Guidance for regulating digital platforms: a multistakeholder approach

Draft 1.1

December 2022

Disclaimer: This document presents the overall UNESCO initiative and an initial draft proposal of the guidance document for regulating digital platforms. A further developed draft of this document will be circulated by the end of January 2023 ahead of the Global Conference Internet for Trust, which will provide a space for debate about the broader issues behind the paper, the proposals themselves, and future actions. A glossary of key concepts, based on international human rights law, UN system resolutions and declarations, other soft law documents produced by the Universal System of Human Rights, is being developed and will be added in the next version of this draft.

All comments should be sent to the e-mail: internetconference@unesco.org, mentioning the specific paragraph number the comment refers to.

Introduction

1. UNESCO is developing, through multistakeholder consultations, a guidance document for actors seeking to regulate, co-regulate and self-regulate digital platforms, with the aim of supporting freedom of expression and the availability of accurate and reliable information in the public sphere, while dealing with content that potentially damages human rights and democracy. The scope of this guidance covers digital platform services that can disseminate users’ content to the wider public, including social media networks, search engines and content sharing platforms. While this guidance is developed for those platforms whose services have the largest size and reach, minimum safety requirements should be applied to all platform service companies regardless of size.

2. This document aims to provide high-level guidance for those Member States and other relevant stakeholders that are considering how to regulate online content. It sets standards to help them in the development of legislation and policies that are consistent with international human rights standards, and which enhance the availability of accurate and reliable information in the public sphere. It also intends to serve as a guidance for co-regulatory and self-regulatory processes, as well as a concrete tool for a process of checks-and-balances, through which companies, civil society organizations, academics, the technical community, and journalists can hold accountable the players in charge of regulating, co-regulating and self-regulating this space.
Why UNESCO?

3. UNESCO has a global mandate to promote the free flow of ideas by word and image. As part of the Organization’s Medium-Term Strategy for 2022-2029 (41 C/4), Strategic Objective 3 is to build inclusive, just and peaceful societies by promoting freedom of expression, cultural diversity, education for global citizenship and protecting heritage. Strategic Objective 4 is to foster a technological environment in the service of humankind through the development and dissemination of knowledge and skills and the development of ethical standards.

4. The development of guidance for Members States (including a diverse range of public entities, among which may be different types of independent regulators) and digital platform services themselves, to secure information as a public good, contributes to all five of UNESCO’s functions as a laboratory of ideas, a clearing house, a standard-setter, a catalyst and motor for international cooperation, and a capacity-builder.

5. More specifically, the development of guidance for the regulation of digital platform services builds on the Organization’s work in the domain of broadcast regulation developed over several decades.

6. This guidance for regulation of digital platform services focuses on the structures and processes to help users have a safer, critical, and self-determined interaction with online content, dealing with content that is potentially damaging democracy and human rights, while supporting freedom of expression and the availability of accurate and reliable information in the public sphere.

7. This guidance will:
   - **7.1** Take forward the Windhoek+30 Declaration on Information as a Public Good, as it calls on all parties to mainstream media and information literacy, as well as promoting increased transparency of relevant technology companies and media viability, principles unanimously endorsed by UNESCO’s Members States during its 41st session of its General Conference.
   - **7.2** Create a multistakeholder global shared space for the debates on regulation, co-regulation and self-regulation of digital platform services, through an inclusive consultative process and research ahead of the conference.
   - **7.3** Enable a network of regulators and regulatory systems to draw upon this guidance, and facilitate the creation of an international community of practice, capable of exchanging good practices on how to approach regulation
of digital platform services to secure information as a public good while protecting freedom of expression and other human rights.

7.4 Serve as an advocacy and accountability tool for all the relevant stakeholders, who will be able to advocate for smart regulation aligned with human rights, where it is missing, and to hold relevant players (parliaments, regulators, companies) accountable, guaranteeing that any regulatory, co-regulatory and self-regulatory measures discussed and implemented are in line with international human rights standards.

7.5 Offer inputs to “Our Common Agenda”, including the Global Digital Compact and the UN Summit of the Future to be held in September 2024.

7.6 Feed into discussions about the upcoming 20-year review in 2025 of the World Summit on the Information Society (WSIS) and the review of the Internet Governance Forum (IGF).

