ARTICLE 19's Recommendations for the EU Digital Services Act

Twenty years after the adoption of the E-Commerce Directive, a cornerstone of Internet freedom in Europe, the EU institutions are set to review whether it is still fit for purpose and adopt a new set of rules governing online platforms as part of a new Digital Services Act (DSA). As the EU is poised to launch its consultation on the DSA, ARTICLE 19 proposes ten key recommendations for the regulation of digital services, in particular social media platforms.

At the outset, ARTICLE 19 makes clear that we do not support the models of regulation that we have seen emerging in Western Europe, particularly in countries such as Germany,[1] France,[2] or in the European Union[3] as a whole. These proposals generally delegate censorship powers to companies, mandate or encourage the use of upload filters, and require the removal of content within excessively short timeframes. We have long opposed these measures and as such, we believe that our position and recommendations, as set out in our policies on intermediaries,[4] still stand. In addition, ARTICLE 19 has advocated for oversight of social media platforms by an independent multi-stakeholder body such as social media councils.[5] We have also launched the #Missing Voices campaign that demands transparency and fair appeals processes to social media platforms.[6]

At the same time, ARTICLE 19 recognises that over the last decade, large tech companies have proved themselves unwilling or too slow to address challenges for the protection of freedom of expression and other rights on their platforms. We also recognise that greater regulation of the tech sector and social media platforms in the EU has been called for by many stakeholders and is envisaged in early proposals for the DSA. Yet, lawmakers should be cautious before abandoning the existing legal framework, which remains largely fit for purpose. Key provisions in the E-Commerce Directive such as immunity from liability and a ban on general monitoring obligations have been a driving factor in the development of innovative online services and the promotion of freedom of expression online. It is important to retain them but also to clarify notice and action procedures. Equally important is for tech companies to be afforded a certain amount of flexibility

[1] ARTICLE 19, Germany: The Act to Improve Enforcement of the Law in Social Networks, August 2017.
[2] ARTICLE 19, France: Analysis of draft hate speech bill, July 2019.
[3] ARTICLE 19, Comments on leaked draft Terrorist Content Regulation, March 2020.
[4] ARTICLE 19, Internet intermediaries: Dilemma of liability, August 2013; and ARTICLE 19, Side-stepping rights: Regulating speech by contract, June 2018.
[5] ARTICLE 19, Protecting freedom of expression in content moderation on social media.
[6] ARTICLE 19, MissingVoices: A campaign calling for better accountability and transparency.

to address challenges as they arise. How they do so could be subject to oversight by an independent regulator with strong transparency and due process obligations. An intermediate multi-stakeholder forum could also facilitate the elaboration of a shared understanding of appropriate solutions between companies and regulators.

Finally, central to concerns over platform power in our democracies is the reality that a handful of companies have become so vital to how we communicate and share information with one another. For this reason, ARTICLE 19 proposes to open the market to more competition by unbundling content moderation from other services offered by social media platforms. Our proposals for how this could work are set out below.

Recommendation 1: Overarching principles of any regulatory framework must be transparency, accountability and the protection of human rights

ARTICLE 19 believes that any regulation of digital services must have transparency, accountability and the protection of human rights at its heart. The latter means that the legality and proportionality principles must be upheld throughout. User choice must also be central to technology design and policy solutions. In addition, any regulation in this area must be based on robust evidence in order to adopt the most appropriate responses to the challenges posed by technology and the platform ecosystem in particular. It must also provide sufficient flexibility for the development of technical and practical solutions that meet the requirements of international standards on human rights. By contrast, we believe that the objectives of any new regulatory framework cannot be limited to 'tackling illegality'. The 'prevention of harm' is also much too broad a concept to be meaningful, let alone a legitimate objective of any such framework.

