



House of Lords Select Committee on Communications – The Internet: To Regulate or Not To Regulate? Inquiry

ARTICLE 19 written evidence

1. Is there a need to introduce specific regulation for the Internet? Is it desirable or possible?

1. ARTICLE 19: Global Campaign for Free Expression (ARTICLE 19) is an independent human rights organisation that works around the world to protect and promote the rights to freedom of expression and freedom of information. ARTICLE 19 has significant experience working on intermediary liability issues. We intervened in *Delfi v Estonia* before the Grand Chamber of the European Court of Human Rights¹ and have responded to numerous EU consultations on this issue.² We have also dealt with intermediary liability and related online content regulation issues in a range of countries, from Brazil³ or Tanzania⁴ to France⁵ and Germany⁶.
2. In ARTICLE 19's view, it is unnecessary to introduce new or specific regulation of the Internet in the sense of online content regulation. Though the current legal framework in this area could be further improved to better protect freedom of expression, we believe that rolling back immunity from liability for social media platforms (and introducing further regulation) would only diminish freedom of expression. To the extent that Internet regulation is thought necessary or desirable, however, ARTICLE 19 believes that its focus should be on strengthening data protection law, online political advertising during elections and competition matters rather than restricting content *per se*.

Social media platforms are already subject to a range of laws and regulations

3. At the outset, we note that the 'Internet' is far from unregulated. Indeed, a great many laws already govern various *activities* on the Internet, from e-commerce to cybersecurity, cybercrime or data collection and retention. In our experience, many of those who call for 'internet regulation' do not take account of this. Instead, what they appear to refer to is the more specific idea of 'online *content* regulation'. Indeed, most of these calls seem to concern proposals for regulating 'social media platforms', particularly in relation to 'fake news', 'extremism' or hate speech. Our submission focuses on these latter aspects.
4. As noted by this Committee in the call for evidence, some degree of online content regulation already exists in the form of the E-Commerce Regulations 2002 (ECRs), which transposed the E-Commerce Directive 2000 ('ECD') into English law.⁷ The original purpose of the Directive was to provide a balance between (i) providing a suitable

¹ ARTICLE 19's intervention is available from [here](#).

² See e.g. [ARTICLE 19's analysis of the EU Code of Conduct on Combatting Illegal Hate Speech](#), 2016.

³ See e.g. ARTICLE 19's country report on Brazil and the Marco Civil DA Internet, 2015 available from [here](#).

⁴ See our [analysis](#) of the Tanzania Electronic and Postal Communications (Online Content) Regulations 2018

⁵ See ARTICLE 19's intervention before the Conseil d'Etat regarding website blocking of 'terrorist' content, available [here](#).

⁶ See ARTICLE 19's [analysis](#) of the 'NetzDG' law or Law on the enforcement of the law on social networks, 2017.

⁷ The text of the ECD is available from here: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0031&from=EN>

environment for the development of information society services; **(ii)** tackling illegal content online; whilst **(iii)** protecting freedom of expression.

5. The Directive does not focus on ‘platforms’⁸ as such but on various *activities* of information society service providers, including ‘mere conduit’, ‘caching’ and ‘hosting’. Of greater relevance to social media platforms is Article 14 ECD, which provides conditional immunity to information society providers for hosting illegal content. If social media platforms fail to remove illegal third-party content ‘expeditiously’, their immunity falls away and they *may* be held liable if the aggrieved party decides to sue them and wins. As such, social media platforms may be held liable for a wide range of content, from privacy laws, to defamation or intellectual property laws.⁹ In this regard, it is worth noting that the ECD applies horizontally, i.e. regardless of the type of content at issue, whether civil or criminal. In practice, however, the position is less clear where the content at issue is criminalised, such as incitement to racial or religious hatred. Generally speaking, the author of the content may be prosecuted and convicted. However, it is unclear that companies should be held criminally liable for content that otherwise constitutes ‘an offence’ if they fail to remove it. There has never been any serious suggestion up until now that this should be the case.
6. In addition, Article 15 ECD prohibits MS from imposing a “*general obligation on providers, when providing the services covered by Articles 12, 13 and 14, to monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity*”. Article 15 therefore provides an important safeguard for Internet intermediaries since any monitoring requirement would immediately fix them with knowledge. Moreover, the prohibition under Article 15 constitutes a vital safeguard for the protection of freedom of expression online as it effectively prohibits Member States from requiring intermediaries to adopt filters as a means of preventing access to potentially unlawful content. Such filters are inherently incapable of distinguishing lawful from unlawful information online, so that there is always a risk that they may restrict access to perfectly lawful content.
7. In essence, both Articles 14 and 15 provide the backbone for the protection of freedom of expression online in the EU. As such, any attempt to undermine or reverse these provisions would have a serious chilling effect on freedom of expression. This is especially so as the scheme of the ECD already has serious shortcomings for the protection of freedom of expression.

