CIVIL LIBERTIES UNION FOR EUROPE

Berlin, September 2020

The Civil Liberties Union for Europe[1] (Liberties) is a non-governmental organisation promoting the civil liberties of everyone in the European Union. Here we put forward our recommendations related to the upcoming revision of the Directive 2000/31/EC of the European Parliament and of the Council on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('eCommerce Directive').
We will refer to this new piece of legislation as Digital Services Act (DSA).

The DSA should not only be an updated version of the 20-year-old eCommerce Directive but it should also be a new concept of a data economy and the gatekeeper functions of information society service providers (ISSPs)[2]. These big platforms have changed significantly in the past 20 years. There are a variety of online platforms, information society services or intermediaries offering their services worldwide. There are different business models and some of them have grown into state-size companies with a huge impact on democratic institutions, such as public debate, the outcome of elections, freedom of expression, and privacy. **These big tech companies enjoy the ability to monitor users' activities and create profiles to sell or use for targeted advertising, while at the same time they are able to set the rules for content curation, access to information and, ultimately, freedom of expression.** These companies also limit the possibility of free competition and for newcomers to enter the market. These companies are not only similar size-wise and financially to the state, but they also impose regulation on users through terms of services and act as privatized law enforcement entities to regulate speech and avoid liability set out in laws. Of course, not all of these companies have such a substantial impact, and the legislator should differentiate between these services according to size, financial status, and market share.

When we talk about regulation, we are not only talking about state regulation, but also hybrid state-industry regulatory regimes, where civil society and individuals are involved.

[1] Currently, we have member organisations in Belgium, Bulgaria, the Czech Republic, Croatia, Estonia, France, Hungary, Ireland, Italy, Lithuania, Poland, Romania, Spain, Slovenia, the Netherlands and associated partners in Germany, France and Sweden.
[2] We will use different names for platforms referring to eCommerce Directive. We consider online platforms, services, intermediaries as ISSPs.

Instead of overly broad regulations, it is more important to focus on harm reduction, such as consumers' rights, fundamental rights, and market distortion and apply state regulation in these fields. While speech regulation differs from the above-mentioned harm-oriented regulation and therefore less regulation is needed. To some extent it is unavoidable to apply hybrid state-industry regulation and industry-regulation (self-regulation) in order to avoid overregulation.

While Europe is trying to regulate the online ecosystem, most of the big tech companies were established in the US[3]. Europe and the US have different approaches and values with regard to online free speech and data protection and privacy — the two most relevant groups of fundamental rights that are at stake when regulating the online ecosystem. While the United States is primarily focusing on free speech, Europe has a special focus on human dignity; as a consequence, it must balance conflicting rights, namely free speech and data protection. While the EU introduced strong data protection legislation with Regulation (EU) 2016/679 of the European Parliament and of the Council 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) (GDPR)[4], the US approach is to find different forms of liability to ensure the privacy for users. One of these ideas is the information fiduciary model[5] to set up new rules for those companies, a business model that is based on harvesting and profiting from users' personal data. These big tech companies have become virtually indispensable and they enjoy the ability to monitor users' activities and create profiles to sell or use for targeted advertising. These companies, according to the information fiduciary model, are similar to older fiduciaries, and are legally obliged to be trustworthy. This would ensure stricter privacy rules but without similar legislation to the GDPR. Elements of the fiduciary model can be found in models Facebook, Google, or Twitter suggest.

In the following chapters we will discuss some of the most important parts that will be covered by the Digital Services Act, such as

1) the definition of services
2) the country of origin principle
3) prohibition on general monitoring obligation - avoid mandatory filtering
4) GDPR requires human intervention in cases of automated data processing
5) Limited liability for user-generated content
6) Harmonised, transparent and rights-protective notice-and-action framework

[3] China is also starting to play an important role# in the digital ecosystem in Europe, see TikTok and WeChat.
[4] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679&from=HU&lang3=choose&lang2=choose&lang1=EN
[5] https://lawreview.law.ucdavis.edu/issues/49/4/Lecture/49-4_Balkin.pdf

7) Enforcement
8) Establish minimum requirements for meaningful robust transparency mechanism
9) Mandatory Human Rights Impact Assessment

We will discuss these chapters from the fundamental rights perspective of the users.

1. Definition of services

The definition of services, and the stakeholders of the online ecosystem, should be reconsidered in order to fit to the existing online environment. Not only are new definitions needed, but **different obligations are needed to apply to companies according to their size, number of users, and their dominance of the market.** As it is stated in the policy of the EU, the main objective of the EU competition rules is to enable the proper functioning of the Union's internal market as a key driver for the well-being of EU citizens, businesses and society as a whole. The Treaty on the Functioning of the European Union (TFEU) contains rules that aim to prevent restrictions on and distortions of competition in the internal market.[6] The prohibition of abuse of a dominant position (Article 102 TFEU) is clearly applicable, and therefore DSA and competition rules should be closely linked using the same market definition and references, where possible.

2. The country of origin principle has to be revised

The territorial scope of the DSA should include all third-country companies that provide services to EU denizens.

To effectively regulate the digital ecosystem, Liberties suggests using the solution established under Article 3 of the GDPR. The GDPR solution could serve as a legal basis for procedures against third-country services.

What we learned from the problems of GDPR enforcement and from having the same ongoing debate over the Copyright DSM Directive, is that when a regulation heavily relies on a single European country's justice system it can never be effective and will make users even more vulnerable to the big tech companies. As an example, the Irish Data Protection Authority (Data Protection Commission[7]) is the lead supervisory authority with regard to cases related to Google or Facebook in Europe and therefore heavily overburdened by complaints.