7.7 Build on and gain insights from the work linked to the development and implementation of the UNESCO Recommendation on the Ethics of Artificial Intelligence, adopted by the UNESCO General Conference in November 2021, and particularly regarding its guidance on digital platforms and online media under the Policy Area for Communication and Information.

Independent Regulation

8. Online content represents a new regulatory challenge that many actors, including states, are struggling to deal with. Existing regulatory systems vary from country to country. In some jurisdictions, there may be an existing broadcast regulator which is being granted new powers over digital platform regulation. In other states a new regulator may be established to regulate online content. There are other cases in which more than one regulatory body or institution oversees these issues. taking into account the wider implications of digital content for our societies. For instance, there are contexts where we have audio-visual, electoral, telecom, data protection regulators that deal with different aspects of the digital platform’s services. This is why this text is using the concept of a regulatory system. Whichever is the case, this guidance outlines the importance of establishing the independence of the regulatory system, however constituted, as well as ensuring that regulators have the necessary skills.

9. In 2006, the World Bank published its Handbook for Evaluating Infrastructure Regulatory Systems¹ in which it says the following about independent regulation:

---

¹ World Bank Handbook for Evaluating Infrastructure Regulatory Systems, p.50
9.1 “The key characteristic of the independent regulator model is decision-making independence. This means that the regulator’s decisions are made without the prior approval of any other government entity, and no entity other than a court or a pre-established appellate panel can overrule the regulator’s decisions. The institutional building blocks for decision-making independence are: organizational independence (organizationally separate from existing ministries and departments), financial in-dependence (an earmarked, secure, and adequate source of funding), and management independence (autonomy over internal administration and protection from dismissal without due cause).”

10. In a guiding document commissioned by UNESCO (2016), the expert on broadcasting independent regulatory systems, Eve Salomon, highlighted:

10.1 “An independent authority (that is, one which has its powers and responsibilities set out in an instrument of public law and is empowered to manage its own resources, and whose members are appointed in an independent manner and protected by law against unwarranted dismissal) is better placed to act impartially in the public interest and to avoid undue influence from political or industry interests. This ability to operate impartially is vital to protect freedom of expression, which is necessary in a functioning democracy. Independence is also required for the proper operation of all of the major functions of broadcasting regulation, including licensing, applying content standards and positive content obligations, and ownership and competition regulation.”

11. This guidance is principle-based – with the regulator (or the regulatory system when there is more than one regulatory entity or body), setting the overall goal for regulation which the digital platform services must fulfil.

12. Importantly, this guidance recommends that any regulatory system focuses on the structures and processes that services use to make content available, rather than seeking to intervene in actual content decisions.

13. In addition to setting out the primary regulatory goal which is to support freedom of expression and the availability of accurate and reliable information in the public sphere while dealing with content that damages human rights and democracy, the guidance suggests that the regulator or the regulatory system could specify a number of issues that the digital platform services should address when reporting. These are currently set out in several separate sub-items3 –. The guidance goes on to set out the constitution, powers, and scope of the regulatory system.

---

2 [https://unesdoc.unesco.org/ark:/48223/pf0000246055](https://unesdoc.unesco.org/ark:/48223/pf0000246055) see also [Principle 17 of Declaration of Principles on Freedom of Expression and Access to Information in Africa](https://www.achpr.org/legalinstruments/detail?id=69)

3 10 in total, although this may change as the result of further consultations
The benefits of this guidance

14 The guidance should be helpful to a range of stakeholders: for policymakers in identifying some of the stakeholders, objectives, principles, processes and procedures that might be considered in legislation; for regulators in implementation; for companies in their policies and practices; and for other stakeholders in their advocacy and accountability efforts.

15 Specifically, this guidance will:

- **15.1** provide guidance in developing regulation that can help Member States address moderation and curation processes of content that potentially damages democratic discourses and structures and human rights, while protecting freedom of expression and other human rights;
- **15.2** provide guidance on the strengthening and constitution of an independent regulator(s) or regulatory system;
- **15.3** facilitate the ability of digital platforms to align to a common framework and develop coherent systems across regions to minimise internet fragmentation to protect freedom of expression and enhance the availability of accurate and reliable information in the public sphere;
- **15.4** help develop regulatory stability and a more coherent global governance (which would benefit the development of a more diverse range of companies and local economies) and help Member States be “future ready” anticipating new challenges;
- **15.5** support platforms by providing practical and implementable guidance with a view to realizing the regulatory goal.