Recommendation 2: Conditional immunity from liability for third-party content must be maintained but its scope and notice and action procedures must be clarified

ARTICLE 19 believes that intermediaries must continue to benefit from conditional immunity from liability for illegal content. Removing or unduly limiting immunity from liability would either give an incentive to filter and remove as much users' content as possible, or it would give them an incentive to be entirely neutral and not remove any content at all. In other words, it would either lead to increased censorship or remove any incentives for companies to engage in content moderation. In our view, both these outcomes would be highly undesirable for the protection of freedom of expression online. We suggest a different liability regime in respect of different types of activities and services as follows:

Provision of infrastructure services, including 'mere conduit' and neutral 'hosting'

ARTICLE 19 believes that companies providing essential Internet infrastructure services, such as content delivery networks, should benefit from broader immunity from liability than services that are engaged in content moderation at the application layer. They should only be required to remove content

by order of a court. In practice, this means that infrastructure providers, such as Cloudflare, should not be penalised for hosting websites such as 8chan or DailyStormer, unless they have failed to comply with a valid court order requiring them to discontinue their services to such a website. Equally, they should not be required to host such a website if they do not wish to do so, except in circumstances where no other alternatives are available. In other words, infrastructure providers should not be mandated to carry content, save where the service they provide is deemed essential by a court for the promotion of pluralism and diversity and there is no other alternative for that content to be hosted. When infrastructure service providers decide to discontinue the provision of their services, they should at least clearly set out the reasons why they have done so.

Provision of 'hosting' services coupled with content moderation

Notice and action procedures

ARTICLE 19 submits that the current standard of knowledge required to benefit from immunity from liability must be maintained, i.e. it should remain 'actual' rather than 'constructive knowledge', and that actual knowledge of *illegality* can only be obtained by an order of a court. To hold otherwise would be to accept that content is illegal simply because a third party, such as a copyright holder, said so.

In the alternative, we believe that a regulatory framework could clarify the different types of notice and action procedures applicable to different types of content. ARTICLE 19 has previously set out how this could work in practice.[7] We believe that our suggested processes remain valid today and provide the best way forward to protect the right to freedom of expression. In summary:

- **Notice-to-notice for private disputes** (such as copyright or defamation): under this procedure, the complainant or 'trusted flagger' would be required to give their name and set out in a notice why they believe that their rights have been infringed, the legal basis for their claim, the location of the allegedly infringing material, and the time and date of the alleged infringement. The hosting provider would be required to pass on the notice to the alleged wrongdoer (i.e. the content provider) as soon as practicable but no more than within e.g. 72 hours. The content provider would have a choice to remove the content or file a counter notice within a reasonable period of time (e.g. 14 days). Again, the hosting provider would be required to pass on the counter-notice as soon as practicable but within a maximum period of time (e.g. 72 hours). The complainant would then be given a period of time (e.g. 14 days) to decide whether they want to take the matter to court. The content would be removed following a court order. A hosting provider could be held liable for statutory damages if they failed to comply with their 'notice-to-notice' obligations, or if they failed to remove the content following a court order. By contrast, if the content provider failed to respond or provide a counter-notice within a given period of time, the hosting provider would lose immunity from liability. They could either remove the allegedly unlawful content or may be held liable for the content at issue if the complainant decides to take the matter to court or other independent adjudicatory body. In order to protect freedom of expression, any new notice-and-notice framework should also provide for penalties for abusive notices.

[7] *C.f.* Dilemma of liability, *op.cit*. See also Manila Principles of Intermediary Liability, 2015, that further provide useful guidance on how 'notice and action' procedures should work.

- **'Notice and takedown' for allegations of serious criminality**: under this procedure, a hosting provider would be *required* to takedown content when it receives a court order to that effect. In other words, they would be liable for failing to comply with such an order. In practice, this would mean that if law enforcement authorities believe that a piece of content should be removed and the matter is not urgent, they should seek a court order, if necessary on an ex parte basis. If, however, the situation is urgent, e.g. someone's life is at risk, law enforcement should be given statutory powers to order the immediate removal or blocking of access to the content at issue. However, any such order should be confirmed by a court within a specified period of time, e.g. 48 hours. The use of informal mechanisms - e.g. phone calls or emails requesting the host to remove content - should not be permitted.

  By contrast, if hosting providers receive notice from an ordinary user about suspected criminal content, the host or platform should in turn notify law enforcement agencies if they have reasons to believe that the complaint is well-founded and merits further investigation. The host or platform *may* also decide to remove the content at issue as an interim measure in line with their terms of service. However, they would not be required to do so and failing to remove the content at issue would not attract liability.