[6] https://www.europarl.europa.eu/factsheets/en/sheet/82/competition-policy
[7] https://www.dataprotection.ie/

The country of origin principle could also mean that the extent of the fundamental rights of users depends on one single member state. Even in the case of the GDPR it is critical, even though harmonized rules and networks of Authorities are protecting the personal data of users. The Data Protection Commission is reluctant to issue a decision in many cases and has a different understanding of strong data protection enforcement compared to countries where Data Protection Authorities are more active, such as Germany or France. In the scope of the Digital Services Act it will be even more problematic, **because neither common rules in the field of freedom of expression nor enforcement or EU-wide oversight exist. Therefore, derogations[8] of the country of origin principle are needed.** Without derogations, we could end up in a situation where all legal debates are decided under Irish law, causing not only extra burden on the Irish justice system but also extra burden on users to file cases anywhere other than their home countries.

Liberties supports the common position of national authorities within the CPC Network concerning the protection of consumers on social networks, which argues that platforms "cannot deprive consumers in the EU of the right to bring proceedings in the Member State of the consumer's habitual residence and the consumer may not be deprived of the protections of EU consumer law".

Accordingly, the contract concluded by a consumer with a social network operator shall be governed by the law of the country where the consumer has his habitual residence. (…) Any different choice of law should deprive the consumer of the protection afforded to him by EU Consumer Law. Choice of law clauses must be sufficiently transparent, in that they should specify unambiguously that consumers still have the possibility to invoke mandatory provisions of the laws of their own country (under Article 6 (2) Rome I[9]). Choice of law clauses which convey the incorrect impression that the contract is governed only by a distant and non-accessible jurisdiction and a foreign and unclear applicable law is unfair pursuant to Directive 93/13/EC and it is not valid under EU law. The contract cannot exclude or hinder the consumer's right to take legal action or exercise any other legal remedy (e.g. participate in a class action).

3. Prohibition on general monitoring obligation – avoid mandatory filtering

[8] The country of origin principle is not absolute. Derogations under Article 3 (3) of eCommerce Directive is possible, one of them is contractual obligations, that would support the technical procedure of accepting 'terms of conditions' of the platforms.

[9] https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32008R0593

Online platform liability has been on the table of the Commission for years now. The Directive 2010/13/EU of the European Parliament and of the Council on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive)[10], the Proposal for a Regulation of the European Parliament and of the Council on preventing the dissemination of terrorist content online Prevention of Online Terrorist Regulation[11], the Directive (EU) 2019/790 of the European Parliament and of the Council on copyright and related rights in the Digital Single Market (DSM Directive)[12], the Action Plan against Disinformation[13], the Commission Recommendation (EU) 2018/334 on measures to effectively tackle illegal content online[14] regulate the liability for online content. The EU has been struggling to find a proper way to reframe the existing liability regime set out by the eCommerce Directive 20 years ago, a new solution that would fit the new environment and also fully respect the Charter of Fundamental Rights: most importantly, freedom of expression, freedom to access information and data protection.

Liberties is of the opinion that any liability regime to be introduced should not in any way impose general monitoring obligations on service providers, and should avoid mandatory upload filters. The general prohibition on a monitoring obligation is also underpinned by the European Court of Human Rights (ECtHR). In *Payam TAMIZ v. the United Kingdom (2017)*[15] the ECtHR noted that "the Council of Europe, the European Union, the United Nations and the Organisation for Security and Co-operation in Europe have all indicated that ISSPs should not be held responsible for content emanating from third parties unless they failed to act expeditiously in removing or disabling access to it once they became aware of its illegality. Indeed, the EU Directive on Electronic Commerce expressly provides that Member States shall neither impose a general obligation on ISSPs which are storing information provided by a recipient of their services to monitor the information which they store, nor a general obligation actively to seek facts or circumstances indicating illegal activity. (para 84)."

The Court of Justice of the European Union (CJEU) had a different understanding of monitoring in the case *Glawischnig-Piesczek v. Facebook Ireland Limited*. The Court found that monitoring for identical content to that which was declared illegal, would fall within the allowance for monitoring in a "specific case" and thus not violate the Directive's general

[10] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02010L0013-20181218
[11] https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52018PC0640
[12] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0790
[13] https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=CELEX:52018JC0036
[14] https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32018H0334
[15] Payam TAMIZ against the United Kingdom (Application no. 3877/14)
https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-178106%22]}

monitoring prohibition. This allowance could also extend to equivalent content, providing the host was not required to "carry out an independent assessment of that content" and employed automated search tools for the "elements specified in the injunction."

The ruling required ISSPs to apply automated filters to remove identical and equivalent posts all over the platform.[16] It is problematic because "[t]he ruling also means that a court in one EU member state will be able to order the removal of social media posts in other countries, even if they are not considered unlawful there. This would set a dangerous precedent where the courts of one country can control what Internet users in another country can see. This could be open to abuse, particularly by regimes with weak human rights records."[17]

The *Glawischnig-Piesczek v. Facebook Ireland Limited* ruling also put false faith in the accuracy of the filters, even though false positives are widely documented through organizations such as the Lumen Database.[18] Content monitoring not only breaches Article 8 and Article 11 of the Charter of Fundamental Rights, but the automated filtering software used for this purpose is notoriously inaccurate and is likely to catch lawful content that does not breach any law and may in fact be essential for societal and political debate.