The ECD has serious shortcomings for the protection of freedom of expression

8. Article 14 ECD effectively forms the basis of what is known as ‘notice and takedown procedures’ (‘NTD’). The interpretation of this provision has given rise to a great deal of regulatory uncertainty, particularly around what constitutes sufficient notice for the purposes of gaining actual knowledge of “illegality”. In particular, ARTICLE 19 and many other human rights and digital groups argue that knowledge of *illegality* can only be obtained by a court order, since only a court or independent adjudicatory body can be in a position to determine the legality of content.¹⁰ However, in practice or in law depending on the country at issue, notice may be given by law enforcement agencies or other public authorities or private parties. In the absence of more detail in the Directive or the ECRs, the level of detail required to file a notice is unclear. This is a matter of

⁸ The European Commission attempted a definition of platforms for the purposes of its [Communication on Online Platforms](https://ec.europa.eu/digital-single-market/en/news/full-report-results-public-consultation-regulatory-environment-platforms-online-intermediaries) but that definition was criticized by many as being too vague: <https://ec.europa.eu/digital-single-market/en/news/full-report-results-public-consultation-regulatory-environment-platforms-online-intermediaries>

⁹ See e.g. *Tamiz v Google* [2013] EWCA Civ 68

¹⁰ See for instance the Manila Principles on Intermediary Liability, a global civil society initiative, which has been endorsed by over 100 organisations around the world: <https://www.manilaprinciples.org/>

concern for organisations such as ours as we believe that the balance of incentives in the ECD is such that social media platforms are more likely to remove content on the flimsiest of accusations lest they face liability. This has a chilling effect on freedom of expression.¹¹ Equally, we are concerned that ‘expeditious’ removal of content is increasingly interpreted by governments as a matter of as little as one hour in relation to certain types of content (usually ‘terrorist’ content).¹² Quite apart from the fact that companies should not be put in the position of determining the legality of content, this is plainly an insufficient timeframe in which to make a properly-informed and carefully-considered determination.

But current regulatory alternatives are worse

9. Despite these shortcomings, ARTICLE 19 believes that it is vital to at least maintain the basic principles underpinning Articles 14 and 15 ECD. The regulatory alternatives currently proposed to deal with illegal content online are, in our view, palpably worse for the protection of freedom of expression online.
- **Current EU self-regulatory or co-regulatory initiatives are unsatisfactory:** Governments regularly put pressure on companies ‘to do more’ to tackle illegal or undesirable content online. At EU level, the European Commission has led the adoption by social media platforms of a “voluntary” Code of Conduct on Combatting Illegal Hate Speech.¹³ The Commission is also looking to put ‘hate speech’ regulation within the purview of broadcast regulators under the revised Audio-Visual Media Services Directive (‘AVMS’).¹⁴ More recently, the Commission published its Communication on tackling illegal content online.¹⁵ ARTICLE 19 is deeply concerned about these initiatives. They effectively deputise censorship powers to online platforms, which are tasked with putting in place mechanisms to remove content as fast possible, usually on the basis of their Terms of Service or community standards and without any of the safeguards provided under international human rights law.

