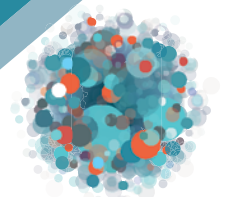


# CONTENT REGULATION AND HUMAN RIGHTS

ANALYSIS AND  
RECOMMENDATIONS

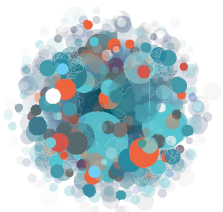


POLICY BRIEF

GLOBAL  
NETWORK  
INITIATIVE

# CONTENT REGULATION AND HUMAN RIGHTS

## POLICY BRIEF



GLOBAL  
NETWORK  
INITIATIVE

# TABLE OF CONTENTS

## **1. EXECUTIVE SUMMARY 4**

---

→ OUR STARTING POINT 4

→ WHAT WE FOUND 5

→ WHAT WE RECOMMEND 5

---

## **2. CONTEXT 7**

---

## **3. EVALUATING RECENT MEASURES UNDER THE INTERNATIONAL HUMAN RIGHTS FRAMEWORK 8**

→ LEGALITY 8

→ LEGITIMACY 16

→ NECESSITY 17

→ PRIVACY 26

---

## **4. CONCLUSION 29**

---

APPENDIX A - RECOMMENDATIONS 30

APPENDIX B - INDEX OF LAWS 31

# 1. EXECUTIVE SUMMARY

## OUR STARTING POINT

---

Principles of good governance and human rights impel governments to understand and address public and private harms within their jurisdiction. Since policymakers and regulators around the world are increasingly concerned about various forms of online content and conduct, it is no surprise that many are considering how different forms of state action may help or hinder efforts to address those concerns.

The multistakeholder Global Network Initiative (GNI) reviewed more than twenty recent<sup>1</sup> governmental initiatives that claim to address various forms of online harm related to user-generated content — a practice we refer to broadly as “content regulation.” We focused on proposals that could shift existing responsibilities and incentives related to user-generated content. Our analysis illustrates the ways that **good governance and human rights principles provide time-tested guidance** for how laws, regulations, and policy actions can be most appropriately and effectively designed and carried out. Because content regulation is primarily focused on and likely to impact digital communication and content, we use international human rights principles related to freedom of expression and privacy as our primary lens.

These historically validated human rights principles can help lawmakers find creative and appropriate ways to engage stakeholders, design fit-for-purpose regulations, and mitigate unintended consequences. **Governments that actively place human rights at the forefront of their deliberations and designs are not only less likely to infringe on their own hallowed commitments, they can also achieve more informed and effective outcomes**, balancing public and private responsibilities, designing appropriate incentives, enhancing trust, and fostering innovation.

---

1. This brief includes analysis of many, but not all of the content regulation initiatives that GNI members have identified as noteworthy up until the brief went to print in mid-September 2020.

## WHAT WE FOUND

---

Although there are important differences between the various content regulation efforts examined in this brief, many share certain key characteristics. By definition, such initiatives alter the balance of responsibilities in the information and communications technology (ICT) ecosystem, introducing a degree of **legal uncertainty**, which can shift user understanding and expectations, disrupt information value-chains, and risk unsettling the playing field for ICT companies of all sizes and business models. While this is not, in and of itself, a reason to refrain from regulation, few governments have demonstrated sufficient efforts to fully understand the social and economic impacts of such disruption.

Many content regulation efforts also **require or otherwise strongly incentivize intermediaries to further rely on automated filtering systems** to proactively identify illegal or otherwise inappropriate content or conduct, notwithstanding the fact that such systems, in their current state, may result in over-removal and increase the risk of self-censorship.<sup>2</sup> Beyond this, a number of the initiatives reviewed would **force intermediaries to rapidly adjudicate the legality or permissibility of third-party content on their services**, creating unintended consequences and complicated implications for the rule of law, democratic process, accountability, and redress.

In addition, some of these initiatives implicitly or explicitly **require tracing and/or attribution of content, raising significant privacy concerns**. Lawmakers have been particularly challenged in their efforts to regulate private messaging services, many of which feature strong end-to-end encryption, which protects user content and security but can make content moderation by intermediaries challenging.

Finally, a number of these efforts **apply more broadly than necessary**. Some seek not only to address illegal expression more effectively, but also to regulate legal but harmful content. Others, whether explicitly or due to unclear or vague language, apply to companies of varying sizes across various layers of the ICT sector, unnecessarily creating the potential for liability among companies that are not well positioned to effectively or proportionately address content. And yet others assert the authority to regulate content extraterritorially, and even globally, heedless of the implications for users' rights in other jurisdictions and international comity.

## WHAT WE RECOMMEND<sup>3</sup>

---

In order to identify effective and proportionate approaches to content regulation, **public authorities need to recognize that the ICT sector is perpetually evolving**. Services that facilitate sharing of user-generated content differ in important ways, and

---

2. See, Natasha Duarte and Emma Llansó, "Mixed Message? The Limits of Automated Social Media Content Analysis," November, 28, 2017, <https://cdt.org/insights/mixed-messages-the-limits-of-automated-social-media-content-analysis/>.

3. Note: A complete set of recommendations can be found in Appendix A at the end of this paper.

the ICT sector features an ecosystem of interrelated components upon which multiple industries, initiatives, and possibilities depend. This complexity counsels careful consideration of what state actions are most appropriate and narrowly tailored to address which specific challenges. Lawmakers must be clear about the priorities that inform their efforts and open to diverse approaches to achieving them.

Fortunately, many actors agree on the need to address legitimate public policy concerns around harmful content and conduct online while respecting human rights. Many ICT companies have come to recognize the value of clear, publicly defined laws and obligations, while civil society actors continue to provide constructive and often prescient advice drawn from the real-world experiences of the most vulnerable and marginalized communities. **Processes for legislative deliberation should therefore be open and non-adversarial, drawing on broad expertise** to ensure results are well thought out and evidence based. Unelected regulatory or oversight bodies should also prioritize transparency and consultation with diverse constituencies.

Furthermore, while governments can and should learn from each other, they should also recognize that **there are no off-the-shelf solutions to complex regulatory challenges**. Governments need to take the time to understand and consider actions that are consistent with international human rights obligations and appropriate and proportionate for their jurisdiction.

Although it is clear that ICT companies have responsibilities and important roles to play in addressing online harms, **lawmakers should resist the temptation to shift all legal liability from those generating illegal content to intermediaries**. Not only can this misalign company priorities, incentivizing invasive monitoring and over-removal of content, it often does little to address the underlying drivers of harmful content and conduct.

Laws and regulations governing the ICT sector should also be targeted and narrowly framed. Lawmakers should pay careful attention to the ways laws and regulations will impact companies with different business models, seeking to **foster a diversity of digital services and avoid raising barriers to entry**.

For all of these reasons, **when the decision is made to regulate, governments should build strong transparency, remedy, and accountability measures into their efforts**. Such measures allow policymakers and other relevant stakeholders to understand if content regulations are working as intended, including assessing the activities and effectiveness of unelected oversight or enforcement bodies. Where experience demonstrates that content regulation is not working as intended, governments must recognize and expeditiously rectify any issues that emerge.

## 2. CONTEXT

Governments are increasingly considering ways to regulate content and conduct in the digital sphere. These efforts range from expanding traditional forms of offline censorship, through tried-and-true legal demands to intermediaries, to deploying government-ordered network disruptions.<sup>4</sup> Governments are also trying out “new school,” less-direct, and non-legal approaches, including pressuring ICT intermediaries to expand the range of content prohibited under their community standards, as well as their enforcement of those standards — often under the (implicit or explicit) threat of legislation or regulation.<sup>5</sup>

**This paper addresses a distinct but related trend: the introduction, application, and interpretation of legal and regulatory measures to compel or otherwise pressure ICT-sector intermediaries to police user content pursuant to domestic law and/or community standards, under threat of significant fines, service disruption, criminal or civil liability, and/or costly court proceedings.**<sup>6</sup> The recently expressed willingness by some governments to impose intermediary liability for user content stands in contrast to earlier regulatory approaches, such as Section 230 of the Communications Decency Act in the United States and the e-Commerce Directive in the EU, which create clear safe harbors for ICT intermediaries against liability for user-generated content.

GNI acknowledges the legitimate and important role that governments can play in regulating the ICT sector, accepts that regulation can help address illegal content, and recognizes that many of these initiatives attempt to enhance transparency around and accountability for content-related decisions by private companies, which are worthy objectives. Good intentions notwithstanding, these laws must also be consistent with and uphold international human rights law. Carefully considered approaches, narrowly tailored requirements that are targeted at services that pose the greatest risk of harm, relevant exceptions, and appropriate safeguards can help ensure such consistency.

As part of our commitment to serve as a multistakeholder voice for freedom of expression and privacy, GNI has publicly participated in deliberations regarding various legal and regulatory initiatives, in an effort to provide proactive, constructive, and timely recommendations.<sup>7</sup> This paper builds on that work and seeks to provide an evergreen resource for all those grappling in good faith with how to envision content regulation that enables and enhances human rights.

---

4. Global Network Initiative, *Disconnected: A Human Rights Based Approach to Network Disruptions*, May 2018, <https://globalnetworkinitiative.org/wp-content/uploads/2018/06/Disconnected-Report-Network-Disruptions.pdf>

5. Jack Balkin, “Free Speech is a Triangle,” 118 Colum. L. Rev. 2011 (2018), [https://columbialawreview.org/content/free-speech-is-a-triangle/#:~:text=Balkin\\*&text=The%20vision%20of%20free%20expression,to%20protect%20free%20expression%20today.&text=The%20twenty%2Dfirst%2Dcentury%20model,of%20it%20as%20a%20triangle](https://columbialawreview.org/content/free-speech-is-a-triangle/#:~:text=Balkin*&text=The%20vision%20of%20free%20expression,to%20protect%20free%20expression%20today.&text=The%20twenty%2Dfirst%2Dcentury%20model,of%20it%20as%20a%20triangle). Relatedly, some governments have developed official initiatives to monitor and flag content under companies’ community standards. See, Jason Pielemeier and Chris Sheehy, “Understanding the Human Rights Risks Associated with Internet Referral Units,” Feb. 25, 2019, <https://medium.com/global-network-initiative-collection/understanding-the-human-rights-risks-associated-with-internet-referral-units-by-jason-pielemeier-b0b3feeb95c9>.

