



# GPD Comments on UNESCO's "Guidance on Regulating Digital Platforms: A Multistakeholder Approach"

Global Partners Digital submission  
January 2023

## About Global Partners Digital

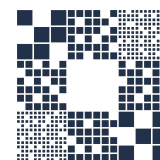
Global Partners Digital is a social purpose company dedicated to fostering a digital environment underpinned by human rights.

## Introduction

We welcome the opportunity to provide comments to UNESCO as part of its consultation on "Guidance on Regulating Digital Platforms: a multistakeholder approach". This guidance document seeks to provide high-level guidance to states and other relevant stakeholders that are considering how to regulate online content. This effort is part of a broader UNESCO initiative "Internet for Trust: Regulating Digital Platforms for Information as a Public Good".

GPD recognises the desire of UNESCO to use this document as a means of guiding regulators, governments, legislatures and companies around the world when they are developing, enforcing or implementing regulation to manage content online, as well as providing a tool for civil society to hold governments and companies accountable and advocate for a regulatory system that safeguards freedom of expression. Based on our analysis, however, we believe that there are aspects of the document that should be modified to better reflect a rights based-approach and mitigate risks to human rights, with a particular emphasis on freedom of expression and privacy.

We set out these comments and a series of recommendations below.



## Response

### General Comments

#### 1. Streamlining and coordinating efforts

We applaud the inclusion of reference to the Santa Clara Principles in the guidance and believe that **the document would be further strengthened by including additional references to relevant frameworks and guidance**. This includes those from regional institutions and multi stakeholder initiatives such as the Council of Europe, the Freedom Online Coalition (FOC), the Forum on Information & Democracy, or the Global Network Initiative (GNI). These institutional partners, bodies and initiatives have, in many cases, already developed detailed guidance on how to approach online platform regulation and content moderation in a rights respecting manner.<sup>1</sup>

It is imperative that UNESCO acknowledge their contributions and sources of best practice within the text to ensure that the guidance builds on and expands upon existing work. This should take place within Sections 1, 2 and 3, as opposed to simply signposting within the glossary.

In addition, going forward, **the initiative would benefit from coordinating with existing bodies and initiatives** that are already committed to developing standards for the regulation of online content from a human rights perspective. In particular, UNESCO should seek to coordinate more closely with institutional partners at the United Nations, particularly the OHCHR, in order to leverage complementary efforts such as those under the B-Tech Project. We further recommend that UNESCO leverage the existing expertise and outputs of relevant Special Procedures of the Human Rights Council.

#### *Recommendation 1:*

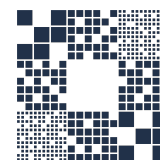
We recommend that the text of the guidance document provide more explicit references to existing relevant frameworks and guidance. This should take place within Sections 1, 2 and 3, as opposed to simply signposting within the glossary.

#### *Recommendation 2:*

We recommend that UNESCO seek to coordinate with existing bodies and initiatives that are already committed to developing standards for the regulation of

---

<sup>1</sup> See, Forum on Information and Democracy, “How to end infodemics” (2020) and “Accountability regimes for social networks and users” (2022); Global Network Initiative, “Content Regulation and Human Rights: Analysis and Recommendations” (2020); and Freedom Online Coalition “Joint Statement on Spread of Disinformation Online”.



online content from a human rights perspective. This should involve outreach efforts with key partners, bodies and initiatives to reinforce the legitimacy of the guidance document as a UN-based tool which builds on and expands upon other efforts at the global, regional and national levels.

## 2. Ensuring robust stakeholder engagement modalities

So far, UNESCO has made important efforts to provide opportunities for stakeholders to shape the guiding document through various consultation rounds. As UNESCO considers next steps, **we encourage further efforts be made to strengthen the stakeholder engagement modalities** to ensure that the process is sufficiently open, inclusive and transparent.

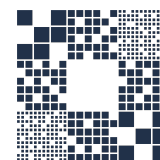
As a first step, given the complex nature of the topic and the wide array of stakeholders affected by it, **we recommend the consultation process for the development of the guidance document is extended** beyond the current timeframe which is set to culminate at the Global Conference in February.