When companies remove content on the basis of their Terms of Service, there is no effective remedy in place to challenge those decisions. To begin with, most online platforms do not have a clear complaint mechanism in place (e.g. Facebook or Twitter). Even when they do, the remedy is entirely within the discretion of the company. Some users have attempted to take online platforms to court over the application of their Terms of Service but apart from jurisdictional issues, the applicable legal standard is that of fairness or reasonableness. To that extent, most removal decisions are likely to be justified.¹⁶ Even when content is removed on the basis of national laws, it is highly unclear that users are notified of an order to remove content and what remedies are available to them. More generally, none of the self-regulatory or co-regulatory mechanisms proposed ever suggest putting in place effective remedies to challenge wrongful removal of content.¹⁷

¹¹ See also the concerns expressed by the UN Special Rapporteur on Freedom of Expression, Frank La Rue, in his 2011 report on freedom of expression on the Internet, [A/HRC/17/27](#).

¹² See Recommendation of the European Commission on measures to effectively tackle illegal content online, March 2018: http://europa.eu/rapid/press-release_MEMO-18-1170_en.htm

¹³ For more details, see ARTICLE 19’s legal analysis of the EU Code of Conduct, *op. cit.*

¹⁴ ARTICLE19’s concerns about proposals for a revised AVMS Directive are detailed here: <https://www.article19.org/resources/new-eu-legislation-must-not-throttle-online-flows-of-information-and-ideas/>

¹⁵ ARTICLE 19’s concerns with the Communication are detailed here: <https://www.article19.org/resources/eu-fails-to-protect-free-speech-online-again/>

¹⁶ See for instance the recent French court decision concerning the removal of the painting *L’Origine du monde*, by Courbet: <https://www.theartnewspaper.com/news/french-court-makes-mixed-ruling-in-courbet-censorship-case>

¹⁷ As an exception, the EU Communication on Tackling Illegal Content makes some weak reference to counter-notices.

- Learning from the French and German regulatory models:** The French and German governments have adopted discreet laws to deal with specific types of content, terrorism and hate speech respectively. In *France*, decree no. 2015-125 lays down rules for the administrative blocking of websites that condone terrorism or distribute child pornography. Under the decree, a special division of the police forces, tasked with combating information technology crimes, can order ISPs to block access to a list of websites without prior judicial authorisation. The division has the power to decide that a website contravenes French criminal laws on terrorism and child pornography, to request that the editors of the website in question remove the allegedly unlawful content, and, where the editors are not identified on the website or refuse to comply with the removal request, to order ISPs to prevent access to the website in question. A magistrate from the privacy public watchdog CNIL is informed of this decision and may recommend its modification or initiate legal proceedings before an administrative tribunal. If internet users access a blocked website, they are redirected to a Ministry of Interior webpage explaining why access has been blocked. The French model raises several concerns for freedom of expression, particularly the use of blocking without judicial approval.¹⁸ At the same time, it is worth noting that some limited safeguards are in place, including the role of the magistrate within the CNIL that can ultimately lead to decisions being challenged in court.

In *Germany*, the Act to Improve Enforcement of The Law on Social Networks (or ‘NetzDG’) came into force in October 2017.¹⁹ The Act establishes an intermediary liability regime that incentivises, through severe administrative penalties of up to 5 million Euros, the removal and blocking of “clearly violating content” and “violating content”, within time periods of 24 hours and 7 days respectively. As regulatory offences, it is possible for the maximum sanction to be multiplied by ten to 50 million Euros. Though the Act does not create new content restrictions, it compels content removals on the basis of select provisions from the German Criminal Code, many of which raise serious freedom of expression concerns in and of themselves, including prohibitions on “defamation of religion”. The threshold at which the failures of a Social Network’s content removal and blocking processes will be considered systemic enough to attract administrative liability is unclear, and ambiguity in the definitions of key terms (including of “Social Network”) is likely to create an environment wherein lawful content is routinely blocked or removed as a precaution. The secondary review that would be provided by “self-regulation institutions”, and the limited oversight provided by the Administrative Courts do nothing to address over-blocking, and provide little protection or due process to Social Networks that in good faith refrain from blocking or removing content in the interests of respecting freedom of expression. Just over 6 months after its coming into force, the new German law has already led to over-vigorous removal of content and discussions are underway to amend it.²⁰