It is crucial to maintain the prohibition on general monitoring obligations. General monitoring would undermine freedom of expression and data protection by imposing ongoing and indiscriminate control of all online content with mandatory use of technical filtering tools. The no-monitoring principle protects free expression and can be maintained while creating oversight and accountability for the use of automated tools in online content moderation.

4. GDPR requires human intervention in cases of automated data processing

Automated filters are widely used to eliminate liability for users' content and also to ensure that no illegal content, such as child sexual abuse content, is available on platforms and other services. Automation is necessary for handling a vast amount of content shared by users, however the consequences are far-reaching. Automated decision-making tools, such as filtering techniques, are contextually blind, and they are therefore unable to assess the context of expressions accurately and differentiate between illegal and legal content.

[16] For further analysis see:
https://globalfreedomofexpression.columbia.edu/cases/glawischnig-piesczek-v-facebook-ireland-limited/
[17] Article 19 statement of the ruling:
https://www.article19.org/resources/cjeu-judgment-in-facebook-ireland-case-is-threat-to-online-free-speech/
[18] https://www.lumendatabase.org/topics/1

Any algorithm-curated content moderation will automatically link to personal data processing. Under Article 22 of the GDPR, users have the right not to be subject to a decision based solely on automated processing which produces legal effects concerning him/her or similarly significantly affects him/her unless it is based on i) a contractual relationship; ii) authorized by law; iii) or it is based on the users' explicit consent. Number i) and ii) are not applicable. For i), accepting terms of services are not considered contractual relationships. **Therefore, data processing in relation to the automated decision-making process can only rely on users' explicit consent under Article 4 (11) of the GDPR. The right of the users to contest an automated decision entitles them *not* to give consent to any kind of automated filtering method without human intervention. Users must be able to understand decisions made about them as well as understand how automated decision-making affects them, and they must also understand how to contest a decision if necessary according to Article 21 (1) of the GDPR. Human intervention is also essential for transparent decision making and transparent appeal mechanisms to balance the imbalance between ISSPs and users. There cannot be an effective remedy without human intervention.**

It is important to avoid considering DSA as authorization of the law. Those who participated in the debate about the DSM Directive know that this problem was not only hotly debated throughout the legislation process, but because of the vague wording of the DSM Directive, it was thoroughly discussed during the stakeholder dialogue as well. There are a few things to be learned from the copyright debate, but one of the most important is that the rules. should be as clear as possible. This is necessary to avoid leaving legislation to the discretion of the stakeholder dialogue involving big US tech firms lobby without proper transparency, and the outcome of the process should not lead to new laws because of the lack of legislative power.

5. Limited liability for user-generated content must be preserved

The freedom to speak and discuss political issues freely is the basic requirement of any democracy. If a strict liability regime is introduced for user-generated content, it would hamper the ability to speak freely. The reason for that is that business-oriented service providers will act against any content that would trigger their liability. To avoid liability, these companies will do anything, including pre-filtering, monitoring or banning users, in order to avoid any liability. And this means they will remove or filter out anything that has the slightest chance of infringing the law.

The limited liability regime has been further interpreted by national and international courts. Safe harbour regimes ensure that intermediaries are immune from liability unless they are aware of the illegality and are not acting adequately to stop it. The differentiation of 'passive' and 'active' roles was elaborated by the Court of Justice of the European Union (CJEU) in two important cases. In C-236/08 to C-238/08 Google France and Google Inc. et al. v Louis Vuitton Malletier et al, the court examined whether the role played by that service provider's conduct was merely technical, automatic and passive, pointing to a lack of knowledge or control of the data which it stores. It implies that that service provider "has neither knowledge of nor control over the information which is transmitted or stored". In the other significant case, C-324/09 L'Oréal SA v eBay L'Oréal v eBay,the activity of the service provider was again at the center of the debate. The active nature of eBay was established by the fact that it exercised control over the data. However, solemnly relying on these court cases would be almost certainly erroneous for the mere fact that both of these decisions are more than 10 years old now and technology, the online economy and the role of the intermediaries have significantly changed in the last decade.

Limited liability requirements have been elaborated by courts, however there are discrepancies between decisions. Therefore, we suggest that the Commission takes a stand for users' rights and upholds and further strengthens limited liability for content shared on online platforms, as it was addressed by the opinion of Advocate General Saugmandsgaard Øe in Joined Cases[19] that online platform operators, such as YouTube and Uploaded, are not directly liable for the illegal uploading of protected works by the users of those platforms. The Advocate General specified that the situations in which the service provider has 'actual knowledge of illegal activity or information' or is 'aware of facts or circumstances from which the illegal activity or information is apparent' refer, in principle, to specific illegal information.

The notification system was intended to offer verification to the illegal nature of information. But such information should only be removed when the illegal nature is manifest. There are only a limited number of cases where the illegal nature is obvious. In most cases the content requires thorough legal, contextual, and factual examination. In the context of copyright, Advocate General Saugmandsgaard Øe stated that "the risk is that in all these ambiguous situations the provider tends towards systematically removing the information on its servers in order to avoid any risk of liability vis-à-vis the rightholders. It will often find it easier to remove information rather than having to claim itself in the context of a possible action for liability that an exception

[19] C-682/18 Frank Peterson v Google LLC, YouTube LLC, YouTube Inc., Google Germany GmbH and C-683/18 Elsevier Inc. v Cyando AGon

applies. Such 'over-removal' would pose an obvious problem in terms of freedom of expression. … I would add that, where the application of an exception is not automatically precluded, the notification must contain reasonable explanations why it should be. In my view, only this interpretation can avert the risk of intermediary providers becoming judges of online legality and the risk of 'over-removal'".[20]

Liberties would like to call the attention of the Commission to how liability questions have changed in the last five years. While the first ECtHR case (Delfi v. Estonia 2015[21]) imposed liability for third-party content, later the arguments changed to a more balanced and more free speech-friendly direction.

