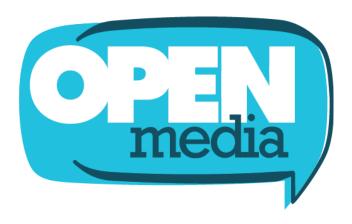**Response to 'The Government's proposed approach to address harmful content online'**
Submitted to the Digital Citizen Initiative at the Department of Canadian Heritage

OpenMedia is a community-based organization that safeguards the possibilities of the open Internet.

September 25, 2021

OVERVIEW OF RESPONSE

A. Introduction

i) About OpenMedia

OpenMedia is a community-driven organization of over 350,000 members that work together to keep the Internet open, affordable, and surveillance-free. We operate as a civic engagement platform to educate, engage, and empower Internet users to advance digital rights around the world.

Our organization and community members have been active participants at the Canadian Radio-television and Telecommunications Commission (CRTC), and have participated in numerous parliamentary review processes and consultations on issues impacting Canada's digital policy. We work to connect those most impacted by policy decisions directly with those making those policies, expanding our democratic processes to maximize public engagement.

For this particular consultation, members of the OpenMedia community have already delivered more than 8,600 unique emails providing individual feedback to the Government of Canada's public consultation on harmful content online.

This formal response on behalf of the organization accompanies and reinforces our community members' individual messages, expanding on the concerns they've raised with the government about the plan for our Internet described in the consultation's discussion guide and technical paper.[1] [2]

ii) Context of the consultation

OpenMedia recognizes this proposal appears to form part of a wider plan from the Canadian government to affect changes to Canada's Internet, covering both illegal content, and other behaviour and content that may be seen as harmful. Heritage Minister Steven Guilbeault has repeatedly spoken of rude speech against public officials as an online harm that is undermining democracy, and the Capitol insurrection in the U.S. as a product of uncontrolled online speech.[3] [4] The technical paper that accompanies this consultation itself frequently uses 'harmful content' as a stand-in for the five forms of illegal content it seeks to place new obligations on platforms to address, further muddying the issue.

We believe there are real problems with both illegal content on the Internet, and legal but in some ways harmful content. But as an organization whose mandate is to fiercely defend the

---

[1] Department of Canadian Heritage (2021). *Discussion Guide*
https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/discussion-guide.html
[2] Department of Canadian Heritage (2021). *Technical Paper*
https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/technical-paper.html
[3] Elizabeth Thompson (2021). "Canada not exempt from social media forces that created U.S. Capitol riot, heritage minister says." *CBC* January 29 2021.
https://www.cbc.ca/news/politics/facebook-twitter-canada-regulation-1.5894301
[4] Canada 2020 (2021). *Democracy in the Digital Age: Addressing Online Harms*
https://canada2020.ca/democracy-in-the-digital-age-addressing-online-harms/

legal expressive and privacy rights of people in Canada, the government's casual disregard for both of these rights is deeply concerning to us.

Minister Guilbeault has claimed that the government only seeks to "reproduce the same framework that exists in the physical world in the virtual world." [5] Yet the proposals within this consultation show a shocking lack of concern for maintaining this balance, or for understanding the real world impact they will have.

If adopted as written, the proposals in this consultation would lead directly and predictably to an unprecedented increase in the removal of considerable legitimate and lawful forms of speech online. It would also lead to the automatic reporting of an enormous volume of lawful content directly to the Royal Canadian Mounted Police (RCMP) and Canadian Security Intelligence Service (CSIS), deputizing online platforms as surveillance agents of the state in a system not seen anywhere else in the democratic world. And it would singularly fail to protect marginalized communities on the Internet, instead empowering their current victimizers in troll communities and law enforcement to more effectively target and harass them.

Policy-makers are responsible for the foreseeable consequences of their policies, not just their intended or desired outcomes. You are accountable for each of these disastrous consequences.