In addition, the following specific suggestions could be considered:

- Setting out a plan for consultation that includes both virtual and in-person consultations with a wide variety of stakeholders, including states, companies and civil society organisations – particularly those with varying mandates and perspectives, including those from the Global South and those representing marginalised communities;
- Ensuring that the plan for consultation provides enough time to allow stakeholders to meaningfully engage, gather evidence and share information and concerns relating to the guidance document;
- Making the consultation plan public and accessible to all relevant stakeholders;
- Providing an online platform where all submissions can be made public and documentation on agreed procedures and outreach;
- Widely publicising the online platform and calling for comments to capture a broad and diverse group of stakeholders (in all UN languages);
- Hosting consultations (virtual and/or in-person) beyond the Global Conference, ensuring broad geographic reach, including by organising regional events in parallel to existing forums and meetings to reduce costs;
- Providing adequate mechanisms and support, including financial support, to ensure the inclusion of the perspectives of all relevant stakeholders, particularly those who are not able to attend in-person events.

### *Recommendation 3:*

We recommend that UNESCO extend the consultation process for the development of this guidance document beyond the Global Conference in February and develop



a more robust plan for stakeholder engagement to ensure that participation in the process is open, inclusive and transparent.

*Recommendation 4:*

We propose the future consultation plan should:

- Set out a plan for consultation that includes both virtual and in-person consultations with all relevant stakeholders;
- Ensure that the plan for consultation provides enough time to allow stakeholders to meaningfully engage;
- Make the consultation plan public and accessible to all relevant stakeholders;
- Provide an online platform where all submissions can be made public;
- Widely publicise the online platform and call for comments to capture a broad and diverse group of stakeholders (in all UN languages);
- Host consultations (virtual and/or in-person) beyond the Global Conference, ensuring broad geographic and stakeholder reach;
- Provide adequate mechanisms and support, including financial support, to ensure the inclusion of the perspectives of all relevant stakeholders.

## **Specific Comments**

### **Section One - The Goal of Regulation**

We are pleased to see a human rights-based approach as one of the five key principles set out at the end of Section 1. However, we believe the guidance would still benefit from additional references to relevant international human rights law and standards, and further clarity on the terminology used throughout the guidance.

**Section 1:** Section 1 would benefit from more robust and specific reference to existing international human rights law and standards. We understand that the requirements for a permissible restriction of expression under Article 19(3) of the ICCPR are provided for in Section 2, but recommend that this language also be reflected in the beginning of the guidance, ideally in the introduction or the first few paragraphs of Section 1.

*Recommendation 5:*

We recommend that the introduction and the first few paragraphs of Section 1 include more robust and specific reference to existing international human rights



law and standards, including the requirements for a permissible restriction of expression under Article 19(3) of the ICCPR.

**Paragraphs 18:** We would appreciate clarification on the last sentence of paragraph 18, which states that “the regulator should also be able to scrutinise the scope of requests to ensure adequate balance between illegality and freedom of expression”. International human rights law does not set out a balancing act between illegality vs freedom of expression, but instead sets out clear standards for when governments must restrict expression under Article 20 of the ICCPR, and when they may permissibly restrict expression under Article 19(3) of the ICCPR. We recommend that these provisions, as well as accompanying guidance (e.g. General Comment No. 34 of the HRC) be included here and that “balance” be removed.

*Recommendation 6:*

We recommend that paragraph 18 remove reference to “adequate balance between illegality and freedom of expression” and suggest that relevant international human rights law and standards on freedom of expression be provided within the text.

**Paragraphs 16, 19 and 33:** We are concerned with the inclusion of legal but harmful content in paragraphs 16, 19 and 33.

- We are unsure if these phrases (“content that potentially damages human rights and democracy” and “potentially harmful content”) are being used interchangeably and recommend that clear and uniform terms be used throughout the guidance.
- We do not believe there is sufficient clarity on the potentially broad scope of these terms, particularly as laid out in paragraph 19, which includes a wide range of potentially illegal and legal but harmful content (e.g. incitement to violence and hostility to human rights defenders), as well as in paragraph 33. Despite noting that such content must be dealt with “while protecting international standards of freedom of expression”, these paragraphs could include additional safeguards to limit subjective interpretation and mitigate risks of governments interpreting the guidance as providing permission to prohibit broad categories of legal content online.

