1. bring their legislation and practice into line with [Recommendation CM/Rec\(2020\)1](#) on the human rights impacts of algorithmic systems, and [Recommendation CM/Rec\(2018\)2](#) on the roles and responsibilities of internet intermediaries;

15.2. consider whether the concentration of economic and technological power in the hands of a few internet intermediaries can be properly dealt with via general and already existing competition regulations and tools;

15.3. use anti-trust legislation to force monopolies to divest a part of their assets and reduce their dominance in the digital markets;

15.4. develop a gradual regulatory approach to accommodate different types of regulations to different types of internet intermediaries, with the aim to avoid pushing new actors outside the market or enabling them to enter the market;

15.5. address the issue of anticompetitive conduct in digital markets by strengthening the enforcement of regulations on merging and abuse of monopolistic positions;

15.6. guarantee that any legislation imposing duties and restrictions on internet intermediaries with an impact on users' freedom of expression be exclusively aimed at dealing with "illegal content" thus avoiding broader notions such as "harmful content";

- 15.7. ensure that mere automated content moderation is not allowed by the legislation; in this connection, encourage internet intermediaries, via legal and policy measures, to:
 - 15.7.1. allow users to choose means of direct and efficient communication which do not solely rely on automated tools;
 - 15.7.2. ensure that where automated means are used, the technology is sufficiently reliable to limit the rate of errors where content is wrongly considered as illegal;
- 15.8. guarantee that legally mandated content moderation provides for the necessary presence of human decision makers, and incorporates sufficient safeguards so that freedom of expression is not hampered;
- 15.9. encourage, via legal and policy measures, the participation of users in the establishment and assessment of content moderation policies;
- 15.10. ensure that regulation enacted to ensure transparency of automated content moderation systems is based on a clear definition of information that is necessary and useful to disclose and of public interest that legitimises the disclosure;
- 15.11. support the elaboration and respect of a general framework of internet intermediaries' ethics, including the principles of transparency, justice, non-maleficence, responsibility, privacy, rights and freedoms of users;
- 15.12. encourage internet intermediaries, via legal and policy measures, to counteract hate speech online by issuing warning messages to persons who spread hate speech online or by inviting users to review messages before sending them; encourage internet intermediaries to add such guidelines to the codes of conduct dealing with hate speech;
- 15.13. consider adapting election legislation and policies tuned to the new digital environment by reviewing provisions on electoral communication; in this respect, reinforce accountability of internet intermediaries in terms of transparency and access to data, promote quality journalism, empower voters towards a critical evaluation of electoral communication and develop media literacy.

B. Explanatory memorandum by Mr Frédéric Reiss, rapporteur

1. Introduction

1. The United Nations Human Rights Council declared in its Resolution A/HRC/RES/32/13 of 1 July 2016 that “[...] the same rights that people have offline must also be protected online, in particular freedom of expression, which is applicable regardless of frontiers and through any media of one’s choice, in accordance with articles 19 of the Universal Declaration of Human Rights and of the International Covenant on Civil and Political Rights.” In doing so, it recalled its Resolutions A/HRC/RES/20/8 of 5 July 2012 and A/HRC/RES/26/13 of 26 June 2014, on the subject of the promotion, protection and enjoyment of human rights on the Internet.

2. The internet is a technological platform that can be utilised for different purposes and for the provision of very different types of services. Therefore, the internet cannot be seen as a medium comparable to the press, radio and television, but rather as a distribution platform which can facilitate the interaction between different kinds of providers and users.

3. The internet is not only a space for what is generally understood as public communication but also a platform that enables the operation of different types of private communication tools (email and private messaging applications in general) and has also created new hybrid modalities which incorporate characteristics of both (for example, applications that allow the creation of small groups or communities and the internal dissemination of content).

4. Besides its communications features, the Internet has also proved to be a very powerful tool to facilitate the deployment of the so-called digital economy. The Covid-19 pandemic has particularly intensified the electronic acquisition or provision of goods and economic services thus creating new spaces for new business models to prosper.

5. In the course of the recent years, a very specific category of online actors has taken a central place in most legal and policy debates.

6. In 2018, the Committee of Ministers of the Council of Europe adopted a Recommendation on the role and responsibilities of Internet intermediaries,³ which described these actors as “(a) wide, diverse and rapidly evolving range of players”, which:

“facilitate interactions on the internet between natural and legal persons by offering and performing a variety of functions and services. Some connect users to the internet, enable the processing of information and data, or host web-based services, including for user-generated content. Others aggregate information and enable searches; they give access to, host and index content and services designed and/or operated by third parties. Some facilitate the sale of goods and services, including audio-visual services, and enable other commercial transactions, including payments”.

7. The intermediaries have become main actors in the process of dissemination and distribution of all types of content. The notion of “intermediaries” refers to a wide range of online service providers including online storage, distribution, and sharing; social networking, collaborating and gaming; or searching and referencing.⁴

8. This report will focus on what are generally known as hosting service providers, and particularly those who tend to engage in granular content moderation. This includes services provided by social media platforms like Facebook or Twitter, content sharing platforms such as YouTube or Vimeo, and search engines like Google or Yahoo.