We're aware that other commentators are providing strong input to the consultation focused on the domestic legal and constitutional implications of the proposal, its compatibility with Canada's obligations under international law, and its potential incompatibility with the USMCA. We share their concerns, and note that a bill bearing striking similarities to the proposals in this consultation was recently struck down on constitutional grounds in France due to the precise issue of over-removal of lawful speech that we discuss below.[6]

Our submission will however focus on where we are best positioned to comment: an analysis of the predictable and damaging consequences of the proposal as described, (Section B), and a nudge towards more potentially more productive directions for future government intervention on these issues that should be explored instead (Section C).

iii) This consultation is not adequate or legitimate

The consultation presented to us does not have the features of a true public consultation, as has been pointed out by those both supportive and skeptical of new government regulation of

---

[5] Elizabeth Thompson (2021). "Canada not exempt from social media forces that created U.S. Capitol riot, heritage minister says." *CBC* January 29 2021.
https://www.cbc.ca/news/politics/facebook-twitter-canada-regulation-1.5894301
[6] Conseil Constitutionnel (2020). *Décision n° 2020-801 DC du 18 juin 2020*
https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm

Internet platforms.[7] [8] This is nothing more than a formal presentation of a predetermined plan, with an unreasonably short time frame for public comment.

This consultation provides absolutely no opportunity to help shape the framework of either the problem at hand, nor any of the proposed solutions. Rather than a solicitation of public and expert input on what the government should do, the technical paper appears to be a list of what the government will do, regardless of what it hears during the consultation period.
 It asks no open-ended questions. It does not solicit any evidence about problems on online platforms, nor does it present any evidence that justifies or explains the systems it proposes. It does not entertain or even reference alternative or complementary approaches to its proposed measures.

This is unacceptable policy-making in a democratic society. But it is particularly egregious as the government considers infringing on our Charter of Rights and Freedoms, and limiting citizens' ability to participate in the primary public spaces of our era, online platforms.

The timing of this consultation is also deeply inappropriate. The deadline for public comment was never published on the consultation page, and the consultation period given in the announcement was too short for substantive public input. But once a federal election was called, this entire consultation should have immediately been rescheduled. This would have comported with Privy Council Office guidance for election periods, as the matters under consideration are very clearly neither routine nor non-controversial.[9]

The overlap with the federal election made public engagement with the consultation significantly more difficult, in part due to regulations placed on third parties in an election, in 2019's Bill C-76. It was further challenged by the limited capacity of experts, academics, public interest groups, and concerned citizens to speak out and mobilize the general public during an election period, and a time-bound requirement for election participation that distracted from the potential to simultaneously participate in this consultation. OpenMedia strongly suspects that the timing of this consultation has significantly reduced the amount of participation from subject matter experts, whose voices are critical in ensuring a fulsome discussion of such issues and proposals.

The consultation's irregularities and deficiencies are major reasons it has drawn widespread criticism from a broad swath of both the academic content moderation-focused community, and

---

[7] Haggart, Blayne and Tusikov, Natasha (2021). *Not much of a consultation, not much of a plan: Our submission regarding the federal government's proposed approach to addressing harmful content online* https://blaynehaggart.com/2021/09/24/not-much-of-a-consultation-not-much-of-a-plan-our-submission-regarding-the-federal-governments-proposed-approach-to-addressing-harmful-content-online/

[8] Internet Society Canada Chapter (2021). Submission to the Department of Canadian Heritage: Consultation on Online Harms https://internetsociety.ca/wp-content/uploads/2021/09/ISCC-Response-Online-Harms-Final-21-9-21-1.pdf

[9] OpenMedia (2021). Open letter: Defer consultations on the Internet until after the election https://openmedia.org/article/item/open-letter-requesting-rescheduling-of-open-internet-consultations

the civil rights community, in Canada and abroad.[10] [11] [12] [13] [14] It compares very poorly to the more serious multi-year consultations that have been held in jurisdictions that have adopted broadly comparable legislation.[15]

Our participation in this consultation should not be read as acceptance or endorsement of this process. We strongly believe this consultation is utterly inappropriate. However, given the government's steadfast insistence on proceeding regardless, we feel we have no choice but to submit an insufficient submission, to ensure that at least some of our comments and concerns can be placed on the public record. If, as we recommend, the consultation is abandoned, it should be replaced by a much more fulsome public discussion about how best to encourage sound content moderation practices on Internet platforms.