6. Far from being a “wild west,” many countries have long applied laws to ICT companies’ handling of user content, including via regulatory and court interpretations that apply older laws or regulations to newer, digital services. In this paper, GNI focuses on more recent efforts that attempt to establish and clarify new legal responsibilities related to user-generated content.

7. See, e.g., Submission to U.K. Online Harms White Paper Consultation, July 2019, <https://globalnetworkinitiative.org/wp-content/uploads/2019/07/GNI-Submission-UK-Online-Harms-Consultation.pdf>; GNI, Letter to Australian Government, April 2019, <https://globalnetworkinitiative.org/wp-content/uploads/2019/04/GNI-Concerns-Australia-Bill-April-2019.pdf>; GNI, Submission to Indian Ministry of Industry and Technology, January 2019, <https://globalnetworkinitiative.org/wp-content/uploads/2019/01/GNI-Submission-Meity-Intermediary-Guidelines-Amendments.pdf>; GNI, Statement on Proposed EU Regulation on Terrorist Content, January 2019, <https://globalnetworkinitiative.org/wp-content/uploads/2019/01/GNI-Statement-Proposed-EU-Regulation-on-Terrorist-Content.pdf>, and Statement on LIBE Committee Amendments to EU Regulation on Terrorist Content, April 2019, <https://globalnetworkinitiative.org/wp-content/uploads/2019/04/GNI-LIBE-Terreg-Statement.pdf>; GNI, Statement on German NetzDG Legislation, April 2017, <https://globalnetworkinitiative.org/proposed-german-legislation-threatens-free-expression-around-the-world/>.

### 3. EVALUATING RECENT MEASURES UNDER THE INTERNATIONAL HUMAN RIGHTS FRAMEWORK

While the Universal Declaration of Human Rights — the first global human rights treaty — and many of its progeny were developed before the advent of mobile telephony and the Internet, their respective provisions on freedom of expression all share language emphasizing that this right must apply “through any media” and “regardless of frontiers.” The UN Human Rights Committee has since clarified that, under the International Covenant on Civil and Political Rights (ICCPR), “[a]ny restrictions on the operation of websites, blogs or any other internet-based, electronic or other such information dissemination system, including systems to support such communication, such as internet service providers or search engines, are only permissible to the extent that they are compatible with [Article 19] paragraph 3.”<sup>8</sup> More recently, the UN Guiding Principles on Business and Human Rights (“UNGPs”) stipulate that, “[i]n meeting their duty to protect [human rights], states should . . . [e]nsure that . . . laws and policies governing the creation and ongoing operation of business enterprises . . . do not constrain but enable business respect for human rights.”<sup>9</sup> In addition, some states have articulated additional commitments regarding the ways in which they will protect digital rights.<sup>10</sup>

Article 19(3) of the ICCPR sets out a framework describing the limited circumstances in which states may legitimately restrict freedom of expression. This framework is replicated, with some distinctions, across a variety of international and regional treaties. The framework consists of three interrelated principles: **legality, legitimacy, and necessity**.

## LEGALITY

The principle of legality establishes two requirements for the regulation of expression. First, it requires that **restrictions on freedom of expression must be provided by public laws “formulated with sufficient precision to enable an**

8. Human Rights Committee, General Comment No 34, CCPR/C/GC/34, 12 September 2011, para 43 (hereinafter General Comment 34); see also, Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, Rep. of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, Frank La Rue, Human Rights Council, U.N. Doc. A/HRC/14/23 (Apr. 20, 2010); U.N. Special Rapporteur on Freedom of Opinion and Expression, OSCE Representative on Freedom of the Media, OAS Special Rapporteur on Freedom of Expression & ACHPR Special Rapporteur on Freedom of Expression and Access to Information, International Mechanisms for Promoting Freedom of Expression, Joint Declaration on Freedom of Expression and the Internet (June 1, 2011), available at <http://www.osce.org/fom/78309>.

9. UN Guiding Principles on Business and Human Rights (hereinafter UNGPs), available at: [https://www.ohchr.org/documents/publications/GuidingprinciplesBusinesshr\\_eN.pdf](https://www.ohchr.org/documents/publications/GuidingprinciplesBusinesshr_eN.pdf).

10. For example, the 32 countries that are members of the Freedom Online Coalition have articulated a wide range of commitments. See, <https://freedomonlinecoalition.com/>.



**individual to regulate his or her conduct accordingly.”**<sup>11</sup> These laws must be validly enacted and publicly available, so that individuals are effectively put on notice as to what conduct and content is prohibited. Second, **they “must provide sufficient guidance to those charged with their execution to enable them to ascertain what sorts of expression are properly restricted and what sorts are not.”**<sup>12</sup> This latter concern is especially important in the context of laws that outsource enforcement of speech regulation to private actors of varying sizes, business models, and capacities. As the Human Rights Committee explained in General Comment 34, laws regulating speech “may not confer unfettered discretion for the restriction of freedom of expression on those charged with its execution.”<sup>13</sup>

The legality requirement is essential to mitigating the chilling effect of ambiguous laws. Whenever expression is prohibited, the mere possibility of being accused of violating the law or being subject to costly court proceedings can cause individuals not to express themselves and companies to refrain from facilitating expression. For individuals to be able to understand and navigate these boundaries, **restrictions on expression must clearly and precisely define both what is prohibited and who can be held responsible** for failing to enforce the prohibition. Any vagueness or ambiguity is likely to cause individuals to refrain from exercising their rights and lead intermediaries to be overly aggressive in censoring expression for fear of being held in violation of the law.

## OPEN, PARTICIPATORY LAW/RULE MAKING

---

**When states consider particular forms of online content sufficiently harmful as to require regulation, they should be deliberated upon openly and defined through legislation, consistent with domestic law.** To the extent that substantial rulemaking authority and discretion is delegated to independent bodies, the scope of the regulator’s duties and corresponding legal safeguards must be set out in primary legislation. States must create robust oversight and accountability mechanisms to ensure that those bodies act pursuant to the public interest and intervene in markets in a non-arbitrary way, consistent with the state’s obligations.

The U.K. Online Harms White Paper (*White Paper*) proposed to allow an independent regulatory body<sup>14</sup> to develop codes outlining the systems and processes companies should have in place to address particular categories of harms, including legal content, in lieu of having such decisions made by a democratically elected parliament.<sup>15</sup> While delegation of rulemaking to independent bodies is not, in and of itself, problematic, and such bodies can be an important part of broader efforts

---

11. General Comment 34, FN 5 *supra*, para 25

12. General Comment 34, FN 5 *supra*, para 25

13. General Comment 34, FN 5 *supra*, para 25

14. DCMS has subsequently clarified that “it is minded to appoint Ofcom,” the UK’s existing communications regulatory, to this role. *Id.*

15. The Australian Government’s discussion paper on “Online Safety Legislative Reform,” while focusing on content that is currently illegal under domestic law, also appears to empower the Minister of Communications and Arts to expand this to include additional types of content without having to go back to Parliament. Australian Ministry of Communications and Arts, “Online Safety Legislative Reform: Discussion Paper,” December 2019, p. 41, <https://www.communications.gov.au/have-your-say/consultation-new-online-safety-act> (hereinafter, Online Safety Discussion Paper).

by public authorities in this space, transferring the power to control, including by significantly restricting, aspects of protected speech to unelected bodies may create the potential for democratic decision making processes and methods of accountability to be circumvented. Provisions that explicitly direct such regulatory bodies to uphold freedom of expression may help ease these concerns somewhat.

As GNI noted in our submission to the U.K. Government on the *White Paper*, “[w]e are concerned that the regulator will not only be charged with developing and enforcing codes for the 23 categories of ‘online harms’ identified in the *White Paper*, but also with enforcing the ‘duty of care,’ ‘even where a specific code does not exist,’ and unilaterally identifying and defining new categories of harm beyond this ‘initial list.’ The degree of discretion and lack of predictability such a system would create is highly concerning.”

The Eliminating Abusive and Rampant Neglect of Interactive Technologies Act of 2020 (*EARN IT Act*), which was recently introduced in the U.S. Senate, would delegate the authority to recommend “best practices” related to the prohibition of child sexual abuse material (CSAM) to a commission led by the Attorney General.<sup>16</sup> Although those “best practices” will not hold the force of law, it is unclear if and when public input and debate would be permitted in this somewhat bespoke process and what bearing those “best practices” will have on courts’ interpretation of the law or future legislation.

These concerns are exacerbated when laws require intermediaries to meet such regulatory codes or standards in order to achieve protections from liability for third-party content, which are critical to protecting freedom of expression. For example, other legislation recently introduced in the U.S. Congress would condition Section 230 liability exemptions on “good faith” or “objectively reasonable” content moderation standards without specifying what exactly those are.<sup>17</sup> Separately, an executive order issued by President Trump calls for liability protections to be removed where social media companies restrict access to content in “bad faith.”<sup>18</sup>

By contrast, Ireland’s Online Safety and Media Regulation Bill (*Online Safety Bill*),<sup>19</sup> which seeks to transpose the EU Audiovisual Media Services Directive, sets out a national framework to regulate harmful content online that takes a much narrower approach by establishing four clear categories of harmful content (barring one

16. Eliminating Abusive and Rampant Neglect of Interactive Technologies Act of 2020 (the “EARN IT Act”) S.\_\_\_\_, 116th Cong. (2020) (hereinafter, the EARN IT Act), <https://www.congress.gov/bill/116th-congress/senate-bill/3398/>;

17. Limiting Section 230 Immunity to Good Samaritans Act S.\_\_\_\_, 116th Cong. (2020), <https://www.hawley.senate.gov/sites/default/files/2020-06/Limiting-Section-230-Immunity-to-Good-Samaritans-Act.pdf>; Stopping Big Tech’s Censorship Act S. 4062, 116th Cong. (2020), <https://www.congress.gov/bill/116th-congress/senate-bill/4062/text>; “The Online Freedom and Viewpoint Diversity Act,” S.\_\_\_\_, 116th Cong. (2020), <https://www.commerce.senate.gov/services/files/94D0F3C6-B927-46D2-A75C-17C78D0D92AA>. For a discussion of some of these proposed bills, see Matt Bailey, “Three and a Half Ways to Fix the Internet,” 1 July 2020, <https://pen.org/three-and-a-half-ways-not-to-fix-the-internet/>. See also, Daphne Keller, “Trump-Backed Rules for Online Speech: No to Porn, Yes to Election Disinformation and Hate Speech,” Oct. 1, 2020, <https://slate.com/technology/2020/10/cda-section-230-graham-doj-reform-content-moderation.html>

18. U.S. White House, Executive Order on Preventing Online Censorship, May 28, 2020, <https://www.whitehouse.gov/presidential-actions/executive-order-preventing-online-censorship/>.