Redressing the imbalance of power between social media platforms and other actors

- ARTICLE 19 believes that, to the extent that state intervention might be needed, it should be focused on strengthening data protection law and the legal framework governing online political advertising during elections. Consideration should also be given to any leverage that could be obtained from competition law in order to redress the imbalance of power between platforms and other actors. Obligations related to the portability of data and interoperability of computer systems could potentially contribute to greater competition in this area. Further research should also be conducted into the extent to

¹⁸ For more details about those concerns, please see [ARTICLE 19 supports challenge to lawfulness of administrative website blocking](#), 30 July 2015

¹⁹ ARTICLE 19’s detailed legal analysis of the NetzDG law, *op. cit.*.

²⁰ See Thomson Reuters, [Germany looks to revise social media law as Europe watches](#), 08 March 2018

which the bargaining power of media actors may be strengthened to allow fairer sharing of advertising revenue with social media platforms.

2. What should the legal liability of online platforms be for the content that they host?

11. ARTICLE 19 believes that online platforms should remain broadly immune for the third-party content that they host on their platform unless they directly intervene in that content.²¹ We also believe that the notice-and-notice model of liability should be further explored, for instance in relation to copyright claims.²² We further recognise that different types of content may call for slightly different regulatory approaches.²³ More generally, we would like to see stronger procedural safeguards in place to prevent the wrongful removal of content.²⁴
12. By contrast, we are concerned about current debates in the UK and the EU that either seek to reverse the conditional immunity principles under the ECD²⁵ or actively seek to undermine them.²⁶ Although the current conditional immunity model is not without its problems (see Q1 above), we believe that its core principles should remain in place, i.e. immunity from liability until actual knowledge of illegality is obtained and a prohibition on general monitoring (Article 15 ECD).
13. We also believe that the current focus on liability of social media platforms and analogies with publishers is misguided. Social media platforms engage in three different types of activities: (i) they may produce content of their own, in which case the same liability should apply to them as publishers; (ii) they host content produced by third parties; and (iii) they distribute content, i.e. through the use of algorithms, they make certain types of content more visible and accessible to their users. This is often described as an editorial function or curation of content. However, it does not involve the production of content itself. As such, it should not attract any liability.²⁷ In this sense, this is not unlike newspapers deciding which stories ought to be published on the frontpage of their broadsheets, those that only get a small mention at the back, and those that are never reported. Newspapers are not held *liable* for these kinds of editorial choices - i.e. presentation or selection of content that is placed more prominently for users to read - but for the *content* of their articles. This, however, should not preclude greater transparency and therefore accountability in this area.²⁸
14. In relation to third-party content hosted by platforms, the current position as it has developed in the case-law of the Court of Justice of the European Union ('CJEU') is that in order for an Internet service provider to be considered a host it must be "neutral", i.e. the service provider must not have played an "*active role so as to give it knowledge of, or*

²¹ See for instance Four Special Rapporteurs on Freedom of Expression, [Joint Declaration on Freedom of Expression and the Internet](#) (2011); ARTICLE 19, [Internet Intermediaries: Dilemma of Liability](#) (2013); the [Manila Principles on Intermediary Liability](#), *op. cit.*

²² This is already the case in England and Wales with the Defamation Act 2013. See also our policy brief, [Dilemma of Liability](#), *op. cit.*

²³ *Ibid.*

²⁴ See Manila Principles on Intermediary Liability, *op. cit.*

²⁵ See Committee on Standards in Public Life, *Intimidation in Public Life: a Review by the Committee on Standards in Public Life*, December 2017, page 36.