Recommendation: The government should abandon this inadequate consultation, and the proposals contained within. Instead, it should pursue a genuinely open discussion on these issues, one that solicits evidence from all interested parties on the nature of problems with online content moderation and appropriate solutions that could be entertained to them.

B. Concerns on the proposed legislative remedies

i) Go fast and break things: 24-hr takedowns guarantee over-policing of content

One of the key recommendations made by the consultation's technical paper is to implement a 24-hour timeline requirement for platforms to remove all potentially illegal content under the five categories identified: terrorist content, incitement to violence against people or property, hate

---

[10] Darryl Carmichael and Emily Laidlaw (2021). *The Federal Government's proposal to Address Online Harms: Explanation and Critique*
https://ablawg.ca/2021/09/13/the-federal-governments-proposal-to-address-online-harms-explanation-and-critique/

[11] Daphne Keller (2021). Twitter thread:
https://twitter.com/daphnehk/status/1421118036895961094

[12] *Lawbytes Podcast* (2021). "Episode 99: Cynthia Khoo on the Canadian Government's Online Harms Consultation"
https://www.michaelgeist.ca/2021/08/law-bytes-podcast-episode-99/

[13] Electronic Freedom Frontier (2021). *O (No!) Canada: Fast-Moving Proposal Creates Filtering, Blocking and Reporting Rules—and Speech Police to Enforce Them*
https://www.eff.org/deeplinks/2021/08/o-no-canada-fast-moving-proposal-creates-filtering-blocking-and-reporting-rules-1

[14] Michael Geist (2021). *Picking Up Where Bill C-10 Left Off: The Canadian Government's Non-Consultation on Online Harms Legislation*
https://www.michaelgeist.ca/2021/07/onlineharmsnonconsult/

[15] Haggart, Blayne and Tusikov, Natasha (2021). *Not much of a consultation, not much of a plan: Our submission regarding the federal government's proposed approach to addressing harmful content online*
https://blaynehaggart.com/2021/09/24/not-much-of-a-consultation-not-much-of-a-plan-our-submission-regarding-the-federal-governments-proposed-approach-to-addressing-harmful-content-online/

speech, non-consensual sharing of intimate images, and child sexual exploitation content. A harsh penalty of 3% of global revenue or $10 million dollars would be applied to any platform that fails to meet the standard.[16] This requirement will be in effect from the time a platform becomes aware of the content – which could mean from when it is posted, or any time content is flagged or reported by any user.

The government has presented this as a way of getting 'tough' on platforms who are not doing enough to remove illegal content. But this view ignores the predictable consequences these requirements will have on platform behaviour, and the subsequent impact on Internet users. In practice, this obligation will lead directly and overwhelmingly to the removal of large amounts of user speech which would not be found illegal by a court of law. This problem is especially acute as much of the content being flagged will be identified by individual platform users who object to the content, but are not legal experts, and not necessarily able to identify the difference between what is illegal, objectionable, or just something they dislike.

Handling the volume of content moderation decisions required daily on a major online platform with any degree of fairness to users is extremely challenging.[17] Any content moderation system inevitably produces errors, whether using either human or algorithmic judgment. At present, platforms continually readjust their standards and systems to account for widely criticized mistakes in both failing to remove content, **and** inappropriately removing content.

The one-sided obligations imposed in this proposal will put a heavy thumb on the scale in favour of systematically over-removing lawful content. The platform incentives are clear: there will be a heavy legal and financial risk attached to leaving up content that could conceivably be found illegal under any of the five harms of this proposal, but no counter-balancing incentive to encourage thoughtful or fair consideration of the expressive rights of the posting user.

Put plainly, the inevitable outcome of this obligation will be the removal of all but the most obviously innocuous content flagged under these harms within the 24-hr window, regardless of its legitimacy.