19. Ireland General Scheme Online Safety and Media Regulation Bill, Oct. 1, 2020, [https://www.dccae.gov.ie/en-ie/communications/legislation/Documents/154/General\\_Scheme\\_Online\\_Safety\\_Media\\_Regulation\\_Bill.pdf](https://www.dccae.gov.ie/en-ie/communications/legislation/Documents/154/General_Scheme_Online_Safety_Media_Regulation_Bill.pdf).

concerning cyberbullying that could use further clarification) and requiring the government's approval of any new categories proposed by the Online Safety Commissioner. The regulator would also designate services for statutory regulation, based on an assessment of risk.

## DEFINITIONAL CLARITY

---

**Regardless of who makes the rules and how, without clear definitions, it is difficult for “an individual to regulate his or her conduct accordingly,”** creating a risk of self-censorship. Lack of clarity also makes it difficult, if not impossible, for companies who may be held responsible for such content “to ascertain what sorts of expression are properly restricted and what sorts are not,” creating a serious risk of over-removal. In addition, it can create particular compliance challenges in the ICT sector, where companies' operations are often transnational.

Unfortunately, many recent proposals to regulate online content fail to provide sufficient guidance to users or to the platforms charged with their enforcement. For example, the U.K. *White Paper* outlines an approach that would require companies to address both “harms with clear definitions” and “harms with less clear definitions,” raising significant questions about whether any regulator could “provide sufficient guidance” to enable users and the companies charged with enforcing the latter set of restrictions “to ascertain what sorts of expression are properly restricted.” Perhaps more concerning, the *White Paper* proposes to open online intermediaries to enforcement action for failing to properly address content posted by their users that is not prohibited by law. This creates clear tension with the principle of legality.

As the French government's report on “Creating a French framework to make social media platforms more accountable” (*Interim Mission Report*) points out, “the scope of content that is not ‘manifestly unlawful’ (grey zone) varies according to geography and does not easily lend itself to European or international harmonization...”<sup>20</sup> The potential for conflicts of laws stemming from the regulation of online content was recently underscored by a Court of Justice of the European Union ruling holding that EU Member States may require platforms to remove illegal content, as well as future “equivalent” content, globally.<sup>21</sup>

Other content regulation efforts use vague and reductive definitions that would be very difficult to enforce in a manner perceived as fair and non-discriminatory. For instance, the draft EU regulation on “preventing the dissemination of terrorist content online” (*Terror Regulation*)<sup>22</sup> and the proposed amendments to the Indian Intermediary

---

20. Government of France, “Creating a French framework to make social media platforms more accountable: Acting in France with a European vision,” May 2019, [https://www.numerique.gouv.fr/uploads/Regulation-of-social-networks\\_Mission-report\\_ENG.pdf](https://www.numerique.gouv.fr/uploads/Regulation-of-social-networks_Mission-report_ENG.pdf) (hereinafter, French Interim Mission Report).

21. European Court of Justice, Judgement of the Third Chamber, *Eva Glawischnig-Piesczek v. Facebook*, ECLI:EU:C:2019:821, October 3, 2019, <http://curia.europa.eu/juris/document/document.jsf?text=&docid=218621&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=1965965>

22. Regulation of the European Parliament and the Council on Preventing the Dissemination of Terrorist Content Online, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018PC0640&from=EN>

Guidelines (*Intermediary Guidelines Amendments*)<sup>23</sup> focus on notoriously difficult to interpret terms like “terrorism,” without providing adequate definitional clarity.<sup>24</sup>

Similarly, the government of Pakistan’s Citizens Protection (*Rules Against Online Harm*) 2020 define “extremism” to include not only violent, but also “vocal or active opposition to fundamental values of the state of Pakistan...”<sup>25</sup> Meanwhile, the stated purpose of Singapore’s new Protection from Online Falsehoods and Manipulation Act (*POFMA*) is “to prevent the communication of false statements of fact,” which it defines circularly as “a statement [that] is false or misleading.” The likelihood for subjective interpretation and application of this particular law is enhanced by the fact that *any* government minister can order a company to act if they personally consider a statement to be false. Tanzania’s Electronic and Postal Communications (Online Content) Regulations implement an even broader list of prohibited content without clear definitions, including content that “promotes annoyance,” leads to “confusion about the economy,” “uses bad language,” and “harms the prestige or status” of Tanzania, among several other vague categories.<sup>26</sup>

Even where laws refer to categories of expression that are already illegal, the range of categories is often quite broad. For instance, there are 22 provisions under the German Act to Improve Enforcement of the Law in Social Networks (*NetzDG*),<sup>27</sup> 14 under the French law on “Countering online hatred” (*Avia’s law*),<sup>28</sup> and 13 in the U.K. *White Paper*. Notwithstanding the fact that these provisions are already enshrined in law, a good number of them remain poorly defined and understood under domestic law.<sup>29</sup>

23. Indian Ministry of Electronics and Information Technology, “The Information Technology [Intermediary Guidelines (Amendment) Rules] 2018,” December 24, 2018, available at: [https://meity.gov.in/writereaddata/files/Draft\\_Intermediary\\_Amendment\\_24122018.pdf](https://meity.gov.in/writereaddata/files/Draft_Intermediary_Amendment_24122018.pdf) (hereinafter, *Intermediary Guidelines Amendments*).

24. See, GNI, Statement on Europe’s Proposed Regulation on Preventing the Dissemination of Terrorist Content Online, January 2019, <https://globalnetworkinitiative.org/gni-statement-draft-eu-regulation-terrorist-content/>

25. Government of Pakistan, Ministry of Information Technology and Telecommunication, Citizen Protection (Against Online Harms) Rules, 2020, January 21, 2020, [https://moitt.gov.pk/SiteImage/Misc/files/CP%20\(Against%20Online%20Harm\)%20Rules%2c%202020.pdf](https://moitt.gov.pk/SiteImage/Misc/files/CP%20(Against%20Online%20Harm)%20Rules%2c%202020.pdf) (hereinafter, *Pakistan Rules Against Online Harm*). See also, Asif Shazad, “Pakistan’s government approves new social media rules, opponents cry foul,” Reuters (Feb. 13, 2020) (quoting Nighat Dad, Executive Director of Digital Rights Foundation, saying: “The worrying part for me is that the definition around extremism, religion or culture is so wide and ambiguous and that means they have these unfettered power to call any online content illegal or extremist or anti-state.”)

26. The Electronic and Postal Communications (Online Content) Regulations, 2020, available at: [https://www.tcra.go.tz/document/The%20Electronic%20and%20Postal%20Communications%20\(Online%20Content\)%20Regulations,%202020](https://www.tcra.go.tz/document/The%20Electronic%20and%20Postal%20Communications%20(Online%20Content)%20Regulations,%202020)

27. Act to Improve Enforcement of the Law in Social Networks, 2017, [https://www.bmjbv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG\\_engl.pdf?\\_\\_blob=publicationFile&v=2](https://www.bmjbv.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG_engl.pdf?__blob=publicationFile&v=2) (hereinafter *NetzDG*)

28. European Commission, Notification Detail: Law aimed at combating hate content on the Internet (France), August 21, 2019, <https://ec.europa.eu/growth/tools-databases/tris/en/search/?trisaction=search.detail&year=2019&num=412> (draft text in English) (hereinafter *Avia’s Law*).

29. See, e.g., Letter from UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression to Government of Germany, available at <https://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf> (expressing concern that “A prohibition on the dissemination of information based on vague and ambiguous criteria, such as ‘insult’ or ‘defamation’, is incompatible with article 19 of the ICCPR. The list of violations is broad, and includes violations that do not demand the same level of protection.”); U.K. Law Commission, “Abusive and Offensive Online Communications: A Scoping Report,” [https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2018/10/6\\_5039\\_LC\\_Online\\_Comms\\_Report\\_FINAL\\_291018\\_WEB.pdf](https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2018/10/6_5039_LC_Online_Comms_Report_FINAL_291018_WEB.pdf) (noting that many of the criminal provisions in the area of abusive and offensive online communications - such as those relating to harassment and disclosing private sexual photographs or films (“revenge pornography”) - are unclear, ambiguous or overly complex.)

## RISKS OF PRIVATIZED ENFORCEMENT

Putting aside concerns about the compatibility of underlying domestic laws with the principle of legality and the delegation of rulemaking to administrative bodies, **the outsourcing of enforcement of these laws to private companies without appropriate guidance on interpretation and application raises separate, significant legality concerns** (this also raises serious issues with regard to “necessity,” which are explored in the “necessity” section below). Determinations that certain content or conduct are illegal should be made by independent adjudicators, and where such responsibilities are passed along to private entities, it is imperative that they be shielded from liability for such decisions.<sup>30</sup> While adjudications by judicial authorities tend to be publicly transparent and easily accessible, content-related decisions by intermediaries are often not made public and related moderation processes are less well understood. Notwithstanding this, many recent proposals delegate adjudication of restricted categories of expression to companies under the threat of heavy penalties for poor compliance, without accompanying measures related to transparency, accountability, or remedy.

