September 24, 2021

Digital Citizen Initiative
Department of Canadian Heritage
25 Eddy St
Gatineau QC K1A 0S5

SUBMITTED VIA EMAIL TO pch.icn-dci.pch@canada.ca

Re: The Government of Canada's proposed approach to address harmful content online

Dear Department of Canadian Heritage:

Access Now[1] writes to express its concerns regarding the Government of Canada's (the "Government") proposed approach to address harmful content online released on July 29, 2021.[2] The Government's goals are laudable as everyone, including the Government, should seek to reduce harmful speech, including hate speech, online. However, the Government's proposal will not achieve these goals. Instead, the proposed framework threatens fundamental freedoms and human rights.

The Government should ensure any legislative framework enacted into law protects human rights, including the rights to freedom of expression and speech, while also making it easier to address illegal content, hate speech, and other harmful online content. With this in mind, Access Now offers human rights-centered guidelines for content governance and urges the Government to substantially revise its approach to comply with international standards. Specifically, Access Now argues that the Government should reconsider the scope of vague definitions and overly broad categories of "harmful content," provide adequate time frames for content removal, avoid imposing proactive monitoring or filtering obligations, make fines and other sanctions proportionate, and refrain from mandating overly broad website blocking at the internet service provider level.

---

[1] Access Now is an international human rights organization that defends and extends the digital rights of people at risk across the globe. This includes ensuring that individuals and groups, particularly those marginalized, do not become victims of censorship, whether through government policies or corporate practices. *See* https://www.accessnow.org/.

[2] *Have your say: The Government's proposed approach to address harmful content online*, Government of Canada, (July 29, 2021) https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content.html.

Reconsider the scope of vague and overly broad categories of "harmful content"

Legal clarity, precise definitions, and a narrowly tailored scope of a legislative proposal are essential preconditions to the proper functioning of the rule of law. Yet, the broad categories established in the proposal are too vague and overly broad. The technical paper incorporates five categories of harmful content: terrorist content; content that incites violence; hate speech; non-consensual sharing of intimate images; and child sexual exploitation content. It states that definitions for various categories of "harmful" content will be borrowed from the Criminal Code and "adapted to the regulatory context."[3] While the proposal seeks to tackle "potentially illegal content falling within the five categories of speech identified as harmful," it is also aimed at potentially harmful but *legal* categories of user-generated content, such as "material relating to child sexual exploitation activities that may not constitute a criminal offence, but when posted on an OCS is still harmful to children and victims."[4]

Legislation that imposes burdens on online platforms' content moderation practices should be limited to illegal content only. The principle of legality requires that offenses should be "clearly defined in the law" and "foreseeable for any person."[5] The proposal, however, falls short of satisfying the principle of legality because of its overly broad definitions and scope. Potentially legal but "harmful" content is an inherently vague concept that is difficult for platforms to define and the government to enforce; thus, a government mandate to police such content is highly prone to human rights abuse as companies take down more content than is necessary.

For example, the technical paper defines terrorist content as "content that actively encourages terrorism and which is likely to result in terrorism."[6] This is a very broad definition of (illegal) terrorist content that omits the element of intent, which should be a facet of all elements constituting terrorist criminal offences and, in fact, is an element of Canada's definition of terrorist activity.[7] Without considering the intention of the poster, there is a serious risk that any user-generated content related to terrosism, including news reporting, academic research, or historical resources, will be automatically deleted when flagged. Such a measure does not comply with the constitutional safeguards of a democratic society.

---

[3] *Technical Paper*, Para. 8.

[4] *Technical Paper*, Para. 8.

[5] Daniel Gradinaru, *The Principle of Legality, Proceedings of the 11th International RAIS Conferenc*e, (Nov. 19-20, 2018), http://rais.education/wp-content/uploads/2018/11/044DG.pdf; *see also Practice Relating to Rule 101: The Principle of Legality*, IHL Database, https://ihl-databases.icrc.org/customary-ihl/eng/docs/v2_rul_rule101.

[6] *Technical Paper*, Para. 8.