This outcome thoroughly undermines the government's stated objective of merely translating our offline speech standards to the Internet. And it cannot and will not be remedied by appealing to the government's proposed Digital Recourse Council to reinstate content. Platforms have no obligation or clear incentive to ever reinstate content; and returning speech to a platform months or years after it was posted is not meaningfully equivalent to allowing it in the first place. Further, studies have shown that having any content removed has a demonstrated chilling effect on

---

[16] Department of Canadian Heritage (2021). *Technical Paper* Para. 11(A), 108[J].
https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/technical-paper.html
[17] Michael Masnick (2021). *Masnick's Impossibility Theorem: Content Moderation At Scale Is Impossible To Do Well.*
https://www.techdirt.com/articles/20191111/23032743367/masnicks-impossibility-theorem-content-moderation-scale-is-impossible-to-do-well.shtml

further speech, both of the affected user and those who see their content removed, directly discouraging participation in public conversation.[18] [19]

A very wide range of lawful user speech could potentially fall afoul of the necessarily broad interpretation platforms will make of what could constitute illegal content, including but most certainly not limited to:
- Satire and humour;
- Support for or participation in protest movements;
- Documentation of human rights abuses;
- Artistic expression;
- Research and journalism on sensitive or violent topics;
- Voluntary adult sexual expression;
- Conversation by or within marginalized communities about their lived experience.

This potential mistargeting of lawful and important user speech is not hypothetical. Currently, platforms' content moderation that is intended to protect against hate speech frequently leads to unintended censorship of targeted groups.[20] [21] Similarly, attempts to remove content that glorifies violence frequently misfire and censor critical reporting and documentation of real world atrocities.[22] [23] Pressure from states has even platforms to directly interfere in critical, lawful social discourse about the justice and legality of government actions.[24]

A more thoughtful assessment of current online platform takedowns of illegal content should examine the average time verified illegal content remains online, and the reasons why, which

---

[18] Jonathan Penny (2017). "Internet surveillance, regulation, and chilling effects online: a comparative study", *Internet Policy Review* Volume 6:2. https://policyreview.info/articles/analysis/internet-surveillance-regulation-and-chilling-effects-online-comparative-case

[19] J. Nathan Matias, Jonathan Penney, Merry Ember Mou and Max Klein (2020). "Do Law Enforcement Bots Reduce Freedom of Expression Online? Study Results". *EAT Lab*. https://citizensandtech.org/2020/09/chilling-effect-automated-law-enforcemen/

[20] Conor Murray (2021). "TikTok algorithm error sparks allegations of racial bias." *NBC* July 9 2021. https://www.nbcnews.com/news/us-news/tiktok-algorithm-prevents-user-declaring-support-black-lives-matter-n1273413

[21] ACLU (2021). *Time and Again, Social Media Giants Get Content Moderation Wrong: Silencing Speech about Al-Aqsa Mosque is Just the Latest Example.* https://www.aclu.org/news/free-speech/time-and-again-social-media-giants-get-content-moderation-wrong-silencing-speech-about-al-aqsa-mosque-is-just-the-latest-example/

[22] Malachy Browne (2017). "Youtube removes Videos showing Atrocities in Syria." *The New York Times* August 22 2017. https://www.nytimes.com/2017/08/22/world/middleeast/syria-youtube-videos-isis.html

[23] Betsy Swan (2017) "Exclusive: Facebook Silences Rohingya Reports of Ethnic Cleansing" *Daily Beast* September 18 2017. https://www.thedailybeast.com/exclusive-rohingya-activists-say-facebook-silences-them

[24] Cat Zakrzewski (2020). "The Technology 202: Instagram faces backlash for removing posts supporting Soleimani." *The Washington Post* January 13 2020. https://www.washingtonpost.com/news/powerpost/paloma/the-technology-202/2020/01/13/the-technology-202-instagram-faces-backlash-for-removing-posts-praising-soleimani/5e1b7f1788e0fa2262dcbc72/

could lead to further suggestions on how to shorten the window without snaring overwhelmingly lawful speech in the process.