  The same process would apply to private bodies that work with law enforcement agencies and operate hotlines that individual internet users can call if they suspect criminal content has been posted online (see e.g. the Internet Watch Foundation in the UK or SaferNet in Brazil). In other words, the hotline would report the content at issue to both the host and law enforcement agencies. The host would use the same process that it uses for complaints from ordinary users, i.e. it would remain free to decide whether to remove content on the basis of its terms of service. The same model could be applied to other bodies, whether public or private, which receive complaints from the public concerning potentially criminal content online, or to notices issued by 'trusted flaggers' (see below for further details on trusted flagger programmes). Whichever option is pursued, it is important that the authorities are notified of any allegation of serious criminal conduct so that it may be properly investigated and dealt with according to the established procedure of the criminal justice system.

We believe that this is the most proportionate and rights-respecting way in which 'notice and action' procedures can be operated, particularly against small companies.

Content moderation measures applied by company of its own motion

ARTICLE 19 believes that platforms and other tech companies should not be held liable simply because they adopt community standards, and use human moderators or other tools to enforce them. In this sense, we support the adoption of a Good Samaritan rule that would encourage 'good' content moderation efforts made in good faith. In our view, failure to do so would prevent the adoption of innovative technical solutions and tools, such as demonetisation or the removal of certain platform features, that would strike a more proportionate balance between the protection of freedom of expression and tackling illegal or even 'harmful' content. At the same time, companies that use these tools should be subject to stringent transparency and due process requirements in relation to the way in which they use them (see Recommendation 5 below). ARTICLE 19 also suggests that a multi-

stakeholder forum such as the SMC could facilitate the development of such technical or practical solutions in line with international standards on freedom of expression.

Similarly, companies should benefit from broad immunity from liability for the recommendations or suggestions made by their algorithms, even in circumstances where those algorithms recommend illegal content in response to content viewed by users. Whilst system developers and coders define the parameters within which 'algorithms' operate, they do not control or determine the outcome of these automated processes. Algorithms produce results from datasets in ways which are both complex and unpredictable. They are also both generally prone to making mistakes and unable to distinguish between lawful or unlawful content. Holding companies liable for every possible 'mistake' made by their systems would therefore be both unworkable and disproportionate. Insofar as liability deals with specific instances of illegality, it is also a poor instrument to address the systemic challenges thrown up by algorithms. Instead, companies - particularly the ones with significant market share - should be subject to greater transparency obligations and required to carry out human rights impact assessments as outlined below. In our view, the same reasoning should apply to navigation or 'discovery' services, i.e. they should not be penalised if their search engine algorithm returns illegal content but they should be transparent and explain to the public how their algorithm functions to return search results.

By contrast, we accept that companies should lose immunity from liability when they 'promote' - or 'optimise' the presentation of - illegal content in the advertisement section of their platform as a result of commercial agreements.[8]

Recommendation 3: General monitoring of content must be prohibited

ARTICLE 19 believes that governments must continue to prohibit general monitoring of content. Although it may be argued that monitoring merely enables companies to detect potentially illegal or 'harmful' content, in practice, mere detection is almost always coupled with removal or other types of actions reducing the availability of such content. This is deeply problematic given that content monitoring technology is not nearly as advanced as it is sometimes suggested. In particular, hash-matching algorithms and natural language processing (NLP) tools are currently incapable of distinguishing content whose legality may vary depending on context, such as news reporting or parody. Vast amounts of legitimate content may therefore be removed. Moreover, these technologies interfere with the privacy rights of users, as they require an analysis of individuals' communications.

In addition, if a law were to make immunity from liability conditional on 'general monitoring' or the adoption of 'proactive measures' or 'best efforts' to tackle illegal content - such as the EU Copyright Directive in the Digital Single Market, companies would inevitably err on the side of caution and remove content by default in order to avoid legal risks and enforcement costs. As noted

[8] See, *mutatis mutandis,* the Court of Justice of the European Union, *L'Oreal v eBay*, C-324/09, 12 July 2011.

by scholars[9], this could lead to platforms only allowing pre-screened speakers or using their Terms of Service to prohibit controversial content. It could also deter new market entrants from challenging incumbents.

At the same time, ARTICLE 19 recognises that 'specific' monitoring and removal of videos or other images that contain incontrovertibly unlawful child pornography, i.e. the depiction of sexual activity such as penetration between a child and an adult, may be compatible with the rights to freedom of expression and privacy. We do so given the gravity of the conduct at issue and the fact that this type of content can reliably be recognised as unlawful regardless of context. We do not, however, agree that such specific monitoring obligations should be applied to any other kind of content.