The proposed Guidance for regulating digital platforms: a multistakeholder approach

Section One - The goal of regulation

16 The objective of this guidance is to protect freedom of expression and enhance the availability of accurate and reliable information in the public sphere, while dealing with content that potentially damages human rights and democracy.

17 The guidance sets out how the regulatory system can oversee the conduct of digital platform services in respect to content issues.

18 The guidance also outlines government responsibilities to be transparent and accountable about the requirements they place upon digital platform services, particularly regarding the alignment with international human rights standards. For example, governments should be open, clear, and specific about the type and volume of requests they make to companies to remove and block content. In the case of sensitivities about publicising these requests – for instance, content relevant to national security or the prevention of serious crime – then the regulator or regulatory system should be able
to examine, based on human rights standards, the validity of such requests and be able to report publicly on their findings and actions. The regulator should also be able to scrutinize the scope of requests to ensure adequate balance between illegality and freedom of expression.

19 This will also require finding a means to deal with the potentially harmful content that may damage democracy and human rights – current examples include hatred of defined groups; incitement to violence; harassment; mis- and disinformation; and hostility directed at women, racial and minorities, human rights defenders or vulnerable groups - while protecting international standards of freedom of expression. But we should recognise that new dangers may arise that are not foreseen now, and that any regulation to protect human rights must be flexible enough to adapt to new or changing circumstances.

20 Finally, this guidance shows that this goal can be achieved only if there is a cooperation among the companies providing services and the regulatory systems, while being effective and implementable and providing real accountability.

21 This guidance for regulation will be based on five key principles: platform, policies and operations need to be (1) human rights-based, (2) transparent, (3) empowering, (4) accountable and (5) verifiable, to help ensure:

- **21.1 Platforms have content governance policies and practices consistent with human rights standards**, implemented algorithmically or through human means (with adequate protection for the well-being of human moderators);
- **22.2 Platforms are transparent**, being open about how they operate (taking into account commercial confidentiality) with policies being explainable;
- **22.3 Platforms empower users** to use digital services in a self-determined and empowering manner, including being able to assess the quality of information received;
- **22.4 Platforms are accountable** to users, the public, and regulators in implementing terms of service and content policies, including giving users rights of redress against content-related decisions;
- **22.5 There is independent oversight** and assessment of the impact that regulation has on companies’ rules and practices, with a view to adjusting regulation to more effectively protect information as a public good.

Section Two - Fulfilling the goal

22 Before setting out the responsibilities of digital platform services in respect to the regulator, it is helpful to set out the responsibilities of governments that are considering legislation to regulate processes impacting content moderation and/or curation, so that such legislation fulfils the regulatory goals of providing and ensuring information as a public good, while protecting freedom of expression.
Governments should:

- **23.1** Protect and respect users’ rights to freedom of expression, the right to information, equality and non-discrimination;
- **23.2** Respect the requirements of Article 19(3) of the International Covenant on Civil and Political Rights (ICCPR), in that any restrictions applied to content should have a basis in law, have a legitimate aim, and be necessary and proportional to the harm that is being restricted;
- **23.3** Ensure that any restrictions are also consistent with Article 20 of the ICCPR;
- **23.4** Be transparent about the requests they make to companies to remove or restrict content, and be able to demonstrate how this is consistent with Article 19 of the ICCPR;
- **23.5** Guarantee that any content removals are subjected to the adequate due process of law, including independent judicial review;
- **23.6** Not impose indirect restrictions to companies (for example, internet shutdowns) for alleged or potential breaches of regulations;
- **23.7** Not subject staff of companies to criminal penalties for an alleged or potential breach of regulations, as this will have a chilling effect on freedom of expression;
- **23.8** Regulators with responsibilities in this area should be structured as independent regulators, with the proper accountability systems in place.