²⁶ See EU Communication on Tackling Illegal Content, *op. cit.*

²⁷ This is unless the platforms have sufficiently intervened in the content such that it might be understood to be their own: see Graham Smith, *The Electronic Commerce Directive - a phantom demon?* 30 April 2018: <https://www.cyberleagle.com/2018/04/the-electronic-commerce-directive.html>

²⁸ Nor would it preclude liability under the ECD if the platforms have sufficiently intervened in the content so as to give it control over it – see Case C-324/09 *L’Oreal and others* [2011] ECR I-06011 (*L’Oreal v eBay*), para. 123.

control over, the data stored.”²⁹ For instance, when an information society provider such as eBay knowingly provides assistance to sellers by optimizing the presentation of their goods, it loses immunity from liability in relation to this content.³⁰ At the same time, acting non-neutrally in relation to some user content does not affect hosting protection for other user content, which has not been controlled.³¹

15. In other words, the current model of legal liability *already* takes into account whether or not online platforms are active or passive. The mere fact that social media platforms have terms of service and policies for the removal of content is generally not enough for immunity from liability to fall away and for them to be considered as publishers in the absence of notification.³² Moreover, it is important to remember that the very architecture of the ECD is designed so as to encourage a degree of self-regulation by platforms whilst protecting them from liability when they try to act as ‘Good Samaritans’.³³
16. Ultimately, ARTICLE 19 argues that at a minimum, the current conditional immunity from liability model should be retained as the least damaging to freedom of expression compared to current proposals. At the same time, we are concerned that under pressure from governments, companies have been encouraged to deploy the use of algorithms to take down content - often in opaque ways and such that content may be prevented from even being published in the first place without any scrutiny.³⁴ ARTICLE 19 therefore suggests exploring the possibility of establishing new models of self-regulation for social media (e.g. ‘social media council’), inspired by the effective self-regulation models created to support and promote journalistic ethics. With some adjustments, the models could be explored for a variety of content regulation issues. For more details, please see our response to Q3.

3. How effective, fair and transparent are online platforms in moderating content that they host? What processes should be implemented for individuals who wish to reverse decisions to moderate content? Who should be responsible for overseeing this?

17. ARTICLE 19 notes that companies have become somewhat more transparent about their internal content moderation processes over the years. We now know for example that they use algorithms to identify e.g. terrorist content or child abuse images. They have also become more upfront about the use of trusted flaggers, whose content takedown notices are fast-tracked for review. Similarly, companies such as Twitter and Facebook have updated and sought to clarify their Terms of Service and online content policies.³⁵ They have also improved their Transparency Reports so that Twitter, for example, publishes government takedowns requests on the basis of its Terms of Service.³⁶

²⁹ See *Google France, SARL and Google Inc. v. Louis Vuitton Malletier SA and Others*, Cases C-236/08 to C-238/08 *Google France & Google* [2010] ECR I-2417.

³⁰ See *L’Oreal v eBay, op.cit.* at para. 123.

³¹ *Ibid.*

³² See *Tamiz v Google*, [2013] EWCA Civ 68. For a case comment on the decision, see e.g. [here](#).

³³ See Recital 40 ECD “*this Directive should constitute the appropriate basis for the development of rapid and reliable procedures for removing and disabling access to illegal information; such mechanisms could be developed on the basis of voluntary agreements between all parties concerned and should be encouraged by Member States; it is in the interest of all parties involved in the provision of information society services to adopt and implement such procedures (...)*”. The same reasoning underpins section 230 of the US Communications Decency Act 1996.

³⁴ This is usually the case of videos, which have been previously flagged as being e.g. ‘terrorist’ content.

³⁵ For instance, Facebook updated its community standards in April 2018.