More appropriate in this regard may be some of the approaches in France’s “Law against manipulation of information”<sup>31</sup> (*Fake News Law*). While it is an imperfect legislative approach,<sup>32</sup> it does not create new categories of prohibited content, instead focusing on a single, defined type of content already prohibited under existing law (thus avoiding creating new, poorly understood categories, or prohibiting speech that is not illegal). The law also establishes clear limiting criteria for when such content should be restricted in the online context (i.e., manifestly illegal content that is viral and leading to a disturbance of the peace or compromise of an election), and it leaves the determination of when those criteria are met to a judge, via an expedited procedure, instead of outsourcing determinations to private platforms.<sup>33</sup>

In contrast, *Avia’s law* would have effectively removed judges from content adjudication, and otherwise departed from the broader framework recommended in the French *Interim Mission Report*,<sup>34</sup> notwithstanding the request from the European Commission to adopt more proportionate measures.<sup>35</sup> Partly for these reasons, the French Constitutional Council struck down several key parts of *Avia’s law* in July 2020,

---

30. See, Emma Llansó, “The Digital Services Act: An Opportunity to Build Human Rights Safeguards into Notice and Action,” Aug. 17, 2020, <https://medium.com/global-network-initiative-collection/the-dsa-an-opportunity-to-build-human-rights-safeguards-into-notice-and-action-by-emma-llans%C3%B3-e0487397646f>.

31. Loi organique n° 2018-1201 du 22 décembre 2018 relative à la lutte contre la manipulation de l’information [Organic Law No. 2018-1201 of 22 December 2018 Regarding the Fight Against Information Manipulation], [https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=3EA914DFE69980E3FBB01324A666B5D1.tplgfr22s\\_1?cidTexte=JORFTEXT000037847556&dateTexte=&oldAction=rechJO&categorieLien=id&idJO=JORFCONT000037847553](https://www.legifrance.gouv.fr/affichTexte.do;jsessionid=3EA914DFE69980E3FBB01324A666B5D1.tplgfr22s_1?cidTexte=JORFTEXT000037847556&dateTexte=&oldAction=rechJO&categorieLien=id&idJO=JORFCONT000037847553) (hereinafter French Fake News Law).

32. For instance, the law creates a broad requirement to act upon content flagged as “fake news” by users. See, Rim-Sarah Alouane, “Macron’s Fake News Solution Is a Problem,” Foreign Policy, May 29, 2018, <https://foreignpolicy.com/2018/05/29/macrons-fake-news-solution-is-a-problem/>

33. <https://www.gouvernement.fr/en/combating-the-manipulation-of-information>

34. *Avia Law*, FN 24 *supra*.

35. See, Chloé Berthélémy, “France’s law on hate speech gets a thumbs down,” EDRI, Dec. 4, 2019, <https://edri.org/frances-law-on-hate-speech-gets-thumbs-down/>



finding that they amounted to “an attack on the exercise of the freedom of expression and communication that is not necessary, appropriate, and proportional.”<sup>36</sup>

In countries where lawmakers can permissively rely on private companies to adjudicate prohibited expression, governments should ensure that such companies have sufficient guidance to adjudicate the content in a wide variety of contexts. In this sense, provisions, such as the one provided for in the *NetzDG*, which allow companies to refer particular decisions to an independent, expert-staffed “self-regulation institution,”<sup>37</sup> may help to address some legality concerns. Another logical approach would allow for the referral and/or appeal of particularly difficult decisions to independent judicial or quasi-judicial public authorities.

## SAFEGUARDS

**Regardless of who is ultimately charged with adjudication, any laws that restrict expression should provide for sufficient transparency, oversight, and remedy** so as to avoid “confer[ring] unfettered discretion for the restriction of freedom of expression on those charged with its execution.”<sup>38</sup> This is particularly important when adjudication is outsourced to private platforms that are not otherwise democratically accountable. Here, the French *Interim Mission Report*’s focus on “transparency obligations” and provisions such as those in the EU *Terror Regulation* regarding transparency requirements for both “hosting providers” and “competent authorities” are helpful. The U.K. Government’s “Online Harms White Paper – Initial Consultation Response,” (*Consultation Response*) also acknowledges the importance of these principles, hopefully foreshadowing greater attention to transparency and remedy in its forthcoming legislation.<sup>39</sup> Meanwhile, the U.S. Platform Accountability and Consumer Transparency Act (*PACT Act*), which was introduced this year in the U.S. Senate, requires standardized, quarterly content moderation transparency reports, focusing on complaint and appeal mechanisms.<sup>40</sup>

**In addition, language requiring human review of content flagged by automated tools, as well as provisions ensuring the right to an effective remedy in response to content restrictions, are also very welcome** (for more on the importance of these qualities, see the “necessity” section).

36. Decision n° 2020-801 DC du 18 Juin 2020: Loi visant à lutter contre les contenus haineux sur internet [Decision No. 2020-801 DC of June 18, 2020: law to combat hate speech on the Internet] <https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm>

37. *NetzDG*, FN 23 *supra*.

38. General Comment 34, FN 5 *supra*, para 25; see also Molly K. Land, “Against Privatized Censorship: Proposals for Responsible Delegation,” *Va. J. Int’l L.* (forthcoming 2020), at 46 (“Delegated censorship unaccompanied by safeguards is unlawful under international human rights law because it will inevitably be overbroad.”).

39. UK Initial Consultation Response, Feb. 12, 2020, <https://www.gov.uk/government/consultations/online-harms-white-paper/public-feedback/online-harms-white-paper-initial-consultation-response>.

40. Platform Accountability and Consumer Transparency Act (“PACT Act,”) S.\_\_\_\_, 116th Cong. (2020) (hereinafter, the PACT Act), <https://www.schatz.senate.gov/imo/media/doc/OLL20612.pdf>

## Legality – Do's and Don'ts

DO	DON'T
✓ Conduct law/rulemaking openly, in a participatory manner that allows for diverse and expert inputs, based on empirical analysis, and accompanied by impact assessments.	⊘ Implement laws or regulations through closed, non-participatory processes.
✓ To the extent substantial rulemaking authority and discretion is delegated to independent bodies, create robust oversight and accountability mechanisms to ensure that such bodies act pursuant to the public interest and consistent with international obligations.	⊘ Use vague and reductive definitions that would incentivize over-removal and be difficult to enforce in a manner that is perceived as fair and non-discriminatory.
✓ Ensure public laws are “formulated with sufficient precision to enable an individual to regulate his or her conduct accordingly.”	⊘ Allow categories of content to be prohibited without transparent and participatory processes for deliberation.
✓ Establish clear limiting criteria and leave the determination of when those criteria are met to a judge.	⊘ Outsource adjudication to private companies without appropriate clarity on interpretation and application of the law.
✓ Clearly and precisely define what is prohibited, as well as who can be held responsible for failing to enforce the prohibition.	⊘ Shift legal liability from authors to intermediaries for illegal content generated by users.
✓ Set clear expectations for responsible company action with regard to reports of illegal content.	⊘ Neglect the importance of transparency and accountability.
✓ Ensure the law requires transparency, oversight, and remedy so as to avoid “confer[ring] unfettered discretion for the restriction of freedom of expression on those charged with its execution.”	

## LEGITIMACY

The principle of legitimacy insists that laws restricting expression can only be justified to achieve specific, enumerated purposes. Article 19(3) of the ICCPR describes these as “respect for the rights or reputations of others” and “the protection of national security or of public order, or of public health or morals.”

While international law gives states significant room to determine what sorts of activities can be understood to sufficiently impact these purposes so as to justify restrictions, that discretion is not unlimited. Indeed, the Human Rights Committee has held that, unlike with some other rights covered in the ICCPR and some regional treaties, “the scope of [freedom of expression] is not to be assessed by reference to a ‘margin of appreciation.’”<sup>41</sup> The Committee has also been clear that “[a]ny such limitations must be understood in the light of universality of human rights and the principle of non-discrimination.”<sup>42</sup>

In considering if and how specific legislation meets this “legitimate purpose” test, it is **important to consider that the right to freedom of expression is broad in its scope, encompassing “even expression that may be regarded as deeply offensive.”**<sup>43</sup> When restricting broad categories of content, lawmakers must consider the likelihood that speech that is controversial but protected will be impacted, and take steps to avoid or mitigate those impacts. While the U.K. *Consultation Response* clarifies that the forthcoming online harms regulation “will establish differentiated expectations on companies for illegal content and activity, versus conduct that may not be legal but has the potential to cause harm, such as online bullying, intimidation in public life, or self-harm and suicide imagery,”<sup>44</sup> it remains unclear how this will work in practice.

Furthermore, **the prohibition of such content when it is expressed in digital form, despite the fact that it may not be illegal in *physical* analog publications or public spaces, raises questions about the potentially discriminatory impact** of such rights. It also creates tension with the well-established maxim, set out in numerous consensus United Nations resolutions, that “the same rights that people have offline, must also be protected online, in particular freedom of expression.”<sup>45</sup> The wisdom behind these calls for equal treatment (and against discrimination based on medium) rests in part on the recognition that such inconsistencies are likely to be exploited by regimes and actors who do not respect democratic norms.

41. General Comment 34, FN 5 *supra*, para 36 (citing communication No. 511/1992, Ilmari Lämsmä, et al. v. Finland, Views adopted on 14 October 1993); but see, [ECtHRs decisions holding that there is a margin of appreciation under the European Convention].

42. General Comment 34, FN 5 *supra*, para 26; see also, Joint Declaration on Freedom of Expression and countering violent extremism adopted by the UN Special Rapporteur on Freedom of Opinion and Expression, the OSCE Representative on Freedom of the Media, the OAS Special Rapporteur on Freedom of Expression and the ACHPR Special Rapporteur on Freedom of Expression and Access to Information, adopted on 4 May 2016, p. 2 (“Restrictions on freedom of expression must also respect the prohibition of discrimination, both on their face and in their application.”).