In *Delfi v. Estonia (2015)*, the European Court of Human Rights (ECtHR) ruled that Delfi, one of the major Estonian news portals, is liable for defamatory comments posted online by its readers. In its decision, the Court took into account the content of the comments, the fact that Delfi is a professional news portal run on a commercial basis, the insufficient measures taken by Delfi to prevent harm being caused to third persons (automatic deletion of certain vulgar words and notice-and-take-down system), and the moderate sanctions imposed on it. According to the Court, the restriction on Delfi's right to freedom of expression was proportionate and holding it liable for comments written by third parties was "necessary in a democratic society".

The next important similar case of the ECtHR was the MTE and Index v. Hungary (2016)[22] case, when the court ruled that the self-regulatory body of the Hungarian Internet Content Providers MTE (and newspaper Index) are not liable for the offensive online comments of their readers. The Court considered that the domestic courts had not properly balanced the rights involved, namely MTE's right to freedom of expression and the plaintiff's right to respect for its commercial reputation. The Court considered four criteria in assessing the proportionality of the interference in a situation that does not involve hate speech or a call to violence: 1) the context and content of the comments; 2) the liability of the authors of the comments; 3) the steps taken by MTE and the conduct of the injured party; 4) and the

[20] http://curia.europa.eu/juris/document/document.jsf?text=&docid=228712&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=18466818

[21] Delfi v. Estonia 2015 (Application no. 64569/09) https://hudoc.echr.coe.int/eng#{%22fulltext%22:[%22delfi%20grand%20chamber%22],%22documentcollectionid2%22:[%22GRANDCHAMBER%22,%22CHAMBER%22],%22itemid%22:[%22001-155105%22]}

[22] Magyar Tartalomszolgáltató Egyesülete and Index.hu Zrt. v. Hungary (Application no. 22947/13) https://hudoc.echr.coe.int/eng#{%22fulltext%22:[%22MTE%20and%20Index%20v.%20Hungary%22],%22documentcollectionid2%22:[%22GRANDCHAMBER%22,%22CHAMBER%22],%22itemid%22:[%22001-160314%22]}

consequences of the comments to the company. In the end, the Court ruled that there had been a violation of Article 10 of the European Convention on Human Rights.

In Pihl v. Sweden (2017)[23], the ECtHR ruled "the Court has previously found that liability for third-party comments may have negative consequences on the comment-related environment of an internet portal and thus a chilling effect on freedom of expression via internet. This effect could be particularly detrimental for a non-commercial website" (para 35).

In *Tamiz v. UK (2017)*, the ECtHR ruled that Google's blog-publishing service is not liable for offensive comments. The ECtHR shared the opinion of the third-party interveners, that although millions of internet users post offensive or defamatory comments online every day, the majority of these comments are likely to be too trivial in nature or their publication too limited to cause any significant damage to another person's reputation (para 80).

In the case of *Magyar Jeti Zrt v. Hungary (2018)*[24], the ECtHR clarified the liability for defamatory content hyperlinked in reports. The Court referred to the very purpose of hyperlinks and added that it cannot accept a strict or objective liability for media platforms embedding, in their editorial content, a hyperlink to defamatory or other illegal content. It found that objective liability, such as applied in the case at issue, "may have foreseeable negative consequences on the flow of information on the Internet, impelling article authors and publishers to refrain altogether from hyperlinking to material over whose changeable content they have no control. It may have, directly or indirectly, a chilling effect on freedom of expression on the Internet." "The ECtHR however did not exclude that, 'in certain particular constellations of elements', the posting of a hyperlink may potentially engage the question of liability, for instance where a journalist does not act in good faith in accordance with the ethics of journalism and with the diligence expected in responsible journalism."[25]

Platform liability and reframing the legal consequences for unlawful content and also for inadequate systems and processes could be designed in different ways. This could serve as a systemic change to the regulation that is more likely to suit the

[23] Rolf Anders Daniel PIHL v. Sweden Application no 74743/14
https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-172145%22]}
[24] Magyar Jeti Zrt v. Hungary (2018) (Application no. 11257/16)
https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-187930%22]}
[25] Carl Vander Maelen, Magyar Jeti Zrt v. Hungary: the Court provides legal certainty for journalists that use hyperlinks, 2019.
https://strasbourgobservers.com/2019/01/18/magyar-jeti-zrt-v-hungary-the-court-provides-legal-certainty-for-journalists-that-use-hyperlinks/

protection of fundamental rights. It is easier and more effective than focusing on content, especially because of the global nature of platforms in light of national liability regimes. Transparency requirements, risk management, and impact assessments of algorithms of content curation and moderation should be revised. We discuss these requirements in the coming chapters.

6. Harmonised, transparent and rights-protective notice-and-action framework

A harmonized notice-and-action framework should enable users to exercise their fundamental rights. A **proper notice-and-action system is compliant with existing regulations and gives users due process. It also allows users to flag potentially illegal content and sets clear and predictable requirements for ISSPs to have processes in place to deal responsibly with such notifications with due regard for users' freedom of expression.**

Any notice-and-action mechanism should protect freedom of expression and data protection by applying transparent and fair due process for intermediaries when they take content moderation decisions. The actual knowledge interpretation is also crucial; the fact that someone issued a notice does not automatically mean that the host has actual knowledge of and is liable for the content. Liability exemption should be sustained and applied according to recent Court rulings. (See cases above.)