Recommendation: No mandatory time window should be put on content takedown decisions by platforms on individual pieces of content.

ii) Content moderation will never be completely unbiased

Content moderation decision-making will always be subjective, and cannot always be distilled down to a clear yes or no answer. Yet the consultation's proposal would require that automated decision making it mandates platforms adopt would not result in "any differential treatment of any group based on a prohibited ground of discrimination;" a requirement that is simply not possible – for online platforms, or for anyone.[25]

Both automated and human moderation have been shown to be rife with errors that are biased against members of protected groups.[26] Moderators, and moderation systems are well-known for making frequent mistakes. Combining algorithmic and human judgement does not undo these errors: it is more likely to conceal and reinforce them.[27]

While there's no 'right', unbiased way to do content moderation, there are many bad ways to do it. The inflexibility and punitive one-sided consequences of the government's proposal guarantees that online platforms will make their existing content moderation systems even worse.

At present, major corrections in content moderation processes on major platforms most often occur following independent journalism or internal leaks.[28] The independent, non-governmental source of these revelations and improvements is welcome and appropriate for monitoring globally relevant online platforms; their piecemeal nature is not.

Recommendation: Mandate independent, non-governmental and public auditing and transparency around content moderation tools and algorithms.

iii) Proactive surveillance obligations are unfit for democratic use

---

[25] Department of Canadian Heritage (2021). *Technical Paper* Para. 10a.
https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/technical-paper.html
[26] Maarten Sap, Dallas Card, Saadia Gabriel, Yejin Choi, and Noah A Smith (2019). "The Risk of Racial Bias in Hate Speech Detection" *Association for Computational Linguistics* (2019:1668-1678).
https://aclanthology.org/P19-1163.pdf
[27] Ben Green (2021). "The Flaws of Policies Requiring Human Oversight of Government Algorithms". *SSRN*
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3921216
[28] Jeff Horwitz (2021). "Facebook Says its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt." *Wall Street Journal* September 13, 2021.

Paragraph 10 of the consultation paper requires platforms to proactively surveil user posts for the five forms of illegal content treated by this proposal, using automated tools. This is an astonishingly overreaching and disproportionate measure that has been roundly criticized and rejected in other jurisdictions, even in much more narrowly scoped form.

Algorithmic detection is the only way to fulfill a proactive detection obligation at any meaningful scale. Yet algorithmic detection is extraordinarily prone to errors in detecting illegal material, particularly for heavily context-dependent speech such as hate speech, incitement to violence and terrorism. Major platforms currently use it judiciously for only the most easily detectable material, such as child sexual exploitation material, precisely because it is so error-prone for more general purposes.

Forcing more generalized adoption of automatic detection of illegal content will sharply increase the misidentification and removal of lawful content, particularly of socially sensitive and political speech. For this reason, multiple UN Special Rapporteurs, the Council of Europe, and the global Manila Principles have all warned against states adopting a proactive content detection or filtering obligation.[29] [30] [31]

Recommendation: Do not mandate proactive surveillance by platforms, especially of more context-dependent harms.

iv) Direct reporting to law enforcement treats all Internet users as criminals

The consultation's technical paper proposes that user posts and account information should be automatically and secretly turned over to law enforcement when platforms remove a post as potentially constituting one of the targeted five forms of illegal content. This is one of the most egregious aspects of the proposal, and is an astonishing data and power grab for law enforcement. This process directly circumvents the critical checks and balances we have in place to prevent abuse of power, over policing and surveillance of millions of innocent people in Canada.

In effect, it would create a mass surveillance system of much lawful speech by Canadians and non-citizens alike who have committed no crime. It must absolutely **not** be in any proposed legislation.

---

[29] Joseph Cannataci, UN Special Rapporteur on the right to privacy; David Kaye, UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression; Fionnuala Ní Aoláin, UN Special Rapporteur on the promotion and protection of human rights and fundamental freddoms while countering terrorism. Open letter from Dec 2018:
https://spcommreports.ohchr.org/TMResultsBase/DownLoadPublicCommunicationFile?gId=24234
[30] 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/1680790e14intermediaries
[31] *Manila Principles on Intermediary Liability* (2015), https://www.manilaprinciples.org/

The unbalanced platform incentives described earlier in this response will mean that the majority of removed content under the consultation's system will not actually constitute illegal content. It will consist of normal user activity that platforms remove because they can, and because any legal risk to them, even at a relatively low probability, is more important than silencing their user base.