9. These providers play an important role as facilitators of the exercise of the users’ right to freedom of expression. This being said, it is also true that the role and presence of online intermediaries raises two main areas for concern. Firstly, the important presence of economies of scale, economies of scope as well as network effects favours a high degree of concentration which may also lead to significant market failures. Secondly, the biggest players in these markets have clearly become powerful gatekeepers who control

3. Available online at: <https://rm.coe.int/1680790e14>.

4. See the comprehensive and detailed categorisation provided by Joris van Hoboken, João Pedro Quintais, Joost Poort, Nico van Eijk, “Hosting intermediary services and illegal content online. An analysis of the scope of article 14 ECD in light of developments in the online service landscape”, European Commission – DG Communications Networks, Content and Technology and University of Amsterdam-Institute for Information Law, 2018. Available online at: <https://op.europa.eu/en/publication-detail/-/publication/7779caca-2537-11e9-8d04-01aa75ed71a1/language-en>.

access to major speech fora. Like in the case of – in many aspects, still to be solved – concentrated actors exercising a bottleneck power in the field of legacy media (particularly the broadcasting sector), values such as human dignity, pluralism and freedom of expression⁵ also need to be particularly and properly considered and incorporated into the legal and policy-making debates around platform regulation.

10. The object of this report is to elaborate on these issues in more detail, with a particular focus on how to preserve the mentioned values in a context where no unnecessary and disproportionate constraints are imposed, and different business models can thrive to the benefit of users and the society as a whole.

11. My analysis builds on the background report by Mr Joan Barata,⁶ who I warmly thank for his outstanding work. I have also taken account of the contributions by other experts,⁷ and by several members of the committee.

2. Hosting services as an economic activity with gatekeeping powers

2.1. Platforms and market power

12. Online platforms constitute a particularly relevant new actor in the public sphere, first of all in terms of market power. They operate on the basis of what is called network effects. In other words, the greater the number of users of a certain platform, the greater the benefits obtained by all their users, and the more valuable the service becomes to them. These network effects are cross sided as the benefits or services received by a part of the users (individual users of social media, for example) are subsidised by other participants (advertisers). In view of this main characteristic, it is clear that the success of these platforms depends on the acquisition of a certain critical mass and, from there, on the accumulation of the largest possible number of end users.

13. The British regulatory body OFCOM has remarked, in a recent report on “Online market failures and harms”, that besides the mentioned network effects, companies benefit from cost savings due to their size (economies of scale) or their presence across a range of services (economies of scope).⁸ Of special economic importance is also the use of data and algorithms: the collection of data regarding users’ habits and characteristics may not only improve their experience through personalised services but also facilitate a more targeted advertising based on a profound knowledge of consumers’ preferences and needs. This creates the necessity to collect the highest possible amount of data from users, as well as disincentivises the users’ search for alternative services. On top on all these elements, it is also important to underscore the fact that business models are becoming more complex and sophisticated as big online platforms often combine different offers of services: social media, search, marketplaces, content sharing, etc. For this reason, the identification of relevant markets for the purpose of competition assessments can become particularly complex.

14. Still from a strictly economic point of view, network effects and economies of scale create a strong tendency towards market concentration. In this context of oligopolistic competition driven by technology, inefficiencies and market failures may come from the use of market power to discourage the entrance of new competitors and the creation of barriers to switching services or from information asymmetries.⁹

15. In the United States, in June 2019, the House Judiciary Committee announced a bipartisan investigation into competition in digital markets. The Subcommittee on Antitrust, Commercial and Administrative Law examined the dominance of Amazon, Apple, Facebook, and Google, and their business practices to determine how their power affects the American economy and democracy. Additionally, the

5. The term “freedom of expression” is used here in the sense given to it by article 10 of the European Convention on Human Rights, namely “Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers.”

6. Research Fellow, Program on Platform Regulation, Cyber Policy Center, Stanford University.

7. Mr Paddy Leerssen, PhD Candidate at the University of Amsterdam and Non-Resident Fellow at Stanford University’s Centre for Internet and Society; Ms Gabrielle Guillemain, Senior Legal Officer, at ARTICLE 19, London; Mr Paul Reilly, Senior Lecturer in Social Media and Digital Society, Deputy Director of Learning and Teaching, Information School, University of Sheffield; Ms Eliska Pirkova, Access Now, European Policy Analyst and Global Freedom of Expression Lead, Brussels.

8. Ofcom, “Online market failures and harms – an economic perspective on the challenges and opportunities in regulating online services”, 28 October 2019. Available at: www.ofcom.org.uk/__data/assets/pdf_file/0025/174634/online-market-failures-and-harms.pdf.