Data protection is an important element, and therefore the personal data of users should be handled in accordance with the GDPR, and a proper redress mechanism and transparency requirements are needed.

Liberties suggests the following safeguards to establish a proper notice-and-action mechanism:

- Notifiers should be required to make reasonable efforts to contact the user directly.
- Notifiers are required to give information about the content in question, such as location, the reason for notice or a presumption of illegality.
- Service providers and platforms should sustain an easy-to-access, easy-to-understand notification system for users. The notification system should differ according to the different complaints.
- Counter-notices should be simplified and must be submitted with a time limit. The counter-notice could ensure that content stays online with very limited exceptions, such as child sexual abuse content or hate speech.
- A declaration of good-faith must exist.

- Content-decision must be proportionate and well elaborated. The content decision should content the possible redress mechanisms available for the user.
- Simple access to legal redress in court is important.

Notice-and-action procedure is under elaboration by the copyright stakeholder dialogue, which could be different because of the nature of the notice-and-action and the harm that could be caused. The relationship between the DSA and DSM should be clarified. Liberties believes that the DSA will serve as lex generalis and the DSM Directive serves as lex specialis.

7. Enforcement should be the core of regulating online services and platforms

Enforcement is crucial in this field, and this depends on both national and European regulatory authorities and consumer organizations, with the involvement of self- and co-regulatory bodies. The extended work to be put on authorities, such as auditing, dispute resolution and imposing fines, will require extra funding. In the case of the GDPR, one of the factors that hampers effective enforcement is that the budgets of national Data Protection Authorities have not been adjusted to the new requirements.

- Liberties suggests an **extra budget** for authorities and consumer organizations for the increase in work.
- It is also worth considering that, **similarly to the EDPB, a European Media Authority Board should be established that could oversee the national work of the media authorities** and effectively enforce media pluralism and freedom of speech throughout internet regulation and platform economy.
- Independent dispute resolution bodies, both self- and co-regaulation, could be involved to support the effective dispute resolution. However, **everyone should have access to an independent judiciary in all cases.** We call attention to the initiations of the big media platforms, such as Twitter's Social Media Council idea or the Facebook Oversight Board. Both initiatives are trying to offer solutions to the pressing need for proper regulation, while also serving as solutions to adopting the fiduciary model and to eliminating the pressure of governments.[26]
- Besides the new content moderation ideas, "radical transparency" is key for both online platforms and governments. As it is stated in the report of David Kaye, Special Rapporteur of the UN, "transparency includes knowing what rules States and companies use to moderate content, the rules regarding content, how those rules

[26]A survey (2018) by Freedom House , a democracy and rights watchdog organization found 65 percent of the countries it reviewed asked online platforms to restrict content of political, social or religious nature.

are applied, what kind of appeals process exists and what kind of accountability there is for wrongful take down of content."[27]

8. The DSA should establish minimum requirements for meaningful and robust transparency mechanisms

Transparency is a precondition for gathering evidence about the implementation and the impact of existing laws. It enables legislators and judiciaries to understand the regulatory field better and to learn from past mistakes. Only through the combination of comprehensive transparency reports by states, regulators, ISSPs and the civil society will we be able to draw a realistic picture of how online content moderation works.

In the following transparency chapters, we rely on the assessment of Ranking Digital Rights [28]. We recommend that the Commission studies the methodology elaborated by RDR.

8/1 Transparency requirements for advertising

"Companies that enable any type of advertising on their services or platforms should clearly disclose the rules for what types of ad content is prohibited—for example, ads that discriminate against individuals or groups based on personal attributes like age, religion, gender, and ethnicity. **Companies should be transparent about these rules so that both users and advertisers can understand what types of ad content are not permissible and so they can be accountable for the ad content that appears on their services or platforms.**"[29]

Companies should publish data at least once a year, in a structured data file that is available in an easy-to-access, easy-to- understand way, with regular updates including:

- The rules applied to content.
- Policies about the types of advertising content that are prohibited on a platform or service, and its processes for enforcing these rules.
- Data on the total number of advertisements removed as a result of breaches to advertisement content policies, and they should also break out this data by what rule was violated.

27 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

28 Ranking Digital Rights produce Corporate Accountability Index, which evaluates how transparent digital platforms and telecommunications companies are about policies and practices affecting freedom of expression and privacy, based on international human rights standards.

29 2020 RDR Index, Indicator F1b: https://rankingdigitalrights.org/2020-indicators/#F1b

- Targeted advertising policies and the advertisements removed for violating targeting rules, and what rule was violated.
- Information about the rights of the users for their personal data, their rights under the GDPR and a step-by-step description of how to exercise rights ensured by the law.
- Access to dispute resolution and to human intervention.

8/2 Transparency requirements for targeted advertising

The ability for advertisers to target users with tailored content—based on their personal data—can be a breach of the GDPR, especially in cases where that data is used and transferred to or by third parties. Cambridge Analytica showed clearly how targeted advertising could be misused, how it creates filter-bubbles to amplify disinformation[30], and how it can also be overtly discriminatory.