By virtue of all flagged content being directly reported to law enforcement, countless Internet users will exist in databases alongside criminal content, in many cases simply because someone else on the platform flagged their content – often simply because they dislike it. Worse yet, the proposal fails to contain a single adequate indication that there would be any accountability for how law enforcement manages, retains, or deletes the data (if it ever does).

This type of law enforcement lawful access to user data has already been proposed, and rejected, numerous times in the past – perhaps most infamously in the debate surrounding 2011's *Bill C-30, The Protecting Children from Internet Predators Act*.[32] [33] The government must not support or create a surveillance state, proactively monitoring innocent internet users.

It is worth noting that Law enforcement in Canada is already flooded with many times more reports of hate crimes than they have the resources or willingness to act on.[34] Even complaints filed directly by those who feel a crime has been committed against them, are often ignored. Automatic reporting of online takedowns will make this situation many times worse, with agencies deluged with an ocean of online reports from platforms, the great majority of no real use.

OpenMedia is concerned this ocean of mostly lawful speech would serve only one meaningful purpose: the extra-judicial creation of an immense trawling net for law enforcement to target and gather intelligence about individuals who have committed no crime, but nonetheless attract attention from police and the powerful, including Indigenous activists, environmental movements, and members of otherwise marginalized ethnic and religious communities.

It is also worth emphasizing that platforms hold an almost unimaginably rich volume of information about their users, including their website traffic, likes and dislikes, commuting routes and geographic locations, detailed social networks, inferred current emotional states, and more. This is not only dangerous in the hands of a single company – an issue the government seems unwilling to address in its abandonment of its own privacy legislation in the last session of

---

[32] Government of Canada (2012). "An Act to enact the Investigating and Prventing Criminal Electronic Communications Act and to amend the Criminal Code and other Acts"
https://www.parl.ca/LEGISInfo/BillDetails.aspx?Language=E&billId=5375610
[33] OpenMedia (2012). *A look back at our Stop Spying campaign against Canada's Bill C-30*
https://openmedia.org/look-back-our-stop-spying-campaign-against-canadas-bill-c-30
[34] Mike Hager, "Alleged hate crimes rarely investigated by police, report claims," *Globe and Mail*, August 30 2021.
https://www.theglobeandmail.com/canada/british-columbia/article-alleged-hate-crimes-rarely-investigated-by-police-report-claims/

parliament – but is wildly inappropriate information for law enforcement to have about innocent internet users, without needing to demonstrate a clear need and threat, and obtain a warrant.

Until our government restricts the vast data platforms collect on us, a requirement for platforms to turn over user data in any circumstances outside clear and imminent threat to life or a confirmed serious crime presents an enormous threat to the right to privacy of people in Canada.

Canada is a democratic country, which cannot and must not treat all of its citizens as criminals. This proposal directly undermines the criminal justice system, our legal checks and balances on abuse of power, and puts Canada on par with some of the world's most oppressive governments.

Recommendation: Do not require reporting of user posts or information by platforms to law enforcement for anything less than clear and immediate threat to life, or once content has been deemed explicitly illegal. Do not mandate ANY automatic reporting to law enforcement.

v) Website blocking is disproportionate, ineffective, and unwelcome in Canada

The consultation paper proposes exceptional recourse that would require ISPs to block access to platforms if the platform repeatedly fails to remove child sexual exploitation material or terrorist content, and other enforcement mechanisms have been exhausted.[35]

It is assumed that this proposal is not targeted at mainstream online platforms, who generally already make adequate efforts to remove both these types of content. Even for smaller platforms, however, website blocking is deeply ineffective at its stated purpose, being easy to circumvent, and therefore very unlikely to deter highly motivated individuals seeking the abhorrent content described. Technologies such as VPNs, proxies servers, and Tor browser are widely available, and must remain so to allow millions of Internet citizens who live under oppressive regimes to communicate and access information, as their Internet is otherwise highly controlled and censored.[36]

The chief consequence of a website blocking regime would be removing access to mixed use platforms from their users who have no connection to illegal content, and are using the platforms legitimately.