[7] Criminal Code, R.S.C. 1985, c. C-46, s. 83.01 ("terrorist activity means … an act or omission … (i) that is committed (a) in whole or in part for a political, religious or ideological purpose, objective or cause, and (b) in whole or in part with the intention of intimidating the public, or a segment of the public, with regard to its security … and … (ii) that intentionally" causes death, serious bodily harm, endangers someone's life, and causes serious property damage, among other things).

Overly broad and opaque legal definitions will ultimately lead to an unnecessary and disproportionate interference on the right to freedom of expression. One of the dangers of overly-broad definitions for "harmful" content is that a content moderation professional with no legal training could quickly resort to bias in deciding which content to remove. This type of chilling effect has a disparate impact on marginalized communities, including communities of color, religious groups, and LGBTQ+ individuals. For example, a U.S. law intended to penalize sites that hosted speech related to child sexual abuse and trafficking led large and small internet platforms to censor broad swaths of speech with adult content.[8] Instead, the consequences of this censorship devastated the community the legislation sought to protect.[9]

Nevertheless, legislators can still address potentially "harmful" but legal content by regulating Online Communication Service Providers ("OCSPs") processes and systems for content moderation and content curation. This approach includes legally mandated criteria of meaningful transparency, due process requirements to enforce platforms' community standards, independent auditing of these systems, and other due diligence safeguards. However, the role of public regulators should be limited to public oversight, ensuring that content moderation and content curation systems are sufficiently transparent and that people have clear and compelling grievance and redress mechanisms available to them. The Government can find an example of such a novel approach in the proposed Digital Services Act of the European Union currently being debated in the European Parliament or in the PACT Act in the United States.[10]

Access Now urges the Government to reconsider the scope of its definitions for "harmful" content to ensure clarity in the law and avoid overly broad content takedowns.

Provide adequate timeframes for removing flagged content

The timeframes for removing flagged content are onerous and will lead to significant impacts on freedom of expression and speech. The technical paper proposes that OCSPs "address all content that is flagged by any person in Canada as harmful content expeditiously."[11] The term expeditiously is

---

[8] Elliot Harmon, *How Congress Censored the Internet*, EFF (Mar. 21, 2018), https://www.eff.org/deeplinks/2018/03/how-congress-censored-internet.

[9] *Is Sex Work Decriminalization the Answer? What The Research Tells Us*, ACLU (Oct. 21, 2020), https://www.aclu.org/sites/default/files/field_document/aclu_sex_work_decrim_research_brief_new.pdf; Karen Gullo and David Greene, *With FOSTA already leading to censorship, plaintiffs are seeking reinstatement of their lawsuit challenging the law's constitutionality*, EFF (Mar. 1, 2019), https://www.eff.org/deeplinks/2019/02/fosta-already-leading-censorship-we-are-seeking-reinstatement-our-lawsuit.

[10] Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC (Dec. 15, 2020), https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-parliament-and-council-single-market-digital-services-digital-services; PACT Act, S. 797, Congress.Gov, https://www.congress.gov/bill/117th-congress/senate-bill/797.

[11] *Technical Paper,* Para. 11(A).

defined as “twenty-four(24) hours from the content being flagged.”[12] The Governor in Council has the authority to adjust this timeframe for different types of harmful content, including the power to shorten the timeline.[13] Within that timeline, OCSPs have two options: if flagged content qualifies as “harmful,” the OCSP must remove the content from its platform; if flagged content does not qualify as “harmful” the OCSP must provide an explanation to the person who reported the content as to why it does not fall under the definition of harmful content.[14] OCSPs that violate the framework are subject to financial penalties of up to three percent of global revenue or $10 million.[15]

A twenty-four-hour deadline to determine whether online speech meets the definition of harmful content and should be removed from a platform is an unreasonable and onerous obligation. Without adequate time to make a content moderation decision, OCSPs will by default remove flagged content regardless of its illegality or harmfulness. Content moderators are often overworked, have many cases to review, and are not truly qualified to make legal determinations. This makes over-reliance on legal criteria and inadequate, biased, or subjective censorship of content inevitable under harsh restrictive time frames for content removals. With such a short timeframe for review, it would be almost impossible for a content moderator to understand the full context of certain content. And for OCSPs that operate in multiple time zones, short time frames allocated for response would likely impose onerous burdens on smaller OCSPs with limited staff. Even worse, the harsh twenty-four hour deadline for content removals could compel OCSPs to deploy automated filtering technologies at a scale that could further result in the general monitoring of online content, ultimately violating the rights to freedom of expression and privacy.[16] Any revisions to the proposal should consider these nuances and the capabilities of smaller OCSPs on the market.