³⁶ <https://transparency.twitter.com/en/gov-tos-reports.html>

18. However, significant problems remain. Community standards are often coined in broad terms that fall below international standards on freedom of expression. They also ban content that may be lawful under national law. It is unclear how algorithms are used and the extent to which legitimate content is removed.³⁷ Appeals processes, when they exist, are not easily accessible and short on detail and procedural safeguards. For instance, Facebook recently announced that it would ‘expand’ its appeals process.³⁸ However, the announcement so far suggests that individuals are not notified that a request has been made to remove their content and therefore given an opportunity to challenge a content takedown request *prior* to a removal decision. Even if a review process takes place *ex post facto* for reasons of practicality, it is unclear that users are told the reason for the removal and what the review entails, e.g. whether the decision is taken by the same person. Ultimately, social media platforms retain huge discretion in relation to content removal and whether to grant a remedy.
19. For this reason, we believe that social media platforms should at a minimum comply with the UN Guiding Principles on Business and Human Rights and the standards outlined in the Manila Principles on Intermediary Liability.³⁹ The Santa Clara Principles on Transparency and Accountability in Content Moderation are also a helpful starting point.⁴⁰ See also our response to Q5 and 6 below.
20. At the same time, ARTICLE 19 believes that other solutions are needed to provide greater transparency and accountability for platforms’ decisions to remove content and the way in which they distribute content. For this reason, we suggest the creation of independent self-regulatory bodies for social media (e.g. ‘social media councils’ or ‘SMCs’). Our proposal is set out in more detail elsewhere⁴¹ and remains open for discussion but the councils would essentially present the following features:
- SMCs would deal with content moderation issues (whether one or more), including user complaints about wrongful removal;
 - SMCs would be funded by social media companies and relevant stakeholders;
 - SMCs would be established at national level with some international coordination;
 - SMCs would elaborate ethical standards that would be specific to online distribution of content and would cover topics such as the terms and conditions, the community guidelines and the practices of content regulation of social media companies;
 - Through light sanctions mainly relying on transparency, peer and public pressure, these mechanisms would promote and ensure respect of appropriate ethical standards by social media companies;
 - By making its work transparent to the general public, and through appropriate consultative processes, social media councils could provide a public forum for important public discussions on the regulation of online distribution;
 - Their transparency and openness, combined with independence, would give them the credibility they would need to gain public trust.

³⁷ YouTube’s latest transparency report seems designed to showcase the amount of content removed on its platform but it begs the question whether all of that content is illegitimate:

<https://transparencyreport.google.com/youtube-policy/overview>

³⁸ <https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/>

³⁹ *Op. cit.*

⁴⁰ The Santa Clara Principles are available from [here](#).

⁴¹ See ARTICLE 19, [Self-regulation and hate speech on social media platforms](#), March 2018

21. ARTICLE 19 recognises that - as with the development of any new system - the creation of a self-regulatory mechanism for social media is likely to raise a number of difficult questions. As the experience of establishing press councils shows, it can be a lengthy and complex process, as all stakeholders need to agree on a system that they all can make their own. Notwithstanding this, ARTICLE 19 believes that a new system can only come to existence and prove its effectiveness if all participants are willing to make it work. By shifting the focus towards the process rather than trying to impose a solution, a self-regulatory mechanism could allow for the adoption of adapted and adaptable remedies unhindered by the threat of heavy legal sanctions.
22. Developing the new system of independent self-regulation could also provide a solid reference framework to assess the initiatives undertaken by dominant social media companies and their partners so far. It would enable an assessment as to whether they are sufficiently inclusive of all the relevant stakeholders and whether they work in the public interest or are captured by private or special interests; the public would also find out what decisions have been made internally and when they have been subject to external, independent review. Ultimately, the new system would provide greater accountability to the public.

4. What role should users play in establishing and maintaining online community standards for content and behaviour?

23. ARTICLE 19 believes that users and other stakeholders such as civil society organisations can play a useful role in shaping companies' online content policies. As such, initiatives such as Facebook's Hard Questions series,⁴² which sometimes calls for input from users, are welcome. Equally, we believe that users can play an important role in challenging other users' comments, particularly when they amount to incitement to discrimination, harassment etc. or are merely offensive. The controls provided by companies, for example to block users, may also be useful in mediating interactions between users, e.g. in order to prevent harassment or abuse. At the same time, we would caution against giving users a 'hecklers' veto' over what content should stay up or be removed on online platforms. Users are unlikely to be familiar with the intricacies of e.g. 'hate speech'.⁴³ If put in charge of policing online content, it is highly likely that vast amounts of minority opinions that people simply do not like or find offensive would be taken offline.

5. What measures should online platforms adopt to ensure online safety and protect the rights of freedom of expression and freedom of information?