43. General Comment 34, FN5 *supra*, para 11.

44. “UK Initial Consultation Response,” FN 27 *supra*.

45. See, e.g., UN Human Right Council Resolution A/HRC/38/L.10/Rev.1, The promotion, protection and enjoyment of human rights on the Internet, July 2018, <https://undocs.org/en/A/HRC/38/L.10/Rev.1>



## Legitimacy Do's & Don'ts

DO	DON'T
✓ Ensure that content that is prohibited falls within one of the enumerated “legitimate purposes” in ICCPR Art. 19(3).	⊘ Prohibit categories of content that cannot be justified under one of the enumerated “legitimate purposes.”
✓ Ensure that controversial and offensive content is not prohibited simply because it makes certain audiences uncomfortable.	⊘ Create a compliance regime that incentivizes individuals to self-censor or intermediaries to over-remove content.
✓ Ensure that content that is allowed in analog contexts is also permitted in digital form.	⊘ Allow the law to discriminate against content based on medium.

## NECESSITY

The principle of necessity requires that states seeking to restrict expression to **“demonstrate in specific and individualized fashion the precise nature of the threat, and the necessity and proportionality of the specific action taken,** in particular by establishing a direct and immediate connection between the expression and the threat.”<sup>46</sup> Proportionality, in this context, means that any restrictive law, as well as actions of administrative and judicial authorities applying the law, **“must be appropriate to achieve their protective function; they must be the least intrusive instrument amongst those which might achieve their protective function; they must be proportionate to the interest to be protected.”**<sup>47</sup> In the words of the French *Interim Mission Report*, “any state intervention must be strictly necessary, proportionate and transparent whenever it affects public freedoms that are as important as the freedom of expression and freedom of communication.”<sup>48</sup>

### TARGETING RULES APPROPRIATELY

To be “appropriate to achieve their protective function,” there must be a connection between the regulatory action and the purpose to be achieved. Most, if not all, of the recent laws and proposals reviewed here express legitimate concerns about the real and potential harms that certain content or behavior can cause. However, **very few provide sufficient empirical support or argumentative clarity to establish “a direct and immediate connection between the expression and the threat” or “the necessity and proportionality of the specific action taken.”**

46. General Comment 34, FN 5 *supra*, para 35

47. General Comment 34, FN 5 *supra*, para 34

48. French *Interim Mission Report*, FN 17 *supra*.

For instance, although *NetzDG* was proposed in response to public concerns and a government-requested study on the specific issue of “hate speech,” the law delegates enforcement, under threat of liability, of 22 categories of expression — many of which having little to do with hate speech. Meanwhile, the U.K. *White Paper* articulates a series of concerns regarding content and behavior on social media platforms, yet it calls for a regulatory approach that would apply to a wide range of online services, including traditional user comments hosted on digital news media outlets, and many other intermediaries not set out in the consultation response. In Kenya, the Computer Misuse and Cyber Crimes Act (*Kenya Cyber Crimes Act*) criminalizes knowingly publishing false information, not only where it “is calculated or results in panic, chaos, or violence,” but also when “it is likely to discredit the reputation of a person.”<sup>49</sup> In these instances, states have failed to provide compelling evidence as to why the breadth of each regulation is necessary to achieve its stated objectives.

Even where a law is more narrowly tailored, as is the case of Australia’s *Abhorrent Material Law*,<sup>50</sup> it is far from clear that the regulatory approach (in that case, strict liability for certain live-streamed content) is necessary and constitutes the least restrictive means to address the problem. When such laws are rushed through the legislative process or passed without substantial deliberation and debate (the *Abhorrent Material Law* was approved two days after it was introduced), states are less able to articulate and defend the necessity of the particular approaches proposed, experts and advocates are unable to provide critical review and input, lawmakers are hindered from considering more proportionate alternatives, and the resulting laws are more prone to errors and unintended consequences.<sup>51</sup> **As a general matter, lawmakers and regulators would be best served identifying and targeting those services, scenarios, and types of content that pose the greatest risk to users.**

## FINDING THE RIGHT LAYER IN THE TECHNOLOGY STACK

In the ICT context, the principle of necessity may also suggest lawmakers should focus regulation on particular services in order to minimize its impact on expression. As the table below indicates,<sup>52</sup> data flows through various layers and services in the ICT ecosystem. As a general rule, the further away a particular service is from the end user, the less visibility and granular control it has over user-generated content. **Lawmakers and regulators would be well served to carefully consider which types of private services, at which layers in the ecosystem, are most appropriately positioned to address the specific concern(s) at issue and to constrain their approaches to those best positioned to address those concerns.** The

49. Computer Misuse and Cybercrimes Act No 5 of 2018, § 22, Kenya Gazette Supplement (Special Issue) No. 60 (May 16, 2018), <http://kenyalaw.org/kl/fileadmin/pdfdownloads/Acts/ComputerMisuseandCybercrimesActNo5of2018.pdf> ,

50. Australian Criminal Code Amendment (Sharing of Abhorrent Violent Material) Bill 2019, [https://parlinfo.aph.gov.au/parlInfo/download/legislation/bills/s1201\\_aspassed/toc\\_pdf/1908121.pdf;fileType=application%2Fpdf](https://parlinfo.aph.gov.au/parlInfo/download/legislation/bills/s1201_aspassed/toc_pdf/1908121.pdf;fileType=application%2Fpdf) (hereinafter *Abhorrent Material Bill*)

51. See, David Sullivan, “The Consequences of Legislating Cyberlaw After Terrorist Attacks,” *Just Security*, April 9, 2019, <https://www.justsecurity.org/63560/the-consequences-of-legislating-cyberlaw-after-terrorist-attacks/>; Courtney Radsch, “In wake of Christchurch, tech regulation in EU and Australia risks restricting journalism,” *Committee to Protect Journalists*, June 20, 2019, <https://cpj.org/blog/2019/06/in-wake-of-christchurch-tech-regulation-in-eu-and-ph>

52. This table is illustrative and necessarily incomplete.

EU *Terror Regulation* and the Australian *Abhorrent Material law* both apply broadly to internet service providers (ISPs), telecommunications companies, and others, without explaining why or how each type of private service is necessary to the achievement of the objectives of the law.

In addition, **the consequences of a regulation may differ significantly from one type of service to another**. For instance, expecting search engines to prevent certain kinds of content (not just specific pages, as in the “right to be forgotten” context) from surfacing assumes an unrealistic ability of such services to read and assess much, if not all, of the visible, “surface” web.

SERVICE	DESCRIPTION	EXAMPLES	CONTENT VISIBILITY	CONTENT CONTROL
<b>DNS providers / registrars / registries</b>	Manage domain name servers / operate top-level domains (e.g., “.com”) / sell domain names	Verisign / Cloudflare, Google, etc.	Generally none	Can suspend entire service, but not specific content
<b>Web hosts / cloud service providers</b>	Host website content / data storage and remote computing	Bluehost, GoDaddy, etc.	Technically possible but often commercially / contractually restricted	Can suspend entire service, but not specific content
<b>Network operators / ISPs</b>	Provide infrastructure for data transmission; provide end-user mobile phone and/or internet connection as a service	Comcast, NTT, Telefonica, etc.	Limited to non-encrypted content	May be able to block access to specific unencrypted pages, but if encrypted may only be able block entire site
<b>Email services</b>	Facilitate direct one-to-one communications between users, akin to postal communications	Gmail, Hotmail, ISP-provided email services, enterprise mail, etc.	Restricted by privacy laws in the same way as postal communications. Investigatory powers typically require disclosure of content (including attachments) to authorities	Can suspend accounts
<b>Messaging services</b>	Facilitate direct, real-time messaging between one or many users	WhatsApp, Snap, LINE, etc.	Limited to non-encrypted content	Can suspend accounts
<b>Search engines</b>	Crawl and index web sites and facilitate access to content on the web	Google, Kakao, Bing (but not their syndication partners), etc.	Generally, yes	Can demote or remove links in the index they compile and from which results are served to end users
<b>Social media platforms</b>	Host and facilitate user-generated content for public or semi-public transmission	Facebook, Twitter, Tik-Tok, etc.	Generally, yes	Can remove or limit access to specific posts or content; can also restrict or otherwise moderate content

## SIZE MATTERS

---

Minimizing the impact of laws regulating speech may also require lawmakers and regulators to **consider how such requirements may impact start-ups and smaller entities, as well as any unintended impacts they could have on the pluralism of content and providers of consumer services that may be available.** The *NetzDG*'s approach, which applies specifically to for-profit social media companies and excludes platforms with less than two million German users, demonstrates one approach to addressing these considerations. In its Law on Liberty, Responsibility and Transparency on the Internet (Fake News Bill), Brazil, with a population two-and-a-half times the size of Germany, adopted the same threshold.<sup>53</sup> Meanwhile, the U.S. *PACT Act* would exempt internet infrastructure services, as well as companies with fewer than one million active users and less than \$25 million accrued revenue over the most recent 24-month period.<sup>54</sup> By contrast, the French *Interim Mission Report* recommends a differentiated, three-tiered regulatory approach based on platform *reach*.<sup>55</sup> Tailoring laws or regulations based on reach or effect makes sense, since the number of users can be a poor proxy for the actual impact a platform has on the underlying content or conduct at issue.

## PROPORTIONATE ENFORCEMENT

---

Recognizing that perfect enforcement is not feasible, governments should take care to **afford an appropriate degree of flexibility** to those private services that are covered. They should also **refrain from overly stringent enforcement and penalties**, so as to accommodate a diverse range of business models and capacities among covered businesses, as well as to foster innovative approaches to content moderation and guard against over-removal. As the *Interim Mission Report* rightly notes, “[b]y imposing an absolute standard of conformity that does not take into account the volume of the published content, the audience or the statistical nature of the processing, punitive measures risk encouraging over-moderation and thereby infringing freedom of expression...”