*Recommendation 7:*

We recommend that the guidance use clear and uniform terminology when making reference to legal but harmful content. We further suggest that additional safeguards and clarification be provided to limit subjective interpretation and mitigate risks of governments interpreting the guidance as providing permission to prohibit broad categories of legal content online.



**Paragraph 19:** We welcome that the second sentence of paragraph 19 stresses the need for any regulation that protects human rights to be “flexible enough to adapt to new or changing circumstances” or in other words, to be future-proofed. But we are unsure how it would guide legislatures or other relevant actors in the development or revising of legislation, for example, by issuing secondary legislation or codes of practice. We recommend that paragraph 19 provide further information as to what future-proof legislation means in practice, as well as new language on the benefits of primary legislation, as opposed to secondary legislation, which may not receive the same level of legislative scrutiny when compared to primary legislation.

*Recommendation 8:*

We recommend that paragraph 19 provide further information as to what future-proof legislation means in practice, as well as new language on the benefits of primary legislation, as opposed to secondary legislation, which may not receive the same level of legislative scrutiny when compared to primary legislation.

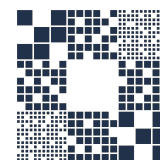
## Section Two - Fulfilling the Goal

**Paragraph 23.1:** We welcome that Section 2 sets out the responsibilities of governments that are considering legislation to regulate processes impacting content moderation and/or curation. We are pleased that paragraph 23 specifically sets out the obligations of governments to respect, protect and promote human rights in this respect. However, while paragraph 23.1 makes reference to freedom of expression, the right to information, equality and non-discrimination, it does not include the right to effective remedy as enshrined in Article 2(3) of the ICCPR. We recommend including the right to an effective remedy here as it supports several of the key principles of the guidance, including that it is human rights-based and that platforms are accountable. It may also be beneficial to replace “users” with individuals to capture a broader set of persons impacted by platforms content and content policies.

*Recommendation 9:*

We recommend including the right to an effective remedy in paragraph 23.1 as it supports several of the key principles of the guidance, including that it is human rights-based and that platforms are accountable. It may also be beneficial to replace “users” with individuals to capture a broader set of persons impacted by platforms content and content policies.

**Paragraph 23.6:** Paragraph 23.6 states that governments should “not impose indirect restrictions to companies (for example, internet shutdowns) for alleged or potential breaches of regulation,” but it is not clear what is meant by indirect restrictions and the scope of these measures besides internet shutdowns. Newer forms of online



platform regulation have increasingly provided disproportionate forms of sanctions, including throttling speeds, blocking access, or subjecting staff to criminal penalties. These types of sanctions are likely to be unnecessary, disproportionate and pose risks to freedom of expression. We recommend that paragraph 23.6 instead stress the need for regulation to provide for a range of proportionate sanctions – whether civil or criminal – for lack of compliance with any respective duties or responsibilities.

*Recommendation 10 :*

We recommend that paragraph 23.6 more clearly stress the need for regulation to provide for a range of proportionate sanctions – whether civil or criminal – for lack of compliance with any respective duties or responsibilities. It should either remove or clarify what is meant by “indirect measures”.

**Paragraphs 25.6, 30, 32.3:** We are pleased that the guidance document seeks to tackle the use of automated processes for content moderation. However, we believe that paragraphs could more effectively address the risks associated with using automated processes for content moderation and their impacts on human rights, particularly freedom of expression and non-discrimination.

Automated tools perform poorly on recognising novel forms of content and considering context, and introduce risks of algorithmic bias, which is inevitable in most automated processes due to discriminatory training data or the types of value judgements built into AI systems. Simply requiring platforms to notify users when automated tools are used or to implement appeals processes for such decisions is insufficient, and as such, we suggest that paragraphs 25, 32 and other relevant paragraphs in Section 2 better reflect the risks associated with use of such tools and propose more specific and robust mechanisms for mitigating these risks. For example, the guidance could recommend that regulations should require regular independent review of any machine learning tools utilised for content moderation for their precision and accuracy across different content types, languages and contexts, regular independent assessment of the impacts of automated content moderation tools on human rights, and well as requiring practical amendments to automated tools and providing users with enhanced transparency and functionality.