9. See Bertin Martens, “An Economic Policy Perspective on Online Platforms”, Joint Research Centre, European Commission, 2016. Available at: <https://ec.europa.eu/jrc/sites/jrcsh/files/JRC101501.pdf>.

subcommittee performed a review of existing antitrust laws, competition policies, and current enforcement levels to assess whether they are adequate to address market power and anticompetitive conduct in digital markets. The subcommittee members identified a broad set of reforms for further examination for purposes of preparing new legislative initiatives. These reforms would include areas such as addressing anticompetitive conduct in digital markets, strengthening merger and monopolization enforcement, and improving the sound administration of the antitrust laws through other reforms.¹⁰

16. In the European Union, EU Commissioners Margrethe Vestager and Thierry Breton presented in mid-December 2020 two large legislative proposals: the Digital Services Act and the Digital Markets Act.¹¹ The Digital Markets Act is aimed at harmonising existing rules in member States in order to prevent more effectively the formation of bottlenecks and the imposition of entry barriers to the digital single market.¹²

2.2. Platforms and informational power

17. The power of online platforms goes beyond a mere economic dimension. Even in the United States, where decisions taken by platforms affecting – including restricting – speech have been granted strong constitutional protection, the Supreme Court has declared that social media platforms serve as “the modern public square,” providing many people’s “principal sources for knowing current events” and exploring “human thought and knowledge.”¹³

18. Same legal and regulatory rules that apply to offline speech must in principle also be applied and enforced regarding online speech, including content distributed via online platforms. Enforcement of general content legal restrictions *vis-à-vis* online platforms – and consequent liability – constitutes a specific area of legal provisions that in Europe is mainly covered by the e-commerce Directive (in the case of the European Union¹⁴) and the standards established by the Council of Europe. The Annex to the already mentioned Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries indicates that any legislation “should clearly define the powers granted to public authorities as they relate to internet intermediaries, particularly when exercised by law-enforcement authorities” and that any action by public authorities addressed to internet intermediaries that could lead to a restriction of the right to freedom of expression must respect the three-part test deriving from Article 10 of the European Convention on Human Rights (ETS No. 5).

19. Besides this, hosting providers do generally moderate content according to their own – private – rules. Content moderation consists of a series of governance mechanisms that structure participation in a community to facilitate co-operation and prevent abuse. Platforms tend to promote the healthiness of debates and interactions to facilitate communication among users.¹⁵ Platforms adopt these decisions on the basis of a series of internal principles and standards. Examples of these moderation systems are Facebook’s Community Standards,¹⁶ Twitter’s Rules and Policies¹⁷ or YouTube’s Community Guidelines.¹⁸ In any case, it is clear that platforms have the power to shape and regulate online speech beyond national law provisions in a very powerful way. The unilateral suspension of the United States former President Donald Trump accounts on several major social media platforms has become a very clear sign of this power. The amount of

10. The final report published in 2020 is available at: <https://judiciary.house.gov/issues/issue/?IssueID=14921>.

11. <https://ec.europa.eu/digital-single-market/en/digital-services-act-package>.

12. The new rules will define objective criteria for qualifying a large online platform as a so-called “gatekeeper” and establish obligations for such gatekeepers in areas including to: allow third parties to inter-operate with the gatekeeper’s own services in certain specific situations; allow their business users to access the data that they generate in their use of the gatekeeper’s platform; provide companies advertising on their platform with the tools and information necessary for advertisers and publishers to carry out their own independent verification of their advertisements hosted by the gatekeeper; or allow their business users to promote their offer and conclude contracts with their customers outside the gatekeeper’s platform.

13. *Packingham v. North Carolina*, 137 S. Ct. 1730, 1737 (2017), as quoted in Daphne Keller, “Who Do You Sue? State and Platform Hybrid Power over Online Speech”, Hoover Working Group on National Security, Technology, and Law, Aegis Series Paper No. 1902 (29 January 2019), available at: www.lawfareblog.com/who-do-you-sue-state-and-platform-hybrid-power-over-online-speech.

14. 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.

15. James Grimmelman, “The Virtues of Moderation”, 17 *Yale J.L. & Tech* (2015). Available online at: <https://digitalcommons.law.yale.edu/yjolt/vol17/iss1/2>.

16. www.facebook.com/communitystandards/.

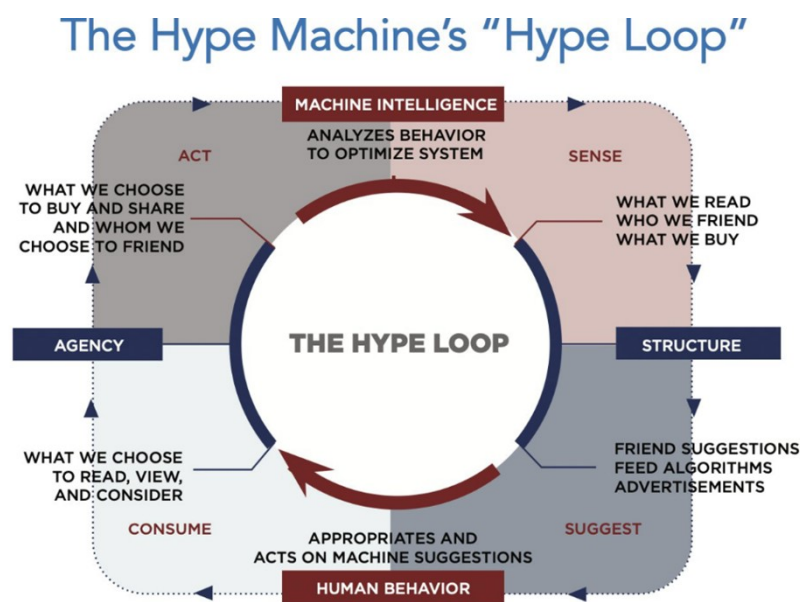
17. <https://help.twitter.com/en/rules-and-policies>.