Overly punitive approaches are particularly likely to create chilling effects for freedom of expression. For instance, under Singapore’s *POFMA*, individuals face significant penalties (up to five years imprisonment) if convicted of issuing statements they know “or hav[e] reason to believe” are false. The Turkish “Law regarding the regulation of publications made in the Internet environment and the fight against crimes committed through these publications” (*Social Media Bill*), requires companies to maintain a representative in Turkey or be subject to fines, advertising bans, or bandwidth reductions that effectively lead to network disruption.<sup>56</sup> It is also not clear

53. Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet, Projeto do Lei 2630 de 2020, <https://legis.senado.leg.br/sdleg-getter/documento?dm=8127649&ts=1593563111041&disposition=inline>.

54. *PACT Act*, FN 33 *supra*, pp. 11.

55. French *Interim Mission Report*, FN 17 *supra*.

56. “The Law regarding the regulation of publications made in the Internet environment and the fight against crimes committed through these publications,” Law No. 7253, <https://www.tbmm.gov.tr/kanunlar/k7253.html> For more on GNI’s concerns on this bill see: <https://globalnetworkinitiative.org/concerns-turkey-social-media-bill/>



that the consideration of criminal liability for company executives under the U.K. *White Paper* or the Australian *Abhorrent Material Law* strikes an appropriate balance between deterrence and the potential for over-removal.

Also problematic are laws that would make companies liable for particular determinations they make, independent of government orders, to allow or remove certain content or conduct.<sup>57</sup> Where courts have determined that a particular piece of content is illegal, it may be reasonable to ask companies to remove repostings of the exact same content. However, deputizing plaintiffs to identify future iterations of similar content<sup>58</sup> or requiring intermediaries to identify and remove content that is not identical, but is similar or of “equivalent meaning,” raises serious concerns.<sup>59</sup>

A better approach is to set clear expectations for intermediaries and then focus enforcement efforts on identifying systematic failures in order to assist with and incentivize better approaches. For instance, the U.K. Government’s *Consultation Response* helpfully clarifies that it plans to follow a “systems based approach” with a focus on companies’ systems and processes, “rather than individual pieces of content,” so that it can “remain effective even as new harms emerge.”

Finally, it is also important that states work to **avoid unnecessary or disproportionate cross-border implications of any content restrictions required under their domestic law, which can also abridge the principle of international comity**.<sup>60</sup> For instance, the Government of Pakistan’s *Rules Against Online Harm* assert its authority to require social media companies to remove content or disable accounts “of citizens of Pakistan residing outside its territorial boundaries,”<sup>61</sup> creating likely conflicts of laws and comity challenges with countries whose laws may differ substantially on matters of freedom of expression. Similarly, the Brazilian *Fake News Bill* does not explicitly state any limitations on users’ nationality or location.

## BEWARE OF ARBITRARY TIMELINES

---

To ensure that laws restricting speech are not enforced in a disproportionate manner, states should also **avoid aggressive and arbitrary timelines for adjudication**. The *NetzDG*, the *Rules Against Online Harm*, the *Intermediary Guidelines Amendments*, the *Turkish Social Media Bill*,<sup>62</sup> and Australia’s *Online Safety Discussion Paper* all require

---

57. See, e.g., Emma Llansó, “The Digital Services Act: An Opportunity to Build Human Rights Safeguards into Notice and Action,” *supra* FN 28.

58. For example, in a May 2020 copyright case, a Swedish court recently required an internet service provider to block not only sites listed by the court in the ruling, but also additional future domain names and URLs whose main purpose is to provide access to the services of the infringing sites and which are reported to Telia by any of the film companies that brought the case. See, <https://www.domstol.se/nyheter/2020/06/internetleverantör-ska-blockera-domännamn-och-webbadresser/>.

59. See, e.g., Daphne Keller, “Facebook Filters, Fundamental Rights, and the CJEU’s Glawischnig-Piesczek Ruling,” *GRUR International*, Volume 69, Issue 6, June 2020, Pages 616–623, <https://doi.org/10.1093/grurint/ikaa047>.

60. GNI, “Statement on Domestic Cases Asserting Global Internet Jurisdiction,” Jan. 29, 2020, <https://globalnetworkinitiative.org/statement-domestic-cases-global-jurisdiction/>.

61. *Rules Against Online Harm*, FN 22, *supra*.

62. “The Law regarding the regulation of publications made in the Internet environment and the fight against crimes committed through these publications,” Law No. 7253, <https://www.tbmm.gov.tr/kanunlar/k7253.html> For more on GNI’s concerns on this bill see: <https://globalnetworkinitiative.org/concerns-turkey-social-media-bill/>



(or propose to require) removal of at least some content within twenty-four hours of notice, while *Avia's Law* and the *EU Terror Regulation* provides only one hour for some content, creating incentives for over-removal. Indeed, this stringent timeline was one of the elements that the French Constitutional Council cited in determining that *Avia's Law* did not pass the necessity and proportionality tests.<sup>63</sup> Similarly, lack of specificity as to how long blocking orders must be implemented can cause unnecessary impacts on freedom of expression.

While certain content, in certain circumstances, may merit quick and decisive moderation, by imposing strict time limits on *all* content adjudication, states may effectively hinder the ability of ICT companies to prioritize resources and make nuanced, content and circumstance-specific determinations. These time limits may also make it difficult for the author to contest the allegation (i.e., issue a counter-notice) or seek injunctive relief or other remedy. Instead, legislation that provides clear guidance as to the precise characteristics of content and circumstances that require prompt or significant action on the part of companies are likely to be both more compatible with the necessity and proportionality principles and more effective in practice.

## AVOID CONTENT BY CATCH

Perhaps even **more concerning** are efforts, such as the Australian *Abhorrent material law*, the U.K. *White Paper*, the *Intermediary Guidelines Amendments*, and an earlier version of the *EU Terror Regulation*, which would impose **requirements for preemptive filtering** of certain categories of content, thus creating tension with the general prohibition against “prior restraint/censorship” of expression.<sup>64</sup> Such filtering may be appropriate for narrow categories of content that are universally condemned and relatively easy to adjudicate, such as child sexual abuse materials. However, given the concerns raised above about how these laws could incentivize companies to err on the side of over-removal in order to avoid liability, it is hard to understand how these approaches can constitute the “least restrictive means” or be “proportionate to” the interests to be protected when applied to difficult-to-adjudicate categories such as hate speech or disinformation.<sup>65</sup>

Instead, laws that focus more broadly on articulating standards for appropriate content moderation based on traditional rule-of-law concepts, such as transparency regarding decision making, due process around content determinations, and remedy for impacted users, will be both more narrowly and appropriately tailored and more likely to provide the flexibility needed to allow platforms to adjust to changing circumstances, norms, and technology. In this regard, the French *Interim Mission Report's* proposed “accountability by design” approach is worthy of further consideration.

63. See, FN 29 *supra* and Dussart, François-Xavier, “The Balancing Act of Content Moderation in Europe: Lessons from the French “Avia Law” for the Digital Services Act,” July 27, 2020, <https://link.medium.com/hmaNOi26L9>

64. See, American Convention on Human Rights, Article 13(2); see also, Eur. Ct. H.R., Case of *The Sunday Times v. the United Kingdom*, Judgment of April 26, 1979, Application N° 6538/74; Graham Smith, “Take care with that social media duty of care,” *Cyberleagle Blog*, Oct. 19, 2018, <https://www.cyberleagle.com/2018/10/take-care-with-that-social-media-duty.html>.

65. Land, FN 20 *supra*, at 47.

## SAFEGUARDS

Requirements for transparency by intermediaries and states can help mitigate the potential for over-removal and self-censorship. Clear rules, procedures, and articulations of decisions can help “enable an individual to regulate his or her conduct accordingly,” as required by the principle of legality. They can also help establish the empirical evidence necessary to determine whether a given approach is in fact narrowly and effectively tailored to a particular legitimate purpose and whether it remains the least restrictive means for addressing its stated purpose(s). **Given how quickly technology and trends in its use evolve, governments would be well placed to pair such transparency with legislative mechanisms requiring periodic review or reauthorization.**<sup>66</sup>

Recognizing freedom of expression concerns, the *Consultation Response* encouragingly states that the U.K.’s forthcoming regulatory framework will respect platforms’ freedom “to explicitly state what content and behaviour they deem to be acceptable on their sites and enforce this consistently and transparently,” emphasizing the proposed requirement to ensure users understand “which content is and is not acceptable on different platforms, and [are able] to challenge removal of content where this occurs.”<sup>67</sup>

Another important, but unfortunately often-overlooked way to mitigate over-removal is to **ensure robust remedial mechanisms for users whose content is restricted.**<sup>68</sup> Whether this remedy is provided through state or corporate mechanisms, or both, it must be “effective,” as set out in the UN Guiding Principles. This means it must be legitimate, accessible, predictable, equitable, transparent, rights-compatible, a source of continuous learning, and in the case of “operational-level mechanisms,” based on engagement and dialogue.<sup>69</sup> In this regard, it is unfortunate that while many of the proposals discussed in this paper require companies to enhance the mechanisms by which users can flag content as either illegal or contrary to platform rules, none of them address in detail the rights of users to receive notice about or contest such notices or related adverse decisions. For instance, Australia’s *Online Safety Discussion Paper* fails to clarify whether existing mechanisms to challenge decisions of the eSafety Commissioner will extend to the new powers proposed in the forthcoming legislation. One notable exception is the “own-initiative report” on the EU Digital Services Act (DSA) from the European Parliament’s Legal Affairs Committee, which outlines some specific criteria for notice and appeal procedures to be included in the DSA.<sup>70, 71</sup>

---

66. One of these evolving technology trends is the use of Artificial Intelligence. The EU has already enshrined a ‘right to explanation’ in the GDPR, and Parliamentary draft reports on the DSA recommend integrating algorithmic accountability and transparency measures. See Heine, Ilse, “Pulling Back the Curtain on the ‘Black Box’: How the Digital Services Act Will Legislate Algorithmic Auditing,” July 22, 2020, <https://link.medium.com/ccARex3iM9>

67. UK Consultation Response, FN 27 *supra*.

68. See, Land, Molly, “Remedy and Enforcement in the Digital Services Act,” 4 August 2020, <https://medium.com/global-network-initiative-collection/remedy-and-enforcement-in-the-digital-services-act-by-professor-molly-land-a37b31b61ed5>

69. UNGPs, FN 6 *supra*, Principle 31.

70. See Articles 7 and 8: [https://www.europarl.europa.eu/doceo/document/JURI-PR-650529\\_EN.pdf](https://www.europarl.europa.eu/doceo/document/JURI-PR-650529_EN.pdf)

71. See, Wingfield, Richard, “The Digital Services Act and Online Content Regulation: A Slippery Slope for Human Rights?” 15 July 2020, <https://medium.com/global-network-initiative-collection/the-digital-services-act-and-online-content-regulation-a-slippery-slope-for-human-rights-eb3454e4285d>

Going forward, we would encourage such laws to **take a balanced approach to digital due process**. This might include requiring that users whose content has been flagged are provided notice and an opportunity to object (counter-notice), except in narrow, exceptional circumstances – such as CSAM. Given that users may fail to take advantage of such due process for a variety of reasons, laws could also require periodic audits or reviews to ensure that particular types of users (especially vulnerable groups) or categories of content are not being unduly impacted. For similar reasons, there may be value in allowing user notices to be anonymized in certain instances, as well as grouped or otherwise associated, so as to demonstrate possible unintended impacts.