Recommendation 4: Any regulatory framework must be strictly limited in scope

As noted above, ARTICLE 19 believes that any regulatory framework aiming to regulate the activities of 'platforms' ought to be limited in its scope, including by reference to its subject-matter, the entities it seeks to cover and its geographical application:

- **Focus on illegal rather than harmful content:** We believe that any such framework should be limited to 'illegal' rather than 'harmful' content for the simple reason that 'harmful' content is an inherently vague concept. This makes it difficult to enforce, prone to abuse and open to challenge on legality grounds. In our view, legal content that is nonetheless prohibited under the community standards of companies should be subject to oversight by independent multi-stakeholder entities such as ARTICLE 19's proposed Social Media Councils. If "legal but harmful content" is included within the scope of the DSA contrary to our recommendations, then it should only impose transparency and due process requirements for the purposes of the company's enforcement of its community standards. The role of the regulator would therefore be limited to ensuring that companies' content moderation systems are sufficiently transparent and that users have clear and effective redress mechanisms available to them.

- **Private messaging services and news organisations should be out of scope:** Similarly, we believe that the scope of application of any regulatory framework should be limited so that below the line comments on news sites and blogs are excluded. Equally, messaging applications and other private channels of communication should be out of scope. In particular, regulators should not have the power to impose obligations on providers where such obligations would entail an unjustifiable interference with users' privacy rights, such as a weakening of end-to-end encryption or mandatory filters.

- **Measures should not have extraterritorial application:** Finally, we believe that the implementation of measures under such a new regulatory framework should be geographically limited to the country mandating such measures, consistent with international principles of comity and the proportionality principle under international human rights law. In other words, no one country

[9] J. van Hoboken & D. Keller, Design Principles for Intermediary Liability Laws, Transatlantic Working Group, October 2019

should be able to issue orders to remove or otherwise restrict content that may be lawful outside its borders.

Recommendation 5: Obligations under any regulatory scheme must be clearly defined

ARTICLE 19 believes that any obligations under a new regulatory scheme governing the activities of platforms and other tech companies must be clearly defined. Below we set out the types of measures that could be included as part of such a framework and those that should not. In particular, we believe that a new regulatory framework could mandate the following:

- **Transparency obligations**: In our view, transparency should be a basic requirement that pervades everything that companies do. In particular, it should apply to:
    - *Distribution of content:* digital companies should provide essential information and explain to the public how their algorithms are used to present, rank, promote or demote content. Content that is promoted should be clearly marked as such, whether the content is promoted by the company or by a third-party for remuneration. Companies should also explain how they target users with (unsolicited) promoted content, whether at their own initiative or on behalf of third parties as a paid service. Equally, companies should clearly highlight content whose reliability is in doubt or content that has been fact-checked.

    - *Companies' terms of service and community standards*: companies should publish community standards/terms of service that are easy to understand and give "case-law" examples of how they are applied. As suggested by the French Government interim report,[10] they should publish information about the methods and internal processes for the elaboration of community rules, which should continue to include consultations with a broad range of actors, including civil society.

    - *Human and technological resources used to ensure compliance*: companies should include detailed information about trusted flagger schemes, including who is on the roster of trusted flaggers, how they have been selected and any 'privileges' attached to that status. They should also publish information about the way in which their algorithms operate to detect illegal or allegedly 'harmful' content under their community standards. In particular, this should include information about rates of false negatives/false positives and indicators, if any, to assess content that is likely to become viral, e.g. by reference to exposure to a wider audience.

    - *Decision-making:* companies should notify their decisions to affected parties and give sufficiently detailed reasons for the actions they take against particular content or

[10] The French government 'Facebook Mission' Interim Report, July 2019.

accounts. They should also provide clear information about any internal complaints mechanisms.

- *Transparency reports*: companies should publish detailed information consistent with the Santa Clara Principles that have been developed by experts in the field. We note that it is particularly important not to limit statistical information to removal of content but also include data about the number of appeals processed and their outcome. Transparency reporting should also distinguish between content flagged by third-parties (including whether they are public bodies or private entities), trusted flaggers (whether public bodies or private entities) or algorithms. Further information should also be provided in relation to the different types of restrictions applied to content as part of content moderation processes, such as demonetisation or downgrading; for every restriction, the company should give information about the rules on the basis of which the decision was made and, where available, the outcome of any appeals.