Companies that enable advertisers and other third parties to target their users should publish data at least once a year in a structured data file in an easy-to-access, easy-to-understand way, with regular updates including:

- The outcome of regular, comprehensive, and credible due diligence, such as through robust human rights impact assessments, to identify how all aspects of its targeted advertising policies and practices affect users' fundamental rights to freedom of expression and information, to privacy, and to non-discrimination, and to mitigate any risks posed by those impacts.
- Targeting policies and their changes, what are the parameters, and what is not permitted in what type of advertisement.
- Evidence of enforcement of ad targeting rules by publishing data on the total number of ads it removes as a result of breaches to ad content policies, and they should also break out this data by what rule was violated. Companies should also provide evidence that it is enforcing its ad targeting policies by publishing data on the number of ads removed for violating targeting rules, and which rules were violated.
- **Clearly disclose that targeted advertising is switched off by default.**
- Information about the rights of the users for their personal data, their rights under the GDPR and a step-by-step description of how to exercise rights ensured by the law.
- Access to dispute resolution and to human intervention.

[30] Access Now, Civil Liberties Union for Europe, European Digital Rights, Informing the disinformation debate, https://dq4n3btxmr8c9.cloudfront.net/files/2r7-0S/online_disinformation.pdf, 2018.

8/3 Transparency about policing content (terms of service)

ISSPs should set clear rules in cases where they prohibit certain content or activities. These limitations must be transparent, especially in case of legal action, to ensure data protection, freedom of expression, freedom of information and to help dispute resolution and effective remedy.[31]

Companies should publish data at least once a year in a structured data file in an easy-to-access, easy-to-understand way, with regular updates including:

- Disclosure of comprehensive and credible due diligence.
- How their policies affect users' fundamental rights to freedom of expression and information, to privacy, and to non-discrimination, and to mitigate any risks posed by rules.
- How they enforce these rules.
- Data on the total number of contents it removes as a result of breaches to ad content policies, and by what rule was violated.
- Information about the right of the users for their personal data, their rights under the GDPR and a step-by-step description of how to exercise rights ensured by the law.
- Access to dispute resolution and to human intervention.

8/4 Transparency about algorithmic system development and use

Algorithmic systems can have adverse effects on fundamental rights such as freedom of expression, access to information, data protection, and non-discrimination. **Algorithmic content curation, recommendation, and ranking systems play a critical role in shaping what types of content and information users can access online. It can alter the information ecosystems and influence political decisions and the outcome of the elections. These systems can also be used to spread misinformation.** [32]

Companies that develop and deploy algorithms should publish data at least once a year in a structured data file in an easy-to-access, easy-to-understand way, with regular updates including:

31 With reference to RDR Index indicators G4b, F3a, F4a, and G6b:. https://rankingdigitalrights.org/2020-indicators/

32 Indicator F12, RDR Index: https://rankingdigitalrights.org/2020-indicators/#F12

- Disclosure of comprehensive, and credible due diligence.
- The features used by these algorithms to optimize content.
- How these systems affect the fundamental rights of users and how risks are mitigated.
- Clear and accessible policy stating the nature and functions of these systems. This policy should be easy to find, presented in plain language, and contain options for users to manage settings.
- Algorithmic systems development policies describing the development and testing of algorithmic systems in a way that users can access, read and understand, so that users can make informed decisions about whether to use a company's products and services
- Disclosure of the data used for developing algorithmic systems.
- Publication of terms of service.
- Information about the rights of the users for their personal data, their rights under the GDPR and a step-by-step description of how to exercise rights ensured by the law.
- Access to dispute resolution and to human intervention.

8/5 Transparency about government demands and the responses

Digital services and platforms frequently receive demands from governments to remove content, accounts, or disclose user information, even access to real-time user communications.
Companies should publish data at least once a year in a structured data file in an easy-to-access, easy-to- understand way, with regular updates:

- Companies must disclose their relationships with governments.
- They should fully disclose their processes for responding to government demands to restrict or block content, or to access user information.
- They should also report data on the number and types of these requests they receive--and from which authorities--and comply with. Transparency reports, similarly to Google's, set a good example.

8/6 Special rules for universal advertising transparency by default[33]

The political campaigning landscape has changed significantly with the digitalisation of our public sphere, which has created new opportunities for political participation, but also

[33] 7/6 Is a joint statement of several non-governmental organizations initiated European Partnership for Democracy.

poses significant risks to the integrity of elections and the political debate. Advertisers can purchase exorbitant amounts of ads and flood people's social media feeds, thereby buying themselves space in public policy and political debates.

At the source of these problems lies the lack of transparency offered by ISSPs. While some platforms have found ways to provide some transparency on political ads (partly due to pressure by the European Commission), their voluntary measures fall short of providing meaningful transparency. One crucial weakness of the status quo is that it leaves platforms to decide what is and is not political advertising - and thus, what advertising will and will not be subjected to platforms' transparency regimes. To avoid this issue and to recognise the kind of behavioural targeting and algorithmic delivery that underlies all types of social media advertising, it is necessary to require meaningful default transparency for all ads.

Why transparency by default for all ads?