Strict and short deadlines for content removals cannot be reconciled with international human rights law especially when other recent proposals that follow the same proposed approach to censoring online speech are under heightened constitutional scrutiny. For example, the Constitutional Council of France declared short deadlines for removing online hate speech and other types of illegal content unconstitutional due to their negative impact on the right to freedom of expression.[17] According to the Council “[t]he shortness of the period left to the operators to proceed with this withdrawal, coupled with the difficulty for them to determine whether or not the comments are manifestly illegal, will encourage them to remove any content flagged as potentially illegal.”[18]

---

[12] *Technical Paper,* Para. 11(B).

[13] *Technical Paper,* Para. 11(C).

[14] *Technical Paper,* Para. 11(B).

[15] *Technical Paper,* Para. 108.

[16] See below, section entitled “Do not impose proactive monitoring or filtering obligations.”

[17] Press Release, Victory! French High Court Rules That Most of Hate Speech Bill Would Undermine Free Expression (June 18, 2020), https://www.eff.org/press/releases/victory-french-high-court-rules-most-hate-speech-bill-would-undermine-free-expression.

[18] *Decision n ° 2020-801 DC of June 18, 2020*, The Constitutional Council of France, https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm.

Another similar and controversial law, the German Network Enforcement Act ("NetzDG"), compels platforms to take down "manifestly illegal" content within 24 hours.[19] Several critics have pointed out NetzDG's severe implications on free speech online.[20] In his critique of the German law, the U.N. Special Rapporteur on Freedom of Expression reaffirmed that "[s]trict time periods of 24 hours … coupled with … severe penalties, could lead social networks to over-regulate expression - in particular, to delete legitimate expression, not susceptible to restriction under human rights law, as a precaution to avoid penalties."[21] In 2018, Human Rights Watch called the German law flawed — critiquing it for being "vague, overbroad, and turn[ing] private companies into overzealous censors to avoid steep fines, leaving users with no judicial oversight or right to appeal."[22]

Under the international human rights framework, Canada must ensure that its policies and laws do not restrict the right to freedom of expression. Unfortunately, the Government's proposal mirrors the most harmful aspects of the worst intermediary liability regimes around the world. It presents a strong incentive for companies to remove speech to ensure compliance, even if it is not harmful or illegal. Access Now recommends that the Government replace the twenty-four content removal timeframe with a system that balances free speech, the capabilities of the OCSPs and protects Canadians against harmful content. At a minimum, the Government should revise the proposal to allow additional time to engage in a contextual analysis of flagged content. Different types of harmful online content may require different responses tailored to the specific type of content.

Do not impose proactive monitoring or filtering obligations

Proactive monitoring and filtering obligations from the government are particularly severe impositions. The technical paper requires OCSPs to "take all reasonable measures, which can include the use of automated systems, to identify harmful content that is communicated on its [platform] and that is accessible to persons in Canada."[23] Thus, the proposal imposes a general obligation to monitor content and places the burden on OCSPs to identify content for removal.[24] General monitoring obligations compel OCSPs to monitor content shared on their platforms indiscriminately and for an unlimited amount of time.

---

[19] *Overview of the NetzDG Network Enforcement Law*, Center for Democracy and Technology (July 17, 2017), https://cdt.org/insights/overview-of-the-netzdg-network-enforcement-law/.

[20] Diana Lee, *Germany's NetzDG and the Threat to Online Free Speech,* Yale Law School Media, Freedom & Information Access Clinic (Oct. 10, 2017), https://law.yale.edu/mfia/case-disclosed/germanys-netzdg-and-threat-online-free-speech.

[21] UN Commission on Human Rights, *Right to freedom of opinion and expression*, June 1, 2017, OL DEU 1/2017, https://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf.