24. ARTICLE 19 believes that the protection of freedom of expression requires companies to be far more transparent and accountable in their online content removal practices. At the minimum, this means that:⁴⁴
- **Community standards should comply with international standards on freedom of expression.** In particular, Internet companies should provide specific examples as to the way in which their standards are applied in practice (e.g. case studies). This should be accompanied by guidance as to the types of factors that are taken into account in deciding whether or not content might be restricted.
 - **Companies should conduct regular reviews of their Terms of Service** to ensure compliance with international standards on freedom of expression both in terms of

⁴² <https://newsroom.fb.com/news/category/hard-questions/>

⁴³ See, for instance, ARTICLE 19, *Hate Speech Explained: a Toolkit*, 2015

⁴⁴ See also the Manila Principles on Intermediary Liability, *op. cit.*

formulation and in practice. In particular, companies should conduct regular audits/human rights impact assessments designed to monitor the extent to which content moderation policies adhere to the principle of non-discrimination. This would at least go some way towards guaranteeing the free expression rights of minority and marginalised groups. Any changes in companies' community standards as a result of such reviews/ human rights impact assessments should be clearly notified to users.

- **Online platforms should not require the use of real names** in order to comply with international standards on privacy. At the very least, Internet companies should ensure anonymity remains a genuine option. Equally, social media platforms should not require their users to identify themselves by means of a government-issued document or other form of identification.
- **Online platforms should ensure that sanctions for failure to comply with their community standards are proportionate.** In particular, companies' should be clear and transparent about their sanctions policy; and apply sanctions proportionately so that the least restrictive technical means should be adopted. In particular, the termination of an account should be a measure of last resort that should only be applied in the most exceptional and serious circumstances.
- **Online platforms must put in place internal complaints mechanisms:** In particular, individuals should be given notice that a complaint has been made about their content. They should also be given an opportunity to respond *before* the content is taken down. In order for them to respond, the notice of complaint should be sufficiently detailed. If the intermediary concludes that the content should be removed or other restrictive measures should be applied, individuals should be notified of the reasons for the decision and given a right to appeal the decision. In circumstances where the intermediary has put in place an internal mechanism, whereby it takes down content merely upon notice, we believe that at the minimum, the intermediary should: **(i)** require the complainant to fill out a detailed notice, i.e. identifying the content at issue, explaining their grounds for seeking the removal of content; provide contact details of the complainant and a declaration of good faith; **(ii)** notify the content producer that their content has been removed or any other measure that has been applied to their account; **(iii)** give reasons for the decision; and **(iv)** provide and explain internal avenues of appeal.
- **Online platforms should collaborate with other stakeholders to develop new independent self-regulatory mechanisms,** as outlined in Q3.

6. In what ways should online platforms be more transparent about their business practices — for example in their use of algorithms?

25. ARTICLE 19 believes that online platforms should be more transparent about their business practices in a number of areas:

- **Clearer terms of services and more accessible complaints mechanisms:** ARTICLE 19 notes that major social media platforms have amended their community standards a number of times over the years. Unlike amendments to their privacy policy, however, users do not generally get individually notified about changes to community standards. These announcements are generally made in a company press release. In our view, this should change. Companies should notify their users about any changes to their policies. Moreover, companies' terms of service continue to be drafted in broad terms. As noted above, it is vital that companies provide case studies / examples of the way in which they apply their community standards in practice. This would at least help users better understand the

rationale behind certain decisions, which may otherwise appear biased or lacking in consistency. Finally, we note that complaints mechanisms for the wrongful removal of content, if any, remain hard to find on companies' websites. In our view, their accessibility should be improved.

- **Use of algorithms:** ARTICLE 19 believes that companies should be far more transparent about the way in which they use algorithms or 'artificial intelligence'.⁴⁵ We note, for example, that the Committee of Ministers of the Council of Europe has called on Member States to take "all necessary measures to ensure that Internet intermediaries fulfill their responsibilities to respect human rights in line with the United Nations Guiding Principles on Business and Human Rights". According to the Committee of Ministers, this means, amongst other things, that:⁴⁶

"Internet intermediaries should clearly and transparently provide meaningful public information about the operation of automated data processing techniques in the course of their activities, including the operation of algorithms that facilitate searches based on user profiling or the distribution of algorithmically selected and personalised content, such as news. This should include information on which data is being processed, how long the data processing will take, which criteria are used, and for what purpose the processing takes place".