18. www.youtube.com/howyoutubeworks/policies/community-guidelines/.

discretion that global reaching companies with billions of users have in order to set and interpret private rules governing legal speech is very high. These rules have both a local and global impact on the way facts, ideas and opinions on matters of relevant public interest are disseminated.

20. Many authors and organisations have warned that intermediaries promote content in order to maximise user engagement and addiction, behavioural targeting, and polarisation.¹⁹ On the other hand, it is also important to note that public understanding of platforms' content removal operations, even among specialised researchers, has long been limited, and this information vacuum leaves policy makers poorly equipped to respond to concerns about platforms, online speech, and democracy. Recent improvements in company disclosures may have mitigated this problem, yet a lot is still to be achieved.²⁰ This being said, it is also worth noting that big platforms already have a long record of mistaken or harmful moderation decisions in areas such as terrorist or extremist content.²¹

21. Platforms do not only set and enforce private rules regarding the content published by their users. They also engage in thorough policing activities within their own spaces as well as play a fundamental role in determining what content is *visible* online and what content – although published – remains hidden or less notorious than other. Despite the fact that users are free to directly choose content delivered via online hosting providers (access to other users' profiles and pages, search tools, embedding, etc.) platforms' own recommender systems are extremely influential inasmuch as they are in a central position among their interfaces and have become key content discovery features.²² Being true that final recommendation results are the outcome of a bilateral interaction between the users – including their preferences, bias, background, etc. – and the recommender systems themselves, it also needs to be underscored that the latter play an important gatekeeping role in terms of prioritisation, amplification or restriction of content.



Aral, Sinan. 2020. *The Hype Machine*. Currency: Penguin Random House

19. Elettra Bietti, "Free Speech is Circular", *Medium*. 1 June 2020. Available online at: <https://medium.com/berkman-klein-center/free-speech-is-circular-trump-twitter-and-the-public-interest-5277ba173db3>.

20. Daphne Keller, Paddy Leerssen, "Facts and Where to Find Them: Empirical Research on Internet Platforms and Content Moderation", in N. Persily & J. Tucker, *Social Media and Democracy: The State of the Field and Prospects for Reform*. Cambridge University Press 2020.

21. See Jillian C. York, Karen Gullo, "Offline/Online Project Highlights How the Oppression Marginalized Communities Face in the Real World Follows Them Online", *Electronic Frontier Foundation*, 6 March 2018 (www.eff.org/deeplinks/2018/03/offlineonline-project-highlights-how-oppression-marginalized-communities-face-real), Billy Perrigo, "These Tech Companies Managed to Eradicate ISIS Content. But They're Also Erasing Crucial Evidence of War Crimes", *Time*, 11 April 2020 (<https://time.com/5798001/facebook-youtube-algorithms-extremism/?xid=tcoshare>), and "When Content Moderation Hurts", *Mozilla*, 4 May 2020 (<https://foundation.mozilla.org/en/blog/when-content-moderation-hurts/>).

22. See a recent and thorough analysis on these matters in Paddy Leerssen, "The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems", *European Journal of Law and Technology*, Vol 11, No 2 (2020).

22. On the basis of the previous considerations it needs to be underscored, firstly, that no recommender system is or can be considered or pre-determined to function on a completely neutral basis. This is due not only to the influence played by users' own preferences but also by the fact that platforms' content policies are often based on a complex mix of different principles: stimulating user engagement, respecting certain public interest values – genuinely embraced by platforms or as the result of policy makers and legislators' pressures –, or adhering to a given notion of the right to freedom of expression. Probably the only case where a set of content moderation policies might deserve the qualification of neutral would be, in fact, the absence of them. However, requesting platforms to host, under no categorisation or pre-established criteria, a pile of pieces of raw content, subjected to the only requirement of respecting the law, would transform them into cesspools where spam, disinformation, pornography, pro-anorexia and other pieces of harmful information would be constantly presented to the user. This scenario is not attractive either for users or for companies. Moreover, experiences of this kind of initiatives – such as 4chan or 8chan in the United States – have shown that among other options, platforms of this nature are often used by groups promoting anti-democratic values, questioning the legitimacy of election processes and disseminating content contrary to the basic principles of human dignity – at least the way this principle is understood and protected in Europe.