[22] *Germany: Flawed Social Media Law*, Human Rights Watch (Feb. 14, 2018), https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law.

[23] *Technical Paper,* Para. 10.

[24] *Id*.

General monitoring requirements imposed by governments violate human rights including the right to freedom of expression. This concept is generally accepted globally. According to the Manila Principles, OCSPs "should never be required to monitor content proactively as part of an intermediary liability regime."[25] The Council of Europe recommendation also warned that governments "should not directly or indirectly impose a general obligation on platforms to monitor the content they merely give access to, or which they transmit or store, be it by automated means or not."[26] Likewise, the United Nations advised against censorship or monitoring of online content, noting that it infringes on the right to privacy; that "such precautionary censorship would interfere with the right to seek, receive, and impart information of all kinds on the internet,"[27] and is likely to amount to pre-publication censorship.[28] Further, "[n]o legal provision should ever mandate, incentivize, or give platforms any sort of indication that they should be proactively filtering content before it is uploaded."[29]

Consequently, the Government should remove the proactive monitoring and filtering mandate from the proposal. The Government should avoid assigning responsibility to OCSPs as adjudicators of online speech and discourse. Instead, the Government should consult with human rights advocates and experts and explore proportionate and effective alternatives that provide OCSPs with a reasonable response time to flagged speech.

Make sanctions for non-compliance proportionate

The sanctions in the proposal are punitive and disproportionate, and instead should be proportionate to the violation. The proposal includes a penalty of 3% or $10 million dollars, whichever is higher.[30] In lieu of a penalty, a potentially-offending OCSP may enter into a compliance agreement with the Digital Safety Commissioner, and violations of that agreement can be up to 5% gross global revenues or $25 million.[31]

Disproportionate sanctions inevitably lead to over-compliance, harming free expression and access to information. Article 19(3) of the Universal Declaration on Human Rights lays down the conditions that

---

25 Manila principles on intermediary liability (2015), https://www.manilaprinciples.org/.

26 Council of Europe (2018), *Recommendation CM/Rec(2018)2 of the Committee of Ministers to member states on the roles and responsibilities of Internet intermediaries*, https://rm.coe.int/1680790e14 intermediaries.

27 UN Commission on Human Rights, *Right to freedom of opinion and expression* (June 1, 2017) OL DEU 1/2017, https://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf; *Decision n ° 2020-801 DC of June 18, 2020*, The Constitutional Council of France, https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm.

28 UN Commission on Human Rights, *Right to freedom of opinion and expression* (April 6, 2016) A/HRC/38/35 Para 67, https://www.undocs.org/A/HRC/38/35.

29 Eliska Pirkova & Javier Pallero, *26 Recommendations on Content Governance*, Access Now (Mar. 3, 2016), https://www.accessnow.org/cms/assets/uploads/2020/03/Recommendations-On-Content-Governance-digital.pdf.

30 *Technical Paper,* Para 108.

31 *Technical Paper,* Para 119.

any restriction on freedom of expression must satisfy: any restrictions on speech must be lawful, necessary to achieve a legitimate aim, the least restrictive means available, and proportionate to the aim pursued.[32] The U.N. Special Rapporteur has already warned that "high fines raise proportionality concerns and may prompt social networks to remove content that may be lawful."[33] Moreover, the International Covenant on Civil and Political Rights states, "governments may only impose restrictions on freedom of expression for reasons of national security or other pressing public need if they are provided by law and are strictly necessary and proportionate for achieving a legitimate aim." Similarly, the principle of proportionality in Canadian jurisprudence requires that a measure be reasonably necessary to achieve an objective and the least intrusive method.[34]

As written, the proposal combines a short content takedown regime with stiff monetary penalties for non-compliance — the perfect cocktail for mass-removal of content. While the system provides harsh fines for OCSPs who leave up "harmful" content beyond twenty-four hours, there is no penalty for taking down legal speech. With those incentives, it will naturally lead to over-removal of content. Therefore, the Government's approach raises proportionality concerns and represents an undue interference with a fair assessment of whether content violates the proposal. Any new or revised regulation should ensure that sanctions imposed on OCSPs operating in Canada are proportionate to the objectives of the legislation.