  More generally, we note that any transparency reporting requirements should aim to provide far more qualitative analysis of content moderation decisions. It is vital that the metric of success in addressing illegal content is not tied to content removal rates as it encourages over-removal. Equally, transparency reporting should not be limited to information submitted by companies but should include information submitted by relevant government agencies. The above is without prejudice to any measures that may be applicable under consumer law.

- *Algorithms transparency audits:* companies should give greater access to datasets for regulators and vetted independent researchers, whether academics, journalists or otherwise, in order for them to verify that the company's systems and algorithms are operating as the company says it does. In particular, auditors should be given access to data about: (i) companies' content moderation programmes; (ii) how companies order, rank, prioritise, recommend or otherwise personalise content; and (iii) how this applies to political advertising. Whilst regulators could be given access to sensitive and commercial data, vetted third-parties could be given access to anonymised datasets. These audits of platforms' operations should take place on a regular basis.

- **Archives of digital and political advertising:** the current practice as set out in the EU Code of Practice on Disinformation[11] relies on the platforms themselves to develop their own 'repositories' containing information about political ads, including actual sponsor identity, amounts spent and targeting criteria used. In the current context, where online political advertising is dominated by Facebook and Google, it is understandable that the repositories are created and managed by individual platforms. In the long-term, however, this raises

[11] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Tackling online disinformation: a European Approach, COM/2018/236, 26 April 2018.

questions about the retention and storage of historical data, since events such as the insolvency of a platform could lead to data loss. Regulation in this area could therefore have the objective of developing a practical, feasible external repository that will permit regulators and interested stakeholders to access cross-platform data for a legally mandated period of time.

- **Internal due process obligations:** without prejudice to Recommendation 8 below, we believe that any regulatory framework regulating the activities of platforms with significant market power, should include a requirement to put in place:
    - Clear notice and takedown rules in line with the Manila Principles on Intermediary Liability;
    - Internal redress mechanisms to deal with complaints about restrictions on the exercise of the right to freedom of expression, such as the wrongful removal of content or the wrongful application of labels that would suggest that a news source is untrustworthy. Conversely, appeals mechanisms should also be able to address a company's refusal to remove content that is arguably in breach of the company's community standards. In all cases, internal complaint mechanisms should respect due process safeguards as set out in our Sidestepping Rights policy.

- **Obligation to promote content diversity:** Given the risks of overly personalised content on social media platforms, large social media companies should be required to take steps to ensure users are exposed to sufficiently diverse content and balanced coverage of issues of public interest on their service by default. This obligation should only be imposed in situations of market dominance and where it is necessary to support the online visibility of the broad diversity of viewpoints and opinions in society and with a view to enabling individuals to make informed decisions. In particular, as noted above, social media companies should provide sufficient information to explain how newsfeeds and the material they promote is selected. At the same time, individuals should be able to opt-out of content diversity default-settings in order to tailor their newsfeeds to their own interests and preferences. We also suggest that a multi-stakeholder forum such as the SMC can serve to elaborate the appropriate approach to content diversity on social media platforms.

- **Obligation to promote media pluralism**: when there is excessive market concentration, a small number of social media platforms act as gatekeepers of the flow of media content. In these circumstances, social media platforms should ensure that a diversity of media actors get their content distributed on their platforms. Regulators or competition authorities should be able to impose appropriate obligations, such as an equivalent to a must-carry duty in legacy media regulation in order to sustain media pluralism on social media platforms.

By contrast, ARTICLE 19 believes that any such regulatory scheme **should not** include the following - non-exhaustive - types of obligations:

- **A broad and undefined 'duty of care'** to prevent an equally undefined notion of 'harm': In our view, such notions would be unlikely to pass the legality test under international human rights law. In

practice, they would both create legal uncertainty and give largely unfettered powers to regulatory authorities, which would be deeply problematic for freedom of expression.

- **A general obligation to monitor content**: as noted above, a new framework should refrain from mandating general monitoring of content or measures that are substantially equivalent to it, such as mandating 'best efforts' or 'proactive measures' to tackle illegal content. Equally, such a framework should refrain from 'nudging' companies towards the adoption of such measures by framing them as purely voluntary or simply 'recommended', when in reality, failure to adopt them could lead to heavy sanctions.