As the Department is well aware, website blocking is not a new or uncontroversial issue in Canada. Despite widespread public opposition, the tactic has been proposed for Canada year

---

[35] Department of Canadian Heritage (2021). *Technical Paper* Para. 120.
https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/technical-paper.html
[36] Electronic Frontier Foundation (2020). *Understanding and Circumventing Network Censorship*
https://ssd.eff.org/en/module/understanding-and-circumventing-network-censorship

after year by media conglomerates who would like to make it harder for Canadians to access media from other countries without paying them for content that they've licensed.

It has been also been rejected by the CRTC and Parliament repeatedly as neither a proportionate nor effective remedy[37] [38]. Yet this year the government again proposed the remedy, in its Consultation on a Modern Copyright Framework for Online Intermediaries and was, again, met with widespread opposition concerned with the inappropriate and ineffective government overreach. OpenMedia expressed our concerns with this proposal in more detail earlier this year during this consultation.[39] [40]

Recommendation: Effective website blocking for highly motivated individuals is not technically feasible. The government should abandon its consideration for these purposes, and focus on developing our relationship with other jurisdictions to address services that intentionally host child sexual abuse material or terrorist content.

vi) Legal remedies must use the court system

Some portions of the government's proposal appear to be efforts to 'simplify' the process of assessing the legality of user posts by circumventing our existing legal process. Not only will this simplification not work, it will directly undermine and overload our existing legal system.

The Digital Recourse Council described in the technical paper consists of an appointed group of 3-5 people, with sensitivity to representation from Canada's diverse populations, but without an expressed requirement for legal or constitutional expertise or counsel.

It seems improbable that this small group will have the capacity or expertise required to deal with the volume of claims they will receive under this system. Countless groups and individuals will have a legitimate interest in having their right to express themselves reinstated by the Council, or illegality of others' content confirmed. The volume of cases brought before a body this small could lead to queues of many years for clear consolidation and response.

Whether the Council can manage the volume of appeal or not, it is unclear what value it is adding to the existing system. If its role is strictly to resolve relatively unambiguous applications

---

[37] OpenMedia (2021). *Thousands of OpenMedia community members just stood up to defend Canada's Internet from website blocking!*
https://openmedia.org/article/item/thousands-openmedia-community-members-stood-up-defend-canada-internet-from-website-blocking
[38] OpenMedia (2018). *Huge win for Canadians as CRTC rejects Bell's website blocking proposal.*
https://openmedia.org/press/item/huge-win-canadians-crtc-rejects-bells-website-blocking-proposal-title_duplicated
[39]Innovation, Science, Economic Development Canada (2021). *Consultation on a Modern Copyright Framework for Online Intermediaries.*
https://www.ic.gc.ca/eic/site/693.nsf/eng/00191.html
[40] OpenMedia (2021). *OpenMedia Submission to the Copyright Consultation on a Modern Framework for Online Intermediaries.*
https://openmedia.org/files/OpenMedia_-_Submission_to_Online_Intermediary_Consultation.pdf

of Canadian law, it is not clear why the Council itself is necessary. If it is intended to issue interpretative judgements, changing or reducing the current understanding of freedom of expression rights on online platforms compared to offline spaces, it would appear to be plainly usurping the rightful and necessary role of our court system.

If that usurpation is recognized and the Council is regularly overruled by our courts, the system will have been a waste of time and money, particularly for the victims and defendants forced to use it. If that usurpation is not recognized, we will have an extra-judicial system setting legal precedents in our country, which would be even more concerning.

Recommendation: Extra-judicial bodies cannot be put in the position of setting legal precedent. If legal clarification is required of how to apply Canada's laws on platforms, that must be a judicial responsibility.

vii) An all-powerful regulator is not the answer

A key mistake in this consultation is attempting to address too many disparate issues on the Internet with the same regulatory agent and power. Direct threats to human life, threats to property, the non-consensual distribution of sexual imagery, sexually exploitative material involving children and hate speech are very different issues. They differ in the immediacy and severity of potential harm, appropriate rights to information, appeal, and decision-making for victims and accused persons, and necessary legal and contextual expertise for a hypothetical regulatory body or agent.