Regulators providing oversight of company compliance should be mandated to examine both under- and over-removal, and to undertake efforts to discourage bad-faith removal requests, both of which were proposed in earlier versions of *Avia's law*. Also **helpful are statutory obligations on regulators to protect freedom of expression**, such as the one contained in the U.K. Communications Act of 2003<sup>72</sup> and the one proposed in the U.K. *White Paper*, which can help ensure an appropriately balanced approach to content regulation.

In addition to tailoring these laws as narrowly as possible, efforts to **carve out or provide affirmative defenses for particularly vulnerable or important groups**, such as human rights documentation groups and journalists, may help ensure that the laws are narrowly tailored to their objectives. For instance, the Australian *Abhorrent Material Law* includes various exceptions, including a carve-out for news reporting “in the public interest” by a “person working in a professional capacity as a journalist.” This represents a positive, if unfortunately narrow, acknowledgment of one unintended consequence of prohibiting even the most abhorrent content.<sup>73</sup>

72. UK Communications Act 2003, § xx available at <http://www.legislation.gov.uk/ukpga/2003/21/contents>

73. Australia Abhorrent Material Bill, § 474.37 “Defenses in respect of abhorrent violent material,” FN 45 supra.



## *Necessity – Do’s and Don’ts*

DO	DON'T
✓ Provide empirical support or argumentative clarity to establish “a direct and immediate connection between the expression and the threat.”	⊗ Be overly motivated by subjective or political rationales.
✓ Conduct careful, public, participatory deliberation to ensure laws are appropriate to achieve their protective function, are the least intrusive instrument amongst those which might achieve their protective function, and are proportionate to the interest to be protected.	⊗ Rush legislation or enact laws that have not been appropriately justified, consulted, and debated.
✓ Carefully consider which types of private services at which layers in the technology stack are most appropriately positioned to address the specific concern(s) at issue, focusing efforts on where the most significant risks/impacts occur and can be most effectively addressed.	⊗ Apply laws broadly to entire sectors, unless a justification for such broad application has been articulated.
✓ Accommodate a diverse range of business models and capacities. Consider how requirements may impact start-ups and smaller entities, as well as any unintended impacts they could have on competition policy.	⊗ Apply legal or regulatory requirements to smaller and large platforms in the same manner.
✓ Provide clear guidance as to the precise characteristics of content and circumstances that require prompt or significant action, as well as clear timelines for and/or periodic reviews of blocking/removal orders.	⊗ Impose strict timelines broadly across all potentially infringing content.
✓ Articulate standards for appropriate content moderation based on traditional rule-of-law concepts such as transparency, due process, and remedy.	⊗ Impose preemptive filtering requirements for categories of content that require nuanced assessment.
✓ Allow for variation and experimentation in approach, including “quarantining” and “downranking” of content. Provide means to guard against intentional misuse and unintentional consequences of content removal measures, including appeal and remedy mechanisms.	⊗ Make intermediaries liable for specific content moderation decisions or require prospective removals, absent clear, appropriate legal orders.
✓ Require courts to adjudicate illegal content and set clear expectations for intermediaries, focusing oversight on assisting compliance and identifying systemic failures.	⊗ Focus exclusively on content removal.
✓ Ensure robust remedial mechanisms for users whose content is restricted in order to avoid incentivizing self-censorship and over-removal. Build periodic reviews or reauthorizations into the law, in order to ensure that it remains relevant and consistent with evolving norms and technologies.	⊗ Impose overly punitive approaches.

## PRIVACY

---

Finally, many recent efforts to address online content also raise serious privacy concerns. Protections against arbitrary or unlawful interference with privacy are established in the UDHR, the ICCPR, and most regional human rights treaties. According to various UN sources, the same legality, necessity, and proportionality considerations discussed above also apply with respect to government infringements on privacy.<sup>74</sup>

## TRACKING AND TRACING

---

**Of particular concern with regard to privacy are those initiatives, such as the POFMA in Singapore and the *Intermediary Guidelines Amendments*, that would require companies to track or trace content across their platforms,** including retrospectively, and potentially across jurisdictional boundaries. Other proposals appear to require the filtering of content that has been previously deemed illegal, which would require companies to develop capabilities for proactively screening content and could limit their ability to encrypt private messages.

It is hard to see how some requirements proposed in the U.K. *White Paper*, such as the duty to “prevent known terrorist or CSEA content being made available to users,” could be complied with if companies do not proactively monitor all content on their platforms. Indeed, the *White Paper* makes clear that the duty of care will apply to private services (including email and messaging services, as well as private storage services), but does not explain how this could be done without requiring companies to access the content of private communications. It is also unclear how this could be achieved while remaining consistent with either data protection laws or the right to privacy. India’s *Intermediary Guidelines Amendments* require that companies “deploy technology based automated tools ... for proactively identifying and removing or disabling” illegal content.<sup>75</sup> Similarly, Pakistan’s *Rules Against Online Harm* requires deployment of “proactive mechanisms” to prevent livestreams of illegal content.<sup>76</sup>

## PROTECT ANONYMITY AND ENCRYPTION

---

**In addition, there has been talk of prohibiting or otherwise limiting anonymity, as well as mandating quicker or more fulsome responses to government requests for user data.** A recent meeting of the Five Eyes ministerial, for example, concluded with a communiqué calling on tech companies to “include mechanisms in the design of their encrypted products and services whereby governments, acting with appropriate legal authority, can obtain access to data in a readable and usable

---

74. See, Report of the United Nations High Commissioner for Human Rights, *The Right to Privacy in the Digital Age*, Human Rights Council, U.N. Doc. A/HRC/27/37, June 30, 2014, ¶¶ 21-27, <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G14/088/54/pdf/G1408854.pdf?OpenElement>; Report of the United Nations High Commissioner for Human Rights, *The Right to Privacy in the Digital Age*, Human Rights Council, U.N. Doc. A/HRC/39/29, Aug. 3, 2018, ¶¶ 34-38, <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/239/58/pdf/G1823958.pdf?OpenElement>.

75. *Intermediary Guidelines Amendments*, FN 17 *supra*.

76. *Rules Against Online Harm*, FN 19 *supra*.

format.”<sup>77</sup> The Indian *Intermediary Guidelines Amendments* would require companies to “enable tracing” of originators of information on their platforms, implying capabilities that may be inconsistent with existing uses of strong encryption for end-to-end communications.<sup>78</sup> Pakistan’s *Rules Against Online Harm* requires companies to provide content to authorities “in decrypted, readable and comprehensible format.”<sup>79</sup> The Brazilian *Fake News Bill* also raises concerns by requiring private messaging services to retain the chain of all communications that have been “massively forwarded.”<sup>80</sup>

As UN Special Rapporteur on Freedom of Opinion and Expression David Kaye has written, “[e]ncryption and anonymity, and the security concepts behind them, provide the privacy and security necessary for the exercise of the right to freedom of opinion and expression in the digital age. Such security may be essential for the exercise of other rights, including economic rights, privacy, due process, freedom of peaceful assembly and association, and the right to life and bodily integrity. Because of their importance to the rights to freedom of opinion and expression, restrictions on encryption and anonymity must be strictly limited according to principles of legality, necessity, proportionality and legitimacy in objective.” Encryption is also an important cybersecurity tool that companies use to protect users and data from hackers and other bad actors.

## AVOID DEPUTIZING INTERMEDIARIES

**In some instances, laws would even require *proactive* reporting of user content to law enforcement** (i.e., without specific requests). For instance, the now-rejected *Avia’s law* and recently proposed amendments to the *NetzDG* would appear to require platforms to proactively provide user data to authorities, including account passwords in the latter context.<sup>81</sup> Brazil’s *Fake News Bill* also obligates messaging services to check with mobile operators which mobile accounts have been terminated and suspend the associated accounts. While the original intent was likely to reduce the proliferation of fake accounts, the provision requires messaging services to proactively investigate user accounts, increasing the risks to privacy.

The *EARN IT Act*, introduced in the United States Congress, raises a different set of concerns. It would remove the liability shield for interactive computer services that are alleged to host third-party child sexual abuse materials, thus exposing them to a variety of potential civil and criminal actions under different federal and state laws.<sup>82</sup>

77. Joint Meeting of FCM and Quintet of Attorneys-General, <https://www.gov.uk/government/publications/five-country-ministerial-communique/joint-meeting-of-five-country-ministerial-and-quintet-of-attorneys-general-communique-london-2019>

78. Intermediary Guidelines Amendment, FN 17 *supra*. See also, Ross Schulman & Nat Meysenburg, “OTI Calls on Indian Government: Don’t Put Citizens’ Privacy at Risk,” Open Technology Institute, Jan. 13, 2020, <https://www.newamerica.org/oti/blog/oti-calls-on-indian-government-dont-put-citizens-privacy-at-risk/>

79. Rules Against Online Hate, FN 22 *supra*.

80. Katitza Rodriguez and Seth Schoen, “5 Serious Flaws in the New Brazilian ‘Fake News’ Bill that Will Undermine Human Rights [UPDATED],” EFF, June 29, 2020, <https://www.eff.org/deeplinks/2020/06/5-serious-flaws-new-brazilian-fake-news-bill-will-undermine-human-rights>.