23. Secondly, no matter how precisely an automated, algorithmic or even a machine learning tool is crafted, it can be arbitrarily hard to predict what it can do. Unexpected outcomes are not uncommon, and factors and glitches creating them can only be addressed once detected. Some of these outcomes might certainly present important human rights repercussions, particularly regarding the consolidation of discriminatory situations: a scientific paper on technologies for abusive language detection showed evidence of systematic racial bias, in all datasets, against tweets written in African-American English, thus creating a clear risk of disproportionate negative treatment of African-American social media users' speech.²³

24. The third remark refers to the idea of computational irreducibility in connection with any set of ethical/legal rules. To put it short, it cannot be reasonably expected that some finite set of computational principles or rules that will constrain automated content selection systems may always behave exactly according to and serving the purposes of any reasonable system of legal and/or ethical principles and rules. The computational norms will always be generating unexpected new cases and therefore new principles to handle them will need to be defined.²⁴

2.3. Disinformation and elections as particular examples

25. This report cannot cover the phenomenon of disinformation, mal-information and mis-information in all its extension.

26. The notion of disinformation, which is the most commonly used, covers speech that falls outside already illegal forms of speech (defamation, hate speech, incitement to violence) but can nonetheless be harmful.²⁵ According to the European Commission, disinformation is verifiably false or misleading information that, cumulatively, is created, presented and disseminated for economic gain or to intentionally deceive the public and that may cause public harm.²⁶ It is in any case problematic as it has direct implications on democracy, it weakens journalism and some forms of traditional media, creates big filter bubbles and echo chambers, it can be part of hybrid forms of international aggression, through the use of State-controlled media, it creates its own financial incentive, it triggers political tribalism, and it can be easily automatised. Tackling disinformation, particularly when disseminated via online platforms, requires undertaking a broad and comprehensive analysis incorporating diverse and complementary perspectives, principles and interests.

27. In addition to this, it is also important to note that the problem is not really the disinformation phenomenon or fake news itself, but the effect that the fake news creates – the manipulation of people, often political, but not necessarily so.²⁷ Therefore, rather than putting the focus on content regulation of

23. Thomas Davidson, Debasmitta Bhattacharya, Ingmar Weber, "Racial Bias in Hate Speech and Abusive Language Detection Datasets", *Proceedings of the Third Workshop on Abusive Language Online*, Association for Computational Linguistics, 2019. Available at: <https://arxiv.org/pdf/1905.12516.pdf>.

24. More details about unpredictability and irreducibility can be found in the transcript of the testimony by Stephen Wolfram "Optimizing for Engagement: Understanding the Use of Persuasive Technology on Internet Platforms" delivered before the United States Senate Subcommittee on Communications, Technology, Innovation, and the Internet on 25 June 2019. Available at: www.commerce.senate.gov/services/files/7A162A13-9F30-4F4F-89A1-91601DA485EE.

25. European Commission, "Final report of the High-Level Expert Group on Fake News and Online Disinformation", 2018. Available at: <https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation>.

26. Code of Practice on Disinformation. Available at: <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>.

disinformation, it is important to concentrate the rules on the causes and consequences of it.²⁸ As the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan, has indicated in her first report to the Human Rights Council on the threats posed by disinformation to human rights, democratic institutions and development processes,²⁹ “disinformation tends to thrive where human rights are constrained, where the public information regime is not robust and where media quality, diversity and independence is weak”.

28. Due to the direct connection with the right to freedom of expression, an excessive focus on legal, and particularly, restrictive measures could lead to undesired consequences in terms of free exchange of ideas and individual freedom. According to the Joint Declaration on freedom of expression, “fake news”, disinformation and propaganda adopted on 3 March 2017 by the UN Special Rapporteur, the Organisation for Security and Co-operation in Europe Representative on Freedom of the Media, the Organisation of American States Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights Special Rapporteur on Freedom of Expression and Access to Information,³⁰ general prohibitions on the dissemination of information based on vague and ambiguous ideas, including “false news” or “non-objective information”, are incompatible with international standards for restrictions on freedom of expression, State actors should not make, sponsor, encourage or further disseminate statements which they know or reasonably should know to be false (disinformation) or which demonstrate a reckless disregard for verifiable information (propaganda), State actors should, in accordance with their domestic and international legal obligations, and their public duties, take care to ensure that they disseminate reliable and trustworthy information, including matters of public interest, such as the economy, public health, security and the environment, and public authorities must promote a free, independent and diverse communications environment, including media diversity, ensure the presence of strong, independent and adequately resourced public service media, and take measures to promote media and digital literacy.

29. Connected to this, many experts and international organisations have particularly warned about how the so-called information disorder (in the terminology used by the Council of Europe³¹) distorted the communication ecosystem to the point where voters may be seriously encumbered in their decisions by misleading, manipulative and false information designed to influence their votes.³² Main risk factors in this context are: the lack of transparency of new forms of advertising online, which can too easily escape the restrictions applicable to advertising on traditional media, such as those intended to protect children, public morals or other social values; the fact that journalists, whose behaviour is guided by sound editorial practices and ethical obligations, are no longer the ones holding the gatekeeper role; and the growing range of disinformation including electoral information, available online, in particular when it is strategically disseminated with the intent to influence election results.