*Recommendation 11:*

We recommend that paragraphs 25, 30 and 32 better reflect the risks associated with use of such tools and propose more specific and robust mechanisms for mitigating these risks.

**Paragraph 26:** We appreciate that paragraph 26 makes references to the UN Guiding Principles on Business and Human Rights (UNGPs), which is a framework already familiar to many tech companies and provides a useful way for online platforms to



assess their actions with respect to their business practices. However, the UNGPs is only mentioned briefly in paragraph 26 and solely in relation to content management policies, as opposed to the policies of digital platforms more broadly (including algorithmic recommendation systems, advertising practices, processing of user data, etc.). It would be beneficial to lift this language to apply throughout Section 2.

*Recommendation 12:*

We recommend that the reference to the UN Guiding Principles on Business and Human Rights (UNGPs) in paragraph 26 be more clearly set out in Section 2 in order to apply to all relevant paragraphs and not solely content management policies.

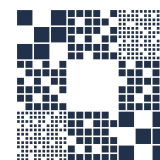
**Paragraph 27:** We also believe that there is room for improvement in paragraph 27 with respect to the risks to freedom of expression and other rights associated with platforms as private entities becoming the arbitrators of what is legal or not. As currently provided in paragraph 27, “Platforms should, in policy and practice .. ensure that, at 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”. But there is no accompanying language on the risks of outsourcing complex legal adjudications – such as determining whether speech should be classed as incitement to hatred – to private platforms without sufficient transparency and accountability, which may lead to the over removal of permissible content. We understand that the guidance focuses on systems and processes, but we recommend that the document better reflect these risks.

*Recommendation 13:*

We recommend that paragraph 27 set out the risks of outsourcing complex legal adjudications – such as determining whether speech should be classed as incitement to hatred – to private platforms without sufficient transparency and accountability, which may lead to the over removal of permissible content.

**Paragraph 32:** Paragraph 32 on user reporting should reflect a more comprehensive approach to reporting mechanisms in support of freedom of expression and the right to an effective remedy. We recommend that the guidance expand its language here to include both users and non-users. This is important because non-users of services are often those directly impacted by content shared on a service, particularly where that service has a smaller user base and lax content moderation policies or processes. Moreover, we recommend that paragraph 32 make reference to children and that it stresses the need for effective user complaints mechanisms that are accessible to younger individuals.





*Recommendation 14:*

We recommend that paragraph 32 reflect a more comprehensive approach to reporting mechanisms in support of freedom of expression and right to an effective remedy. This should include both users and non-users, as well as stressing the need for effective user complaints mechanisms that are accessible for children.

**Paragraphs 32, 33, 46:** We believe that there is inconsistent language used throughout the guidance document with respect to marginalised groups and individuals. For example, paragraph 32.1 provides that user reporting systems should be given high priority content that is threatening or intimidating, particularly to groups with protected characteristics. Whereas 33.3 specifies that 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. Paragraph 46 speaks about the need to protect users (particularly vulnerable users such as children or minorities). It is not clear why certain groups are deemed important for particular requirements but not others. We recommend that the guidance be revised throughout to more carefully consider its references and inclusion of different groups and to refer to these groups in a consistent way throughout the document.

*Recommendation 15:*

We recommend that the guidance, including paragraphs 32, 33 and 46, use uniform and clear language when referring to marginalised or vulnerable groups or individuals.

### Section Three - The Independent Regulatory System

We believe that Section 3 is quite strong and provides useful guidance on the constitution, powers and review of the regulatory system.

However, we believe that the subsection on constitution would benefit from an additional requirement that the regulatory body have (or have access to) sufficient expertise in international human rights law. Moreover, we believe paragraph 45 could be strengthened by carving out a role for civil society in the review of the regulatory system. For example, “the regulatory system should also hold periodic multi-stakeholder consultation on their operation, which should involve the compulsory consultation of civil society organisations”.

*Recommendation 16:*



We recommend that the subsection on the constitution include an additional requirement that the regulatory body have (or have access to) sufficient expertise in international human rights law.

*Recommendation 17:*

We recommend that paragraph 45 provide a role for civil society in the review of the regulatory system. For example, “the regulatory system should also hold periodic multi-stakeholder consultation on their operation, which should involve the compulsory consultation of civil society organisations”.