24. In turn, this guidance recommends that the regulatory system expect digital platform services to have in place structures and processes and to report to them on the following issues:

25. **Transparency of process.** How digital platform services fulfil the principles of transparency, explicable, and reporting against what they say they do in their terms and conditions (T&Cs) and community standards. This should include:

- **25.1** Information about the reasons behind any restrictions imposed in relation to the use of their service being publicly available in an easily accessible format in their terms and conditions;
- **25.2** How content is managed, including through algorithmic decision-making and human review, as well as content that is being removed or blocked under either T&Cs or pursuant to government demands/requests, and

---

4 “1. Any propaganda for war shall be prohibited by law. 2. Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.”
information relevant to complaints about the removal, blocking, or refusal to block content;

- **25.3** Any information about processes used by the platform to enforce their T&Cs and sanction users, as well as government demands/requests for content removal, restriction, or promotion;

- **25.4** Any safeguards applied in relation to any content moderation that are put in place to safeguard freedom of expression and the right to information, including in response to government demands/requests, particularly in relation to matters of public interest, so as to ensure a plurality of views and opinions;

- **25.5** How users can access the complaints process;

- **25.6** Any use made of automated means for the purpose of content moderation, including a specification of the role of the automated means in the review process and any indicators of the benefits and limitations of the automated means in fulfilling those purposes.

**Content management policies.**

**26.** The content management policies of digital platform services should be consistent with the obligations of corporations under the UN Guiding Principles for Business and Human Rights, the International Covenant for Civil and Political Rights and relevant regional treaties. They should also follow best practices as expressed, for example, in the Santa Clara Principles. Five

**27.** Any restriction upon content posted should be clearly set out in the platform rules, which should be implemented consistently, without arbitrary distinctions made between types of content or between users.

- **27.1** Platforms should, in policy and practice, through adequately trained and staffed personnel, ensure that, at a minimum, there is quick and decisive action against child sexual abuse materials, promotion of terrorism, promotion of genocide, clear threats of violence, gender-based violence and incitement to hatred based on protected characteristics.

- **27.2** There is often a tension between national laws and international human rights standards, which poses a challenge for any attempt to define

---

Five Any restrictions upon content being posted, or content removed, should be defined in law, have a legitimate purpose and be necessary in a democratic society and be applied proportionally.
global guidance on regulation. Should illegal content be defined in a jurisdiction that may violate international human rights law, the platform will be expected to report on how it responds to such requests.

28. In addition, platforms should report on systems they have in place that would help enable them to identify the following, while protecting the right to privacy and anonymity:

- **28.1** multiple accounts created by the same source;
- **28.2** [false or inauthentic behaviours that promote misinformation or other damaging content];
- **28.3** [synthetic content designed to mislead or create a false impression (unless clearly identified as such for artistic or creative purposes)];
- **28.4** the use of automated programmes designed to mimic users (bots);
- **28.5** content created by accounts registered by state actors or otherwise credibly determined to be state-affiliated.

29. Platforms should then have explicit processes to deal with these phenomena, whether it is to label and identify such content or accounts, while protecting the right to privacy and anonymity, to restrict the virality of content arising from such accounts, or to flag with a warning that the nature of this content could be misleading or otherwise problematic (similar to splash page warnings now provided by banks before allowing transactions). The purpose for these provisions is to allow users to understand the nature and origin of questionable content and accounts and allow them to make their own judgement as to their provenance.

30. Finally, platforms should notify users when their content is removed or subject to content moderation. This will allow users to understand the reasons that action on their content was taken, the method used (algorithmic or after human review) and under which platform rules action was taken. Also they should have processes in place that permit users to appeal such decisions.

31. **An enabling environment.** This guidance recognises the difficulties of identifying content that is potentially damaging to democracy and human rights. For example, separating misinformation) from disinformation (is complex, bringing with it dangers

---

6 This item was highlighted by different stakeholders consulted as requiring extra safeguards for the balancing of rights, therefore we would appreciate detailed comments on how to address this point.

7 This item was highlighted by different stakeholders consulted as requiring extra safeguards for the balancing of rights, therefore we would appreciate detailed comments on how to address this point.
to free expression of suppressing content legitimately protected under international human rights law.

31.1 Platforms should show what they do to provide an enabling environment that facilitates expression, that challenges false or misleading information, warns of offline consequences to speech that might be dangerous (e.g., hate speech) or simply flags different perspectives or opinions.