In other words, transparency need not be absolute but should be meaningful to ensure fairness and accountability.

- **Use of trusted flagger scheme:** Social media platforms rely on 'trusted flaggers' to report certain types of content. The assumption is that those flaggers can be trusted to identify e.g. 'hate speech', 'terrorist content' etc. and that they will provide more detailed reports of violations of company community standards or the law. As such, notices by trusted flaggers are more likely to lead to prompt removal. However, very little information is available about how the scheme operates, e.g. who those trusted flaggers are in a given country, what criteria are applied to qualify as trusted flaggers, what proportion of content is removed as result of notices filed by trusted flaggers etc.
- **Transparency reports:** Companies' reporting of content removals has improved over the years. For instance, Twitter now reports content removed on the basis of its Terms of Service when the removal has been requested by the authorities.⁴⁷ However, companies continue to shy away from providing data about content removed on the basis of their own terms of service at their own initiative (e.g. through filtering) or upon request from third parties. Companies sometimes oppose the need to protect users' privacy as a reason for not providing this information. However, we believe that this should not apply in the case of lawyers or trusted flaggers, which often include copyright holders associations or other interest groups. Finally, companies should provide information about the number of complaints they receive about alleged wrongful removals of content and the outcome of such complaints (i.e. whether content was restored or not).

⁴⁵ See [ARTICLE19's written evidence to the House of Lords Select Committee on Artificial Intelligence](#), September 2017

⁴⁶ See Recommendation CM/Rec (2018)2 on the roles and responsibilities of internet intermediaries.

⁴⁷ <https://transparency.twitter.com/en/gov-tos-reports.html>

7. What is the impact of the dominance of a small number of online platforms in certain online markets?

26. ARTICLE 19 notes that, at present, there is much more information available online than ever before and that social media platforms have greatly contributed to this state of affairs. However, the dominance of a small number of online platforms remains a matter of concern. In particular, the behaviour of dominant social media platforms has the potential in some instances to become a barrier to entry in the marketplace of ideas. In our view, this might in certain circumstances warrant state intervention under Article 10 of the European Convention on Human Rights, as States' positive obligation to ensure pluralism and diversity of the media" (see also Q1).

8. What effect will the United Kingdom leaving the European Union have on the regulation of the Internet?

27. ARTICLE 19 notes that Article 15 ECD (general monitoring) was not expressly transposed in the E-Commerce Regulations 2002. This raises the prospect that the UK may wish to impose general monitoring obligations in future legislation, particularly as the UK has signaled that it did not wish to fully align with EU legislation in this area.⁴⁸ If that were to be the case, we believe that this would constitute a serious interference with the rights to freedom of expression and privacy. This would be out-of-step with major international standards on freedom of expression and privacy in this area.⁴⁹ More fundamentally, proactive filtering would mean all expression mediated by algorithms, which are inherently incapable of detecting nuance or context, i.e. the very elements that might make the difference between lawful and unlawful speech. As Graham Smith, partner at Bird & Bird as noted, "*Article 19 of the 1948 Universal Convention on Human Rights is not predicated on the assumption of mediated speech.*"⁵⁰ General monitoring would effectively delegate censorship powers to private companies and amount to a form of prior restraint. As such, we strongly urge the Committee to refrain from any recommendation that would undermine the prohibition of general monitoring on the Internet.

⁴⁸ For the implications of such divergence, see e.g. Graham Smith, *The Electronic Commerce Directive - a phantom demon?* *Op. cit.*

⁴⁹ See e.g. UN Special Rapporteur on freedom of expression, [A/HRC/17/27](#) (2011); Four Special Rapporteurs on Freedom of Expression, [Joint Declaration on Freedom of Expression and the Internet](#) (2011) and more recently, the [Joint Declaration on Freedom of Expression and Fake News, Disinformation and Propaganda](#) (2017)

⁵⁰ See Graham Smith, *Time to speak up for Article 15*, 21 May 2017: <https://www.cyberleagle.com/2017/05/time-to-speak-up-for-article-15.html>