- **Unduly short timeframes**: Internet companies should not be required to remove content within unduly short timeframes, particularly when the content at issue may give rise to difficult questions of interpretation, such as 'hate speech' or 'terrorist' content. Short removal timeframes do not incentivise companies to review notices sufficiently carefully. As such, they promote the wrongful removal of content and fail to protect freedom of expression. Moreover, as Facebook itself has noted, removals within short time frames can incentivise companies to allocate resources to removal of notices regardless of their severity or to focus on content simply because it has been posted in the last 24 hours rather than older content that may well be more deserving of attention.

- **Compliance targets**: Equally, lawmakers or regulators should not impose numerical compliance targets that could have the effect of encouraging companies to expand the definition of content they disallow on their platform in order to boost their compliance rate. In other words, numerical targets would encourage the removal of ever-greater amounts of legitimate content. We further note that insofar as lawmakers may be considering various metrics and thresholds to ensure compliance, they should consider the extent to which society can be expected to tolerate a degree of risk of harm online, as it does in the offline world.

- **Obligation to cooperate or report illegal content**: Vague obligations to cooperate are problematic because they could involve serious interference with users' rights, such as access to user data by law enforcement without sufficient safeguards. At the same time, being vague makes it arguable for companies that they have cooperated in other less intrusive ways. In short, such obligations are likely to be difficult to enforce and as such, it is unclear that they are necessary. Obligations to report illegal content would likely give a strong incentive to companies to focus on notices they receive of allegedly illegal content regardless of its severity and report it to law enforcement. They could also disincentivise companies to invest in automated tools to detect illegal content if they could be fixed with knowledge of illegality or found in breach of their obligations for failing to report all the potentially illegal content they identify automatically on their networks. Both these outcomes would be undesirable and would also likely have a negative impact on freedom of expression since vast amounts of legitimate content would be reported.

Recommendation 6: Any regulator must be independent both in law and practice

ARTICLE 19 believes that for any online content regulatory scheme to have any kind of legitimacy, it must be overseen by an independent regulator, i.e. free from political or commercial interference. Its independence and institutional autonomy must be guaranteed and protected by law including through: a clear statement of overall platform and online content regulation policy, clearly laying out the powers and responsibilities of such a body; rules of membership; funding arrangements and accountability to the public through a multi-party body, publication of an annual report and general information about its activities on a website and other communication channels with the public. The government should be kept at arm's length and not be involved in any of those. More specifically, for freedom of expression to be protected by a regulator, the law setting it up should contain:

- An overarching provision stressing the importance of protecting freedom of expression, including expression that may shock, offend or disturb;

- A provision making clear that the mission of any regulator in this area includes the protection of human rights, including freedom of expression;

- A provision requiring any regulator to audit content removal decisions and consider the extent to which companies over-remove or over-restrict content, whether upon request or of their own accord;

- A provision making clear that companies should not be penalised for failing to remove lawful content;

- A provision making clear that the protection of promotion of media pluralism and diversity is one of the essential objectives of regulation.

On a more practical level, the regulator should be equipped with appropriate knowledge, expertise and skills in order to be able to address the specific challenges posed by social media platforms and other internet actors. In our view, this would be particularly important if the remit of a telecom and/or broadcast regulator is expanded to include a wide range of other companies whose industry culture and practices are very different from telecom and broadcast media companies. Moreover, the regulator should have adequate resources in order to fulfil its role.

ARTICLE 19 also notes that the legal and regulatory framework could facilitate the creation of an independent, accountable and transparent multi-stakeholder mechanism (such as SMCs), in a co-regulatory approach, provided that the law includes sufficient safeguards to ensure the independence and effectiveness of such a forum. Under monitoring of the independent public regulator, the intermediate multi-stakeholder body could facilitate the elaboration and development of the appropriate technical or practical remedies that are necessary to achieve the general interest objectives set in the law, in compliance with international standards on FoE.

Finally, any regulator tasked with overseeing the operations of a broad range of providers of digital service should ensure cooperation with other relevant regulators such as data protection, consumer protection and competition authorities, or with relevant multi-stakeholder mechanisms such as SMC.