31.2 Where possible, users should be given the ability to control the content that is suggested to them - platforms should consider ways to encourage users’ control over the selection of the content to be displayed because of a search and/or in news feeds. The availability to choose between content recommendation systems that display different sources and different viewpoints around trending topics should be made available to users in online platforms.

32. User reporting. In supporting freedom of expression and the availability of accurate and reliable information in the public sphere, it is critical to empower users of digital platform services. In this regard, all, companies, governments, civil society organisations and academic institutions have a roll to play. Companies in particular, in addition to the platform providing information about its policies accessible in a digestible format and in all relevant languages, it should show how it allows users to report potential abuses of the policies, whether that be the unnecessary removal of content, the presence of violent or threatening content, or of any other content which is in breach of the policies. Where possible, users should have access to a platform representative in their own country.

32.1 The user reporting system should give high priority to content that is threatening or intimidatory, particularly to groups with protected characteristics, ensuring a rapid response⁸ and, if necessary, by providing specific means of filing the report. This is particularly important when it comes to gendered online violence and harassment. A pre-set template would allow the aggregation of similar complaints that would help identify systemic failings on the platform. At the same time, this guidance recognises that much of this will depend upon local and regional contexts.

32.2 There should also be an effective user complaints mechanism to allow users meaningful opportunities to raise issues of concern. This should include a clear and

---

⁸ One mean could be through an escalation channel for the most egregious threats.
understandable reporting channel for complaints and users should be notified about the result of their appeal.

32.3 There will clearly be issues of scale for platforms with large numbers of users. In such circumstances, platforms may need to deploy automated systems to process and record complaints and the regulatory system will review the operation of these systems.

33. Content that potentially damages democracy and human rights, including mis- and disinformation and hate speech

33.1 For platforms and independent regulators, attempts to identify potentially damaging content that is not manifestly illegal can be a significant challenge, as most freedom of expression legal standards emphasise the importance of context and intent – saying the same words in different contexts and in different ways can have very different legal implications. And sometimes apparently legal speech which constitutes disinformation can be deployed with the intent of causing severe harm. Different opinions and viewpoints on the potential damage posed by content will arrive at very different solutions. This is made even more difficult by the sheer volume of content uploaded continually across all platforms, which can feasibly be managed, at least in the first instance, mainly automatically.

33.2 Platforms should say how they define and respond to a wider set of damaging content through a systematic risk assessment. The regulatory system should assess if platforms are consistently applying their systems and processes to effectively enforce their own standards (including the protection of legitimate speech) which should be aligned to international human rights standards.

33.3 This guidance recognises the formidable challenge of identifying damaging speech which may be legal in one context but damaging in another. For example, it is important to distinguish between content promoting hatred directed at women, children, youth, LGBTTIQ, indigenous groups, people with disability and vulnerable communities and content that is simply offensive to a particular group of people in a particular context.
33.4 Platforms should demonstrate how they would respond to potentially damaging speech – either by providing alternative reliable information, flagging concerns about the quality of this information, curbing its virality or any other means. Content removal or de-platforming of users should be considered only when the intensity, and severity of content that has the intention to harm a group or individual occurs. The platform should also be explicit about whether it partners with outside organizations or experts to help it make these kinds of decisions particularly in countries or regions where the platform itself has little local knowledge.

33.5 Platforms should show whether they apply specific protection measures to particular groups. If they do, these measures might include risk assessment and mitigation processes or the creation of specific products that enable these specific groups to actively participate online.

34. Media and information literacy. Platforms should set out the resources they make available to improve media and information literacy, including digital literacy about their own products and services, for their users. There should be a specific focus inside the company on how to improve the digital literacy of its users with thought given to this in all product development teams. The platform should be reflecting on how any product or service impacts upon user behaviour and not just on the aim of user acquisition or engagement.

34.1 Platforms should implement specific media and information literacy measures for women, children, youth and indigenous groups.

35. Election integrity

35.1 Digital platform services should have a specific risk assessment process for any election event and should engage with the election’s administrator/regulator (and relevant civil society groups), if one such exists, prior to and during an election to

---

9 Several media platforms have instituted “disputed news” tags that warn readers and viewers about contentious content.

10 The frequency of occurrence over a given period and their range.

11 Its scale, scope or irremediability. Scale means the gravity of the impact on the human right(s). Scope means the number of individuals that are or could be affected. Irremediability means the ease or otherwise with which those impacted could be restored to their prior enjoyment of the right(s).