3. Considerations regarding possible solutions

30. I will present in the following paragraphs the main principles, values and human rights standards that I believe could be considered to articulate proper and adequate solutions to the different issues mentioned above.

31. Regarding the economic power of platforms, it is important to assess whether general competition law may be useful in order to partially or totally address such matters, taking also into account, as it has already been mentioned, the fact that the identification of relevant markets for the purpose of competition assessments can become particularly complex. The diversity and flexibility of business models and services offered by a wide range of types of platforms make it necessary a case-by-case analysis as well as the

27. See Paul Bernal, “Misunderstanding misinformation: why most fake news regulations is doomed by failure”, British Association of Comparative Law, 29 January 2021. Available at: <https://british-association-comparative-law.org/2021/01/29/misunderstanding-misinformation-why-most-fake-news-regulation-is-doomed-to-failure-by-paul-bernal/>.

28. See Kate Starbird, Ahmer Arif, Tom Wilson, “Disinformation as Collaborative Work: Surfacing the Participatory Nature of Strategic Information Operations”, University of Washington, 2019. Available at: <http://faculty.washington.edu/kstarbi/Disinformation-as-Collaborative-Work-Authors-Version.pdf>.

29. Available at: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G21/085/64/PDF/G2108564.pdf?OpenElement>.

30. Available online at: www.osce.org/files/f/documents/6/8/302796.pdf.

31. Claire Wardle, Hossein Derakhshan, “Information disorder: Toward an interdisciplinary framework for research and policy making”, Council of Europe, 2017. Available at: <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>.

32. See Krisztina Rozgony, “The impact of the disinformation disorder (disinformation) on elections. Report for the European Commission for Democracy through Law (Venice Commission).” CDL-LA(2018)002. 26 November 2018.

possible consideration of specifically tailored ex post solutions, such as for example data portability and platform interoperability. Interventions need to be carefully considered and defined in order to avoid unintended and harmful outcomes, particularly when it comes to incentives to entry and innovation.

32. It is important to note that a few innovative remedies to mitigate the power of platforms have already been presented by experts in the field of what can be called “algorithmic choice”. These include giving users the possibility of receiving services from third-party ranking providers of their choice (and not those of the platform they are using), which might use different methods, and emphasise different kinds of content; or giving the user the choice to apply third-party constraint providers, which would deliver content following a classification previously made by the user him/herself.³³ In a recent contribution by Francis Fukuyama and other authors, it is proposed that online platforms are forced by regulation to allow users to install what is generally called “middleware”. Middleware is software, provided by a third party and integrated into the dominant platforms, that would curate and order the content that users see. This may open the door to a diverse group of competitive firms that would allow users to tailor their online experiences. In the opinion of the authors, this would alleviate platforms’ currently enormous editorial control over content and labelling or censoring speech and enable new providers to offer and innovate in services that are currently dominated by the platforms (middleware markets).³⁴ This solution raises many interesting questions, although its proper implementation would also need to take into account important implications, particularly in terms of data protection.

33. The law may regulate filtering (by any means, human or automated, or both) by requiring platforms to filter out certain content, under certain conditions, or by prohibiting certain content to be filtered out. Such regulation would ultimately aim to influence users’ behaviour, by imposing obligations on platforms.³⁵ This kind of legislation needs to contain enough safeguards so that the intended effects do not unnecessarily and disproportionately affect human rights, including freedom of expression. In any case, any legislation imposing duties and restrictions on platforms with further impact on user’s speech must be exclusively aimed at dealing with illegal content thus avoiding broader notions such as “harmful content”. In other words, it cannot enable or establish State interventions which would otherwise be forbidden if applied to other content distribution or publication means.

34. Regarding the use of artificial intelligence and automated filters in general for content moderation, it is important to understand that nowadays, and when it comes to the most important areas of moderation, the state of the art is neither reliable nor effective. Policy makers should also recognise the role of users and communities in creating and enabling harmful content. Solutions to policy challenges such as hate speech, terrorist propaganda, and disinformation will necessarily be multifaceted and therefore mandating automated moderation by law is a bad and incomplete solution.³⁶ It is also important to acknowledge and properly articulate the role and necessary presence of human decision makers,³⁷ as well as the participation of users and communities in the establishment and assessment of content moderation policies.³⁸

33. See the already mentioned testimony of Stephen Wolfram.

34. Francis Fukuyama, Barak Richman, Ashish Goel, Roberta R. Katz, A. Douglas Melamed, Marietje Schaake, “Middleware for dominant digital platforms: a technological solution to a threat to democracy”, Freeman Spogli Institute, Stanford University, 2020. Available at: https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/cpc-middleware_ff_v2.pdf.

35. See Giovanni Sartor, Andrea Loreggia, “The impact of algorithms for online content filtering or moderation”, European Parliament’s Committee on Citizens’ Rights and Constitutional Affairs, 2020. Available at: [www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_STU\(2020\)657101](http://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_STU(2020)657101).