Recommendation 7: Any regulatory framework must be proportionate

ARTICLE 19 believes that for any regulatory framework to comply with international standards on freedom of expression, it must be strictly proportionate to the aim pursued:

- **Tiered approach:** States should be extremely cautious about adopting measures that are meant to hold social media companies with a certain degree of market power to account but would ultimately impose an undue burden on other, smaller, service providers. As such, we believe that a tiered approach in this area would be necessary. In other words, social media platforms with a certain degree of market power could be made subject to more stringent obligations than smaller players. In order to assess the degree of market power of a platform, regard could be had, among others, to the following factors: (i) the number of its users, (ii) its annual global turnover and; (iii) the capacity to play a role in access to the market (gatekeeping) or in the functioning of the market ('regulatory role'). Non-profits, such as Wikipedia/Wikimedia Foundation, should be exempt and continue to operate under a broad immunity from liability framework.

  At the same time, we highly recommend that any proposed measures should be the subject of rigorous impact assessments, including in relation to possible undesirable outcomes that would entrench the dominance of certain players. Dominant social media companies are likely to be able to adapt to any demands placed upon them. Such demands could ultimately lead to a perception that they are 'safer', which would almost certainly give them an advantage over smaller entrants who would not be able to engage in the same kind of content moderation exercise as the incumbents.

- **Evaluating systemic failures:** For regulation in this area to be sustainable and proportionate, companies should not be assessed because they have failed to remove a single piece of content or only published a single dataset. Instead, a regulator should evaluate whether they have failed to comply with their obligations under the law on a systemic basis. The threshold for systemic failures should be defined by law by reference to clear criteria that should enable a holistic assessment rather than a purely numerical one.[GG26] For instance, the law should not sanction companies because they have failed to remove a given quantity or percentage of content flagged as either illegal or harmful. Rather, it should contain an overall assessment of the measures they have adopted to mitigate risks to human rights.

- **Proportionate sanctions:** Failure to comply with the obligations outlined above should be meted out with proportionate sanctions. Whilst this may include significant fines, these should not be set so high as to provide a disincentive to protect freedom of expression. In our view,

4% of global turnover is likely to be too high for freedom of expression to be protected. Equally, criminal sanctions imposed on chief executives for failure to comply with these obligations could have a chilling effect on freedom of expression. Faced with the prospect of several years in prison, company executives would almost certainly adopt policies that would favour greater removal or other types of restrictions on content. As such, governments should refrain from adopting such criminal sanctions.

Recommendation 8: Any regulatory framework must provide access to effective remedies

Beyond internal complaints mechanisms, ARTICLE 19 believes that governments should ensure that internet users have access to judicial remedies in order to challenge wrongful removal of content by platforms on the basis of their terms of service. Such remedies should include access to the courts but also to alternative dispute resolution mechanisms, such as e-courts or an ombudsman. In practice, governments should develop proposals for funding such mechanisms, including through a levy on social media platforms, for example.

This should be without prejudice to self-regulatory schemes, such as social media councils, that would, among other things, enable users to challenge the content moderation decisions made by social media platforms by reference to an agreed set of principles, such as a 'Charter of Users' Rights'.

Recommendation 9: Large platforms should be required to unbundle their hosting and content moderation functions and ensure they are interoperable with other services

ARTICLE 19 believes that in order to address the excessive concentration in social media markets, content moderation should be decentralised. In practice, a new legal framework could impose a combination of data portability, interoperability and unbundling requirements, consistent with data protection laws. In other words, regulators could mandate platforms with a certain degree of market power to separate their hosting and content moderation functions, and to allow third parties to access their platform (in practice their Application Programming Interface or API) in order to provide content moderation to users. This kind of functional separation would not impede large social media companies from offering content moderation to their users; however users would decide whether to opt-in if they want to have the same provider offering both hosting and content moderation services. In other words, when creating a profile on Facebook, for example, the user would be asked to select a content moderation provider, and Facebook could remain one of the options to select, but it should not be the default one. Ideally, and to avoid further lock-in, users would remain free to change their choice at any time, through the platform's settings. In our view, these kinds of solutions should be further explored in order to give users real choice and help them take back control; it would ensure healthy competition and innovation in social media markets and return to the promise of a free, diverse and decentralised Internet.

At the very least, we believe that users should be given greater content moderation options by social media platforms, both in terms of the type of content they would like to view more of and according to what criteria, e.g. in chronological order.

Recommendation 10: Data collection in the provision of digital services and digital advertising should be strictly limited

ARTICLE 19 believes that insofar as the business model of tech companies raises challenges in terms of power imbalances and the protection of the right to privacy, governments should ensure that those companies comply with strict data protection laws. In addition, the advertising industry itself should be subject to more robust oversight in relation to its data collection practices.