To allow for public interest scrutiny: Transparency is necessary, first and foremost, to allow for public scrutiny of advertising. As many studies on the implementation of the EU Code of Practice against Disinformation have shown, false negatives and false positives were rife in the political ad libraries of the signatories of the code: non-political advertisements were erroneously included in the libraries, while many political ads were excluded.[34] The lack of a comprehensive repository of all ads made it impossible to verify whether all political ads were included in the libraries, and the political ad libraries and labelling missed a lot of sponsored content. **In a situation where it is difficult to police the labelling of political ads, it is ultimately necessary to ensure the transparency of all ads.**

To overcome diverging definitions of political ads: EU member states have diverging definitions of political advertising, and some have no definition at all. The EU Code of Practice on Disinformation distinguishes between political and issue-based advertising, which introduces a distinction that is not reflected uniformly across member states'

[34] See: Márcio Silva, Lucas Santos de Oliveira, Athanasios Andreou, Pedro Olmo Vaz de Melo, Oana Goga, Fabrício Benevenuto, (2020): Facebook Ads Monitor: An Independent Auditing System for Political Ads on Facebook. Cornell University. Available here.
Privacy International (2019): Social media companies are failing to provide adequate advertising transparency to users globally. Available here.
European Partnership for Democracy (2020): Virtual Insanity: The need to guarantee transparency in digital political advertising. Available here.
European Regulators Group for Audiovisual Media Services (2020): Report of the activities carried out to assist the European Commission in the intermediate monitoring of the Code of practice on disinformation (ERGA Report). Available here.

electoral laws. Introducing mandatory transparency of all advertising helps to address the difficulty of adopting and applying one common definition of political advertising.[35]

To verify the labelling and disclaimers of political ads: Full, meaningful transparency is the only way to verify if political content is labelled and regulated as such, and it allows civil society and other watchdogs to monitor the grey zone between political and commercial ads. Past experience with the Code of Practice has shown that the platforms often incorrectly categorise and label political ads.[36] As the forms of political advertising online will undoubtedly evolve as the technology changes, full transparency creates enough flexibility to account for such changes. Moreover, "political" ads are not the only ones that should be subject to scrutiny and accountability: false or deceptive advertising, hoaxes, and paid disinformation (notably pertaining to public health, in the current pandemic context) should also be subject to scrutiny.

To better understand malign actors: In addition, full transparency of paid-for content will allow for better identification and a deeper understanding of other malign actors' strategies. Currently, it is very easy for malign actors to get into the political campaigning eco-system and hijack the political debate to their own ends in ways that are not possible on television or through other advertising channels. Ads can be used to lure people into Facebook groups that are not initially about a political issue, but eventually become focused on a political cause and are used for malign purposes. Investigations show the platforms' inability to enforce their own policies in this regard. Researchers, civil society and journalists need access to an archive to understand the marketing techniques, networks and origin of these actors.

To protect consumers and strengthen businesses: For commercial advertising, transparency by default benefits both brands and users. Universal ad transparency will help combat discriminatory and potentially illegal advertising practices, and help ensure compliance with privacy and data protection laws as they apply to ad targeting. At the same time, transparency also helps protect consumers - particularly those from vulnerable groups - from advertising for illegal and harmful products, and potentially increases trust in brands and in the platforms.[37] Businesses that act in good faith and comply with regulation (including the GDPR) also benefit, as transparency levels the playing field by

[35] ERGA's noted that the definitions of political ads and issue-based ads adopted by the platforms are inconsistent with the requirements set out in EU Member State laws, where they exist, although they do not exist in many countries. See ERGA, Report of the activities carried out to assist the European Commission in the intermediate monitoring of the Code of practice on disinformation, June 2019, p. 15, available here.
[36] Ibid.
[37] See Paddy Leerssen in Panoptykon (2020): Who really targets you? Facebook in Polish election campaigns. Available here.

preventing bad-faith advertisers from breaking the law with impunity. Transparency on the advertiser, engagement and targeting criteria is only one part of a company's advertising strategy and therefore does not imply a disclosure of trade secrets. Transparency mechanisms would have to be built into the platforms in an easy-to-use format so that it doesn't prove a burden for advertisers.

Such public-facing transparency is a necessary yet in itself insufficient first step for enhancing the accountability of platforms and advertisers. While the measures described below will enhance the transparency of advertisers, this needs to go alongside transparency of the ad optimisation processes on the part of the platforms, as well as user-level transparency explaining why exactly an ad is reaching them individually.[38] Transparency in and of itself is only instrumental to accountability and needs to therefore be backed up with further action to safeguard rights and democratic processes online. For example, transparency may reveal widespread harmful practices that may in fact be prohibited but have escaped meaningful enforcement, or even novel practices that should be regulated.

What universal transparency by default looks like

Mandatory, functional ad libraries: The European Commission should foster the development of and issue minimum technical standards for advertisement libraries for digital platforms, covering both the design and functioning of ad libraries. These minimum technical standards should be developed through a multistakeholder process, and help overcome the numerous problems and bugs that render the existing ad libraries meaningless as transparency tools. The ad repositories should comply with well-defined accessibility and technical standards standards. We suggest the following starting point for these standards[39]:

- Libraries provided by each platform should be compatible with each other. There should be a unique set of standards and protocols provided by the Commission that all platforms are required to use.[40]
- Platforms should assign unique identifiers to each advertisement and advertiser to allow for trend analysis over time and across platforms. Advertisers should keep the same unique identifier no matter what platform they're using.

[38] For further information about these ideas, contact the Panoptykon Foundation.
[39] These recommendations are based on the guidelines for effective ad archives issued by Mozilla and a cohort of independent researchers. Available here.
[40] For an example, see this universal transparency schema Google has created here.