36. Emma Llansó, Joris van Hoboken, Jaron Harambam, “Artificial Intelligence, Content Moderation, and Freedom of Expression”, Transatlantic Working Group on Content Moderation Online and Freedom of Expression. 26 February 2020. Available at: www.ivir.nl/publicaties/download/AI-Llanso-Van-Hoboken-Feb-2020.pdf.

37. A recent study by the Council of Europe Ad hoc Committee on Artificial Intelligence (CAHAI), establishes that human oversight may be achieved through governance mechanisms such as human-in-the-loop (HITL), human-on-the-loop (HOTL), or human-in-command (HIC) approach. HITL refers to the capability for human intervention in every decision cycle of the system. HOTL refers to the capability for human intervention during the design cycle of the system and monitoring the system’s operation. HIC refers to the capability to oversee the overall activity of the artificial intelligence system (including its broader economic, societal, legal and ethical impact) and the ability to decide when and how to use the system in any particular situation. “Towards regulation of AI systems. Global perspectives on the development of a legal framework on Artificial Intelligence systems based on the Council of Europe’s standards on human rights, democracy and the rule of law” (December 2020). Available at: <https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a>.

38. See the proposals and conclusions included in the report by Ben Wagner, Joanne Kübler, Eliška Pírková, Rita Gsenger, Carolina Ferro, “Reimagining content moderation and safeguarding fundamental rights. a study on community-led platforms”, The Greens/EFA. May 2021. Available at: <https://extranet.greens-efa.eu/public/media/file/1/6979>.

35. Regarding the specific case of election processes, the European Commission for Democracy through Law (Venice Commission) has recommended a series of measures such as the revision of rules and regulations on political advertising in terms of access to the media and in terms of spending; enhancing accountability of internet intermediaries as regards transparency and access to data, promoting quality journalism; and empowering voters towards a critical evaluation of electoral communication in order to prevent exposure to false, misleading and harmful information, together with efforts on media literacy through education and advocacy.³⁹

36. Apart from legal instruments, governmental and non-governmental organisations at national, regional and international levels have been developing ethics guidelines or other soft law instruments on artificial intelligence, broadly understood. Despite the fact that there is still no consensus around the rules and principles that a system of artificial intelligence ethics may entail, most codes are based on the principles of transparency, justice, non-maleficence, responsibility, and privacy.⁴⁰ The Consultative Committee of the Council of Europe Convention for the protection of individuals with regards to the automatic processing of personal data (ETS No. 108, “Convention 108”), adopted on 25 January 2019 the “Guidelines on Artificial Intelligence and Data Protection”. They are based on the principles of the Convention and particularly lawfulness, fairness, purpose specification, proportionality of data processing, privacy-by-design and by default, responsibility and demonstration of compliance (accountability), transparency, data security and risk management.⁴¹

37. On 8 April 2020 the Committee of Ministers of the Council of Europe adopted the Recommendation CM/Rec(2020)1 to member States on the human rights impacts of algorithmic systems. It affirms, above all, that the rule of law standards that govern public and private relations, such as legality, transparency, predictability, accountability and oversight, must also be maintained in the context of algorithmic systems. It also stresses the fact that when algorithmic systems have the potential to create an adverse human rights impact, including effects on democratic processes or the rule of law, these impacts engage State obligations and private sector responsibilities with regard to human rights. It therefore establishes a series of obligations for States in the fields of legal and institutional frameworks, data management, analysis and modelling, transparency, accountability and effective remedies, precautionary measures (including human rights impact assessments), as well as research, innovation and public awareness. The recommendation also includes, beyond States’ responsibilities, an enumeration of those corresponding to private sector actors, *vis-à-vis* the same areas previously mentioned, and in particular the duty of exercising due diligence in respect of human rights.

38. The European Commission, upon reiterate requests and demands from other EU institutions, civil society and the industry, recently proposed a new set of rules regarding the use of artificial intelligence. In the opinion of the Commission, these rules will increase people’s trust in artificial intelligence, companies will gain in legal certainty, and member States will see no reason to take unilateral action that could fragment the single market.⁴² It is impossible to describe and analyse the proposal within the framework of this report. In any case it is worth noting that it essentially incorporates a risk-based approach on the basis of a horizontal and allegedly future-proof definition of artificial intelligence. Grounded in different levels of risk (unacceptable risk, high-risk, limited risk and minimal risk) the proposal articulates a series of graduated obligations, taking particularly into account possible human rights impacts. The Commission proposes that national competent market surveillance authorities supervise the new rules, while the creation of a European Artificial Intelligence Board would facilitate their implementation. The proposal also envisages the possible use of voluntary codes of conduct in some cases.

39. Entities that use automation in content moderation must be obliged to provide greater transparency about their use of these tools and the consequences they have for users’ rights.⁴³ Such rules must be capable of incentivising their full and proper compliance.⁴⁴ It is also important that before adopting any legal rule or regulation there is a proper and common understanding about the kind of information that is necessary and useful to be disclosed via transparency obligations and the public interests that legitimise such obligations. In

39. See the already mentioned report “The impact of the disinformation disorder (disinformation) on elections”.

40. On these matters, see the study by CAHAL, already mentioned.

41. Available at: <https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8>.

42. Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), released on 12 May 2021. Available at: <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence>.