12 Although this guidance is aiming to highlight the asks of the regulatory systems towards the platforms, it is important to underline that UNESCO has a series of recommendations for governments regarding media and information literacy policies, which must be implemented by relevant authorities, particularly in the education sector.
establish a means of communication if concerns are raised by the administrator or by users/voters. Within the assessment, they should review whether political advertising products, policies, or practices arbitrarily limit the ability of candidates or parties to deliver their messages.

35.2 Digital platform services that accept political advertising should ensure in their terms of service that to accept the advert, the funding and the political entity are identified by those that place the adverts.

35.3 The platform should retain these advertisements and all the relevant information on funding in a publicly accessible library online. Political advertisements which refer to issues rather than parties or candidates should be scrutinised to ensure they are consistent with the overarching policies of the platform in relation to hate speech or speech targeting people with protected characteristics.

35.4 Digital platform services should adopt transparency measures regarding the use and impact that the automated tools they use, although not necessarily the specific codes with which they operate, may have in practice, including the extent to which such tools affect the data collection, targeted advertising, and the disclosure, classification, and/or removal of content, especially election-related content.

36. Major events – Digital platform services should have risk assessments and mitigation policies in place for “major events” crises such as conflicts, wars, natural disasters, health emergencies, and sudden world events where mis- or disinformation and hate speech are likely to increase and where their impact is likely to be rapid and severe.

37. Language and accessibility. Digital platform services operate globally, and the main language of many such platforms is English. There are over 7,000 languages spoken in the world today, though many are spoken only by small groups of people.\textsuperscript{13} It is critical if regulation is to be effective that users can operate in a language that they understand. Setting a reasonable expectation for which languages platforms should be able to operate in will depend upon the scale, reach, and sensitivity of the service. For global platforms, it would be reasonable to suggest that users can contact them either in one of the six UN languages\textsuperscript{14} or in one of the 10 languages spoken by more than 200 million people.\textsuperscript{15} Automated language translators, while they have their

\textsuperscript{13} \url{https://www.ethnologue.com/guides/how-many-languages}. Around 40% of languages are spoken by fewer than 1,000 speakers.

\textsuperscript{14} Arabic, Chinese, English, French, Russian and Spanish

\textsuperscript{15} English, Chinese, Hindi, Spanish, French, Arabic, Bengali, Russian, Portuguese and Urdu.
limitations, can be deployed to increase the number of languages available. Platforms may wish to ensure the provision of information in additional languages during election events, perhaps by increasing the capacity of moderation in local languages during such events. Consideration should also be given to persons with disabilities, and the ways in which they can interact with, and make complaints in relation to, the platform.

37.1 It is recognised that this signifies an important shift in the way platforms operate, as English is a predominant international language. Nevertheless, as the major force in global communication, platforms must recognise their responsibility to allow people to communicate effectively if they are to be accountable to them.

38. Data access. Platforms should provide stable access, wherever safe and practicable, to non-personal data and anonymised data. Access should be provided to data that is aggregated, or manifestly made public data for research purposes through automated means such as application programming interfaces (APIs) or other open and accessible technical solutions allowing its analysis. They should provide access to data necessary to undertake research on content that is potentially damaging to democracy and human rights and support good faith research that involve their services. There need to be safeguards with providing data access that ensures the protection of privacy and respect of commercial confidentiality. For platforms to build reliable interfaces for data access, there will need to be alignment among regulators that can determine what is useful, proportionate and reasonable for research and regulatory purposes.

Section Three- The independent regulatory system

39. There are vastly different types of bodies involved in online regulation throughout the world. These range from existing broadcast and media regulators who may be asked to take on the role of regulating content online, to newly established dedicated internet content regulators, or general communications regulators given an extended remit. There may also be overlap in some states with advertising or election bodies, or with information commissioners or national human rights institutions. Some regulators may exist independently of government and others might be constituted as government agencies. It is therefore difficult to set out detailed guidance when regulation can take so many varying forms and potentially involve so many agencies.