81. Alexander Hardinghaus, Ramona Kimmich, & Sven Schonhofen, “German government introduces new bill to amend Germany’s Hate Speech Act, establishing new requirements for social networks and video-sharing platforms,” April 6, 2020, <https://www.technologylawdispatch.com/2020/04/regulatory/german-government-introduces-new-bill-to-amend-germanys-hate-speech-act-establishing-new-requirements-for-social-networks-and-video-sharing-platforms/>.

82. EARN IT Act, FN 15 *supra*. See also, Emma Llanos, “Amendments to EARN IT Act Can’t Fix the Bill’s Fundamental Flaws,” 1 July, 2020, <https://cdt.org/insights/amendments-to-earn-it-act-cant-fix-the-bills-fundamental-flaws/>.

In addition, there is a real threat that this exposure, combined with the voluntary “best practices” established by a somewhat arbitrarily empowered commission headed by the U.S. Attorney General, will turn the use of end-to-end encryption into a liability that companies won’t be able to afford.<sup>83</sup> Similarly, a proposal from the U.S. Department of Justice recommends withholding liability protections from providers who do not “maintain the ability to assist government authorities to obtain content (i.e. evidence) in a comprehensible, readable, and usable format,” thereby also threatening to destroy encrypted communications.<sup>84</sup>

## PROTECT THE FREE FLOW OF DATA

Privacy rights are also at risk with the growing number of bills and policies enforcing data localization requirements. Data localization laws force IT businesses to store data from local operations in country, as opposed to on servers elsewhere. The *Social Media Bill* and the *Rules Against Online Harms* also require social media companies to store data in Turkey and Pakistan respectively. Authorities justify these measures under the pretext of safety and security, despite the lack of evidence demonstrating that localizing data storage makes it more secure. Moreover, if data localization is required in a country with weak data protection laws and higher risk of government censorship and surveillance, these laws can have negative impacts on citizens’ fundamental rights and civil liberties.

### *Privacy – Do’s and Don’ts*

DO	DON'T
✓ Think creatively about how to facilitate accountability for those who violate the law, while continuing to strengthen privacy protections for all.	⊗ Require tracing or tracking of content that can only be implemented by weakening or violating users’ right to privacy.
✓ Recognize that anonymity and pseudo-anonymity can help vulnerable users protect themselves from harassment.	⊗ Prohibit or make it difficult for platforms to provide (pseudo-) anonymity to users.
✓ Recognize the value of strong encryption in protecting users, ICT services, and the ICT ecosystem.	⊗ Establish a regime whereby use of strong encryption is disincentivized or prohibited. Allow the law to discriminate against content based on medium.
✓ Ensure that authorities meet due process obligations and evidentiary thresholds before requesting sensitive user data.	⊗ Mandate that companies proactively share data with law enforcement absent specific demands.

83. Id. See also, Matthew Green, “EARN IT is a direct attack on end-to-end encryption,” 6 March 2020, <https://blog.cryptographyengineering.com/2020/03/06/earn-it-is-an-attack-on-encryption/>; Civil Society Letter to Sens. Graham and Blumenthal, March 6, 2020, [https://newamericadotorg.s3.amazonaws.com/documents/Coalition\\_letter\\_opposing\\_EARN\\_IT\\_3-6-20.pdf](https://newamericadotorg.s3.amazonaws.com/documents/Coalition_letter_opposing_EARN_IT_3-6-20.pdf).

84. U.S. Department of Justice, “Section 230 — Nurturing Innovation or Fostering Unaccountability?” June 2020, <https://www.justice.gov/file/1286331/download>

## 4. CONCLUSION

For over a decade, the Global Network Initiative has brought together leading companies, civil society organizations, academics, and investors from around the world in a shared endeavor to enhance freedom of expression and privacy across the ICT sector. Together, our members have built trust and shared expertise with each other in order to help protect rights in the digital age. We have also worked constructively with a wide range of governments to ensure that well- intentioned efforts are fit for purpose and proportionate. The analysis and recommendations in this policy brief draw from this extensive experience and expertise. We look forward to continuing to engage with lawmakers and other stakeholders to keep human rights at the forefront of our collective efforts to reduce harms in the ICT sector.



## APPENDIX A - RECOMMENDATIONS

### LEGALITY

- Law/rule-making should be done openly, in a participatory manner that allows for diverse and expert inputs, based on empirical analysis, and accompanied by impact-assessments.
- To the extent substantial rule-making authority and discretion is delegated to independent bodies, create robust oversight and accountability mechanisms to ensure that such bodies act pursuant to the public interest and consistent with international obligations.
- Ensure public laws are “formulated with sufficient precision to enable an individual to regulate his or her conduct accordingly.”
- More appropriate are approaches that establish clear limiting criteria and leave the determination of when those criteria are met to a judge.
- Clearly and precisely define what is prohibited, as well as who can be held responsible for failing to enforce the prohibition.
- Set clear expectations for responsible company action with regard to reports of illegal content.
- Ensure the law requires transparency, oversight, and remedy so as to avoid “confer[ring] unfettered discretion for the restriction of freedom of expression on those charged with its execution.”

### LEGITIMACY

- Ensure that content that is prohibited falls within one of the enumerated “legitimate purposes” in ICCPR Art. 19(3).
- Ensure that controversial and offensive content is not prohibited simply because it makes certain audiences uncomfortable.
- Ensure that content that is allowed in analog contexts is also permitted in digital form.

### NECESSITY

- Provide empirical support or argumentative clarity to establish “a direct and immediate connection between the expression and the threat.”
- Conduct careful, public, participatory deliberation to ensure laws are appropriate to achieve their protective function, are the least

intrusive instrument amongst those which might achieve their protective function and are proportionate to the interest to be protected.

- Carefully consider which types of private services at which layers in the technology stack are most appropriately positioned to address the specific concern(s) at issue.
- Accommodate a diverse range of business models and capacities. Consider how requirements may impact start-ups and smaller entities, as well as any unintended impacts they could have on competition policy.
- Provide clear guidance as to the precise characteristics of content and circumstances that require prompt or significant action.
- Articulate standards for appropriate content moderation based on traditional rule-of-law concepts such as transparency, due process, and remedy.
- Allow for variation and experimentation in approach, including “quarantining” and “downranking” of content. Provide means to guard against intentional misuse and unintentional consequences of content removal measures, including appeal and remedy mechanisms.
- Ensure robust remedial mechanisms for users whose content is restricted in order to avoid incentivizing self-censorship and over-removal. Build periodic reviews or reauthorizations into the law, in order to ensure that it remains relevant and consistent with evolving norms and technologies.

### PRIVACY

- Think creatively about how to facilitate accountability for those who violate the law, while continuing to strengthen privacy protections for all.
- Recognize that anonymity and pseudo-anonymity can help vulnerable users protect themselves from harassment.
- Recognize the value of strong encryption in protecting users, ICT services, and the ICT ecosystem.
- Ensure that authorities meet due process obligations and evidentiary thresholds before requesting sensitive user data.

# CONTENT REGULATION AND HUMAN RIGHTS

## INDEX OF LAWS AND LEGISLATIVE PROPOSALS

\*No English translation

- [Australia's "Sharing of Violent Abhorrent Material" Bill](#)  
p. 18, 19, 21, 22, 24
- [Brazil's Law on Liberty, Responsibility and Transparency on the Internet \(\*Fake News Bill\*\), 2020](#)  
p. 20, 21, 27
- [U.S. Department of Justice Proposal: "Section 230 — Nurturing Innovation or Fostering Unaccountability," 2020](#)  
p. 28
- [U.S. EARN IT Act, 2020](#)  
p. 10, 27, 28
- [U.S. "Limiting Section 230 Immunity to Good Samaritans Act," 2020](#)  
p. 10
- [U.S. "Stopping Big Tech's Censorship Act," 2020](#)  
p. 10
- [U.S. Online Freedom and Viewpoint Diversity Act, 2020](#)  
p. 10
- [U.S. Platform Accountability and Consumer Transparency Act \("PACT Act"\)](#)  
p. 14, 20
- [EU Digital Services Act \(DSA\)](#)  
p. 24
- [EU regulation on "preventing the dissemination of terrorist content online" \(Terror Regulation\)](#)  
p. 11, 14, 19, 22
- [French "Avia's Law" \(rejected, July 2020\)](#)  
p. 12, 13, 22, 24, 27
- [French Interim Mission Report, 2019](#)  
p. 11, 13, 14, 17, 20, 22
- [Germany's NetzDG Law, 2017](#)  
p. 12, 14, 18, 20, 21, 27
- [India's Draft Intermediaries Guidelines \(Amendment\) Rules, 2018 \("the Draft Amendments "\)](#)  
p. 11, 12, 26, 27
- [Ireland's Online Safety Media Regulation Bill, 2019](#)  
p. 10
- [Kenya's Computer Misuse and Cyber Crimes Act](#)  
p. 18
- [Pakistan's The Citizens Protections \(against Online Harms\) Rules, \(the Rules\), 2020](#)  
p. 12, 21, 26–28
- [Singapore's Protection from Online Falsehoods and Manipulation Act, or "POFMA," 2019](#)  
p. 12, 20, 26
- [Tanzanian Electronic and Postal Communications \(Online Content\) Regulations, 2020](#)  
p. 12
- [Turkey's "Social Media Bill," 2020\\*](#)  
p. 20, 21, 28
- [U.K. Online Harms White Paper, 2019](#)  
p. 9–12, 14, 16, 18, 21–24, 26



GLOBAL  
NETWORK  
INITIATIVE