43. See the recent proposals made in this field by UNESCO’s report “Letting the Sun Shine In. Transparency and Accountability in the Digital Age” (2021). Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000377231>.

44. Daphne Keller, “Some humility about transparency”, *CIS Blog*. 19 March 2021. Available at: <http://cyberlaw.stanford.edu/blog/2021/03/some-humility-about-transparency>.

addition, civil society organisations and regulators need to count on sufficient knowledge and expertise in order to properly assess the information they may request within the context of their functions and responsibilities. The proposed Digital Services Act in the context of the European Union contains remarkable provisions aimed at establishing transparency obligations vis-à-vis online platforms graduated on the basis of size. They include reporting on terms and conditions (article 12), content moderation activities and users (articles 13 and 23), statements of reasons regarding decisions to disable access or remove specific content items (article 15), online advertising (article 24), and recommender systems (article 29).

40. Transparency of social media recommendations has been classified in three main categories: user-facing disclosures, which aim to channel information towards individual users in order to empower them in relation to the content recommender system; government oversight, which appoints a public entity to monitor recommender systems for compliance with publicly-regulated standards; and partnerships with academia and civil society, which enable these stakeholders to research and critique recommender systems. In addition to these areas, experts have also advocated for a robust regime for general public access.⁴⁵

41. Additional areas vis-à-vis transparency in the design and use of artificial intelligence include, firstly, the relationship between the developer and the company: both parties need to share information and a common understanding regarding the system that is to be developed and the objectives it is going to fulfil. Secondly, no matter how precise and targeted legal and regulatory obligations may be, it might be important to count on proper independent oversight mechanisms (also seeking to avoid lengthy litigation procedures before the courts) that could check any specific transparency request in order to properly safeguard certain areas of public or legitimate private interests (for example, the protection of commercial secrets). Thirdly, independent review mechanisms must count on proper outcome assessment tools (both at quantitative and qualitative levels) in order to scrutinise the adequateness and effectiveness of the rules in place. Such tools need to particularly focus on identifying disfunctions and unintended results (false positives) and impacts on human rights.

4. Conclusions

42. Online communication has become an essential part of people's daily lives. Therefore, it is worrying that a handful of internet intermediaries are *de facto* controlling online information flows. This concentration in the hands of a few private corporations gives them huge economic and technological power, as well as the possibility to influence almost every aspect of people's private and social lives.

43. To address the dominance of a few internet intermediaries in the digital marketplace, member States should use anti-trust legislation. This may enable citizens to have greater choice when it comes to sites that protect their privacy and dignity.

44. Crucial issues for internet intermediaries and for the public are the quality and variety of information, and plurality of sources available online. Internet intermediaries are increasingly using algorithmic systems, which can be abused or used dishonestly to shape information, knowledge, the formation of individual and collective opinions and even emotions and actions. Coupled with economic and technological power of big platforms, this risk becomes particularly serious.

45. The use of artificial intelligence and automated filters for content moderation is neither reliable nor effective. Big platforms already have a long record of mistaken or harmful content decisions in areas such as terrorist or extremist content. Solutions to policy challenges such as hate speech, terrorist propaganda and disinformation are often multifactorial. It is important to acknowledge and properly articulate the role and necessary presence of human decision makers, as well as the participation of users in the establishment and assessment of content moderation policies.

46. Today there is a trend towards the regulation of social media platforms. Lawmakers should aim at reinforcing transparency and focus on companies' due processes and operations, rather than on the content itself. Moreover, legislation should deal with illegal content and avoid using broader notions such as harmful content.

45. Paddy Leerssen, "The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems", *European Journal of Law and Technology*, Vol 11 No 2 (2020). See also the report by Damian Tambini, Eleonora Maria Mazzola, "Prioritisation Uncovered. The Discoverability of Public Interest Content Online", Council of Europe, 2020. Available at: <https://rm.coe.int/publication-content-prioritisation-report/1680a07a57>.

47. If lawmakers choose to impose very heavy regulations on all internet intermediaries, including new smaller companies, this might consolidate the position of big actors which are already in the market. In such a case, new actors would have little chance of entering the market. Therefore, there is a need for a gradual approach, to accommodate different types of regulations on different types of platforms.

48. Member States must strike a balance between the freedom of private economic undertakings, with the right to develop their own commercial strategies, including the use of algorithmic systems, and the right of the public to communicate freely online, with access to a wide range of information sources.

49. Finally, internet intermediaries should assume specific responsibilities regarding users' protection against manipulation, disinformation, harassment, hate speech and any expression which infringes privacy and human dignity. The functioning of internet intermediaries and technological developments behind their operation must be guided by high ethical principles. It is from both a legal and ethical perspective that internet intermediaries must assume their responsibility to ensure a free and pluralistic flow of information online which is respectful of human rights.

50. The draft resolution which I prepared builds on these considerations.