40. Nevertheless, in whatever form regulation operates, it will constitute what this guidance calls a “regulatory system” of some kind. The guidance below is therefore meant to be generally applicable to any system of regulation, however established,
however varied. Of course, this guidance recognises that this approach, if adopted, could imply significant changes to the way regulation operates in some Member States.

Constitution

41. Any regulatory system, whether a single body or multiple overlapping bodies, charged with managing online content (overseeing systems and processes) needs to be independent and free from economic or political pressures or any external influences. Its members should be appointed through an independent merit-based process of appointment, overseen by an oversight body (which could be the legislature or an independent board/boards). Its members should not seek or take instructions from any external body, whether public authority or private actor.

42. The dismissal of members of the regulatory system should be based on clear criteria and must also be subjected to a thorough process, guaranteeing that their dismissal is not a result of any political or economic pressures.

43. The members of the regulatory body/bodies should make public any possible conflict of interest.

44. The regulatory system must have sufficient funding to carry out its responsibilities effectively.

45. The regulatory system should make a regular report to an oversight body on its findings and will be accountable to it. The regulatory system should also hold periodic multi-stakeholder consultations on their operation.

Powers

46. Regulation will set the overarching goals for platforms to safeguard information as a public good, empowering and protecting users (particularly vulnerable users such as children or minorities) and specifying expectations as to how the stated goals should be fulfilled. It should not make judgements about individual pieces of content but will focus upon the systems and processes used by the platforms.

47. While the guidance is developed for those platforms whose services have the largest size and reach, minimum safety requirements should be applied to all platform service companies regardless of size.
48. In-scope digital platform services will be required to report regularly on how they are achieving the goals, though regulators may commission off-cycle reports if there are exigent circumstances, such as a sudden information crisis (such as that brought about by the COVID-19 pandemic) or a specific event which creates vulnerabilities (e.g., elections, protests, etc.).

49. The regulatory system will have the power to call in any digital platform service deemed not to be complying with its own policies or failing to protect users and, after discussion, may recommend a specific set of measures to address the identified failings. Any such judgement should be evidence-based; the platform should have an opportunity to make representations and/or appeal against a decision of non-compliance; and the regulatory system should be required to publish and consult on enforcement guidelines and follow due process before directing a platform to implement specific measures. Failing to comply with this stage could lead to penalties which are proportionate, dissuasive, and effective (but excluding personal criminal liability).

50. It will have the power to commission a special investigation or review by an independent third party if there are serious concerns about the operation or approach of any platform or an emerging technology.

51. It is expected that illegal content will be removed solely in the jurisdiction where it is illegal.

52. One option as an additional protection for users is for there to be an ombudsman for complaints about platforms. While in the first instance, complaints should be made directly to the digital platform service itself, in the event of no or an inadequate response, the user could go directly to the ombudsman. This may result in an unmanageable workload, and an alternative for digital platform services with large volumes of content is for them to have independent complaints/appeals/redress processes, which the regulatory system can then evaluate.

53. Given the likely volume of complaints, the regulatory system will be expected to prioritise those complaints that demonstrate importance and relevance, systemic failings and/or substantial user harm. In this event the relevant regulator will have the power to intervene and require action, including on an interim/urgent basis if necessary.

---

16 However, it is important to recognise that no systems and processes will be 100% precise in identifying illegal content (at least not without disproportionate intrusion and monitoring). Therefore, it should not automatically be a breach of the regulations if illegal content is found on the service, unless it can be shown that the platform knew of it and failed to report it, or if the relevant systems and processes can be shown to be inadequate. Moreover, identification of illegal content should be interpreted consistently with international human rights law to avoid unjustified restrictions on freedom of expression.

17 The scope of complaints would be limited to a failure to comply with its regulatory duties, rather than as an additional appeal mechanism where users are unhappy with specific decisions.
Review of the regulatory system

54. There will be provision for a periodic independent review of the regulatory system, conducted by a respected third-party reporting directly to the legislature and subsequent consideration by the legislature.

55. Any part of the regulatory system should act only within the law in respect of these powers, respecting fundamental human rights - including the rights to privacy and to freedom of expression. It will be subject to review in the courts if it were believed it had exceeded its powers or acted in a biased, irrational or disproportionate way.

56. Decisions on eventual limitations of specific content should be taken by an independent judicial system, following a due process of law.