The government should not force removal of websites without clear standards

Government-mandated website takedowns are particularly harmful without especially clear standards. The proposal grants the Digital Safety Commissioner the authority to apply to the Federal Court for an order requiring certain Telecommunications Service Providers to "block access in whole or in part to" an offending OCSP that repeatedly refuses to comply with requests to remove child sexual exploitation or terrorist content.[35]

As a general matter, website blocking is a blunt measure that interferes with freedom of expression and has been condemned as a violation of human rights by the United Nations.[36] Canada recently

---

[32] United Nations (1966), *International covenant on civil and political rights*, Article 19, http://www.ohchr.org/EN/ProfessionalInterest/Pages/CCPR.aspx.

[33] *See* UN Commission on Human Rights, *Right to freedom of opinion and expression* (June 1, 2017) OL DEU 1/2017, https://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf; *see also Decision n ° 2020-801 DC of June 18, 2020*, Constitutional Council of France, https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm.

[34] Baker McKenzie, *Proportionality in Sentencing (Canada): White Collar Offenders Beware,* Lexology (Jan. 11, 2016), https://www.lexology.com/library/detail.aspx?g=67364fc2-9cda-4165-a848-52e074d6a85d.

[35] *Technical Paper,* Para 120.

[36] Michael Geist, *UN Special Rapporteur for Freedom of Expression: Website Blocking Plan "Raises Serious Inconsistencies" With Canada's Human Rights*, Michael Geist Blog (Mar. 31, 2018), https://www.michaelgeist.ca/2018/03/un-special-rapporteur-for-freedom-of-expression-bell-coalition-website-blocking-plan-raises-serious-inconsistencies-with-canadas-human-rights-obligations/; *see* UN Commission on Human Rights, *RE: Application to Disable On-line Access to Piracy Sites, CRTC File No 8663-A182- 201800467* (Mar.

proposed a similar takedown system for "piracy" websites.[37] In response, the U.N. Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression warned the Canadian Government against implementing a website blocking regime, noting that blocking "raises serious inconsistencies with Canada's obligations under Article 19 of the International Covenant on Civil and Political Rights and related human rights standards," particularly the necessity and proportionality of the requirement.[38]

While the proposal requires a federal court to decide legality, website takedowns are limited to terrorist content and child sexual exploitation, which as described above have the broadest and most difficult definitions. An OCSP may have legitimate reasons not to remove certain content, especially that the content was not illegal and did not, in its view, violate any of the vague definitions, but may still result in full takedowns for refusing to take the content down. By combining the threat of website blocking with opaque and overly broad definitions, the Government is incentivizing OCSPs to censor any content related to terrorism or child sexual exploitation to avoid non-compliance. Even the "mere threat of blocking may have a significant and disproportionate chilling effect on its operation" as OCSPs would be inclined to take down lawful content rather than risk being shut down, as has occurred in other countries that have implemented website blocking.[39] The Government should therefore remove these provisions.

Conclusion

The Government should avoid internet legislation, such as the instant proposed legislation, that endangers freedom of expression, speech, and information online. The obligations in the legislation impose unrealistic requirements on OCSPs and pose grave risks to human rights. The Government's proposal has several deficiencies and should be amended consistent with these comments, or abandoned.

29, 2018), https://services.crtc.gc.ca/pub/ListeInterventionList/Documents.aspx?ID=272698&en=2018-0046-7&dt=i.

[37] *See* Michael Geist, The Case Against the Bell Coalition's Website Blocking Plan Part I: Canada's Current Copyright Law Provides Effective Anti-Piracy Tools, Michael Geist Blog (Feb. 12, 2018), https://www.michaelgeist.ca/2018/02/case-bell-coalitions-website-blocking-plan-part-1-canadas-current-copyright-law-provides-effective-anti-piracy-tools/.

[38] UN Commission on Human Rights, *RE: Application to Disable On-line Access to Piracy Sites, CRTC File No 8663-A182- 201800467* (Mar. 29, 2018), https://services.crtc.gc.ca/pub/ListeInterventionList/Documents.aspx?ID=272698&en=2018-0046-7&dt=i

[39] *Id*.