- All images, videos, and other content should be provided in a machine-readable format accessible via an application programmatic interface. This should include any words used in images or in audio provided as searchable text.
- The ability to download a week's worth of data in less than 12 hours and a day's worth of data in less than two hours.
- Bulk downloading functionality of all relevant content. It should be feasible to download all historical data within one week.
- Search functionality by the text of the content itself, by the content author or by date range.
- Up-to-date and historical data access, including the availability of advertisements within 24 hours of publication; the availability of advertisements going back 10 years. In addition, APIs should be promptly fixed when they are broken and APIs should be designed so that they either support or at least do not impede long-term studies
- The API itself and any data collected from the API should be accessible to and shareable with the general public.
- The ad libraries must be **free of charge** and shared under a **permissive open source licence.**
- The ad libraries should include clear **audit trails** for content which has been removed, including the reasoning for its removal while maintaining data on the advertiser, funder, spend, and targeting.

Such ad libraries should become **mandatory** for platforms from a set number of users onwards, to be decided by a European-level regulator or coordination platform between national regulators, and reviewed on a yearly basis.[41]

Such public advertising libraries, which again should include commercial advertising as well as "political advertising" (however defined), must **disclose the following information at minimum**:

- Exact spend: broad spend ranges like 0-100 EUR, 100-1,000 EUR, … are not meaningful information for users and researchers. Enhanced transparency on all aspects of online targeting - including the amount spent - is a necessary price to pay for the increased customer access advertisers gain with behavioural targeting practices.
- Advertiser information: this needs to be accurate and complete. Third parties, like advertising agencies, who advertise on behalf of another entity need to be as transparent as the brand or entity that commissions the advertisement. Information

[41] For an example, see this definition by European Digital Rights of dominant platforms (p.16) here.

on both the third party and the political candidate or party needs to be detailed in the ad library and the disclaimer. Information on the funding entity should also be disclosed and verified.

- Advertiser identification: Platforms should facilitate the linkage with other databases that support verification by displaying official identification such as corporate registers, advertisers' tax ID, political candidates' electoral court declaration, or any other identification number that facilitates enforcement and verification of the advertisers' identity.
- Targeting mechanism: use of lookalike audiences, and which audiences they chose; use of profiling based on imported datasets and the source of this data (such as a newsletter platform, for instance); or other similar mechanisms to improve targeting by the platforms.
- Targeting and delivery criteria, with the same level of granularity as the advertiser can choose from. This must include the optimisation goal selected by the advertiser and general information on the optimisation logic used by the platform (possibly in another layer/interface that is accessible from the ad library).
- Audience reached
- Engagement and reach in absolute and relative terms, e.g. likes, shares, comments
- The ad library should include this information for all ads, including the ads taken down by the platforms because they did not adhere to community standards. With the exception of content judged illegal by the relevant state authorities, banned ads should remain in the ad repository for public scrutiny. For ads that were taken down, information about the kind and category of content, and the reasons and process for take-down should be displayed.
- None of these measures should reveal personal user information and all of them should be GDPR compliant.

Real-time transparency disclosures for individual users: There should be clear, consistently applied on-screen designations of ads, distinguishing them from other content. Users should have easy access to easily comprehensible basic transparency information, as well as easy access to the above-mentioned transparency information. Furthermore, access to a personal ad library showing users who is targeting them and how, would allow users to better hold platforms to account. The design of this access and the information presented should be at least as good as the rest of the platform or service. Companies should show evidence of the design process and provide information on user interactions with it on request.

Verification of advertisers: Platforms and political advertisers need to be held to account for verifying all advertisers' real identity, who's paying (indirectly) for the ads, contact details and for political advertisers a reference to their declaration to the electoral authorities (when applicable in the country context). Such verification needs to be quick and mandatory. It should not rely on self-declaration by the advertiser but require the platforms to verify the information provided. It also needs to be more closely monitored by authorities, to ensure platforms perform better than they did as part of their efforts for the Code of Practice, with appropriate sanctions available for advertisers and platforms that do not stick to the rules.

Anonymity where needed to protect safety: We encourage the European Commission to issue guidelines for platforms to protect advertisers in high risk contexts. The European Commission could make suggestions for a mechanism for advertisers to anonymise their identity on the basis of political threats and risk, for public interest actors such as human rights defenders and activists. Such an anonymity mechanism would for instance protect organisations raising awareness on LGBTQ+ rights in countries where those rights cannot be taken for granted. This mechanism, meant to protect those in need of anonymity, could be abused as a loophole by advertisers trying to hide their identity, even though they do not have to fear prosecution. Therefore, consideration should be given to independent mechanisms to oversee the granting of anonymity. Exemption applications should be carefully scrutinised according to a transparent set of criteria and information should be made publicly available on the number of exemptions requested and granted on an annual/quarterly basis.

Binding requirement and enforcement: The Commission should develop a mechanism for ensuring universal transparency for online advertising meets the standards set out above. This should include relevant penalties for non-compliance up to and including preventing a platform from running any ads if their efforts in this area are deemed insufficient.

8. Mandatory Human Rights Impact Assessments (HRIA)

HRIAs have been used more systematically in the nordic countries[42] , but it is **important to introduce such requirements, similar to the Data Protection Impact Assessment set out in the GDPR, more widely in order to analyze the effect of business activities on users' fundamental rights.** This should be a fundamental rights-based approach that

[42] The Danish Institute for Human Rights has developed a related Human Rights Compliance Assessment tool https://www.humanrights.dk/business/tools/human-rights-impact-assessment-guidance-toolbox

integrates human rights principles such as freedom of expression, data protection, privacy and non-discrimination.

Either consumer organizations or other responsible authorities would be able to require such an assessment from tech firms over a certain size, number of users, or dominance of the market etc.