By attempting to handle all of these harms through a single body and piece of legislation, the government is creating equally invasive powers, detection standards, and potential penalties in each case. This creates a slippery slope in which powers that could be justified for the most extreme potential harm are available to the regulator for very different smaller or contested cases. It is likely to lead to disproportionate procedures used in many cases not justified by their actual harm. This is even more likely given that a single overstretched regulator will lack the capacity to wisely and contextually interpret the full range of cases brought before it.

It also leaves on the table the potential for much more nuanced issue-sensitive remedies tailored to the type of violation. For example, independently managed hashed image databases have proven effective as a non-legislative tool for reducing the spread of child sexual exploitation material on online platforms. They may also have value as a solution for removing non-consensually distributed adult intimate imagery (NCDII), given that the impacted adults could verify their ID to the body and request its removal. Yet hashed databases would not be appropriate for removing evidence of promotion of terrorism or hate speech, since this content is more ambiguous in status and meaning, and society has many valid purposes for accessing it, including journalism, research, and documentation of real-world abuse.[41]

---

[41] Danielle Citron and Neil M. Richards (2018). "Four Principles for Digital Expression (You Won't Believe #3!)", *Washington University Law Review* (Vol 95:1353-1387:2018). https://wustllawreview.org/wp-content/uploads/1353-1387-Citron-Richards_Final.pdf

The astonishing breadth of power assigned to the proposed Digital Safety Commissioner, including the power to compel online platforms "to do any act or thing, or refrain from doing anything," seems a sign of a real lack of clarity about what the creation of the Commissioner's position is actually trying to accomplish.[42] This is not an appropriate approach to creating an extremely powerful new body in a democratic society, particularly one governing an area as sensitive as online speech.

Recommendation: Do not to address all five forms of illegal content through the same system and procedures. Consult with scholars and issue experts about appropriate solutions to each. At a future consultation, publicly discuss all the options suggested and solicit opinions on them.

Recommendation: Any powers granted to a new or existing regulator over online platforms and online speech must be carefully defined, explained, justified, and clearly limited.

viii) The proposal will harm those it claims to help

This proposal is presented as a strategy to combat online hate, and better protect marginalized communities online. That makes it worth reviewing the ways it will significantly harm and worsen the experience of many marginalized people and victims of online attacks.

1) **Censoring the speech of marginalized communities:** It is clear to see how marginalized communities are already targeted online with hate, harassment, and abusive behaviours. Yet their ability to discuss this victimization, share examples of hate speech directed at them, and push back against that speech will be badly damaged by the predictable consequences of the consultation's proposals. Due to the clear incentives the proposal gives platforms to aggressively remove speech without much sensitivity to context, platforms will insensitively remove far more speech from targeted groups around their experience of social marginalization, and descriptions of the attacks others make on them. This is not hypothetical platform behaviour; it is already a common issue, without these new legal incentives that will strongly reinforce it.[43]

2) **Enabling online hate:** Counter-intuitively, forcing rigid content moderation rules on platforms will super-charge troll brigades who already use platform rules to attack marginalized individuals. No one knows the exact limits of a given platform's rules for

---

[42] Department of Canadian Heritage (2021). *Technical Paper* Para. 80.
https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/technical-paper.html
[43] Jessica Guynn (2019). "Facebook while black: Users call it getting 'Zucked,' say talking about racism is censored as hate speech." *USA Today* April 24, 2019.
https://www.usatoday.com/story/news/2019/04/24/facebook-while-black-zucked-users-say-they-get-blocked-racism-discussion/2859593002/

speech like some members of hateful online communities do, and no one is better at communicating hateful views while staying within those rules, while studiously observing their targets for pushback or past posts that might violate them.[44] The only way for platforms not to fall prey to this kind of rules lawyership is to continue giving platforms space to exercise judgment and flexibility in applying their own rules.

3) **Enabling law enforcement surveillance and over-policing of marginalized communities**: It is difficult to imagine a more powerful engine of over-policing of marginalized communities than giving law enforcement who already operate with bias an overwhelming volume of content takedown reports, and letting them pick and choose which to try to criminally enforce. Law enforcement surveillance of lawful marginalized communities is already a problem in Canada; the provisions described in the consultation will make it a much, much larger one.[45]

ix) Setting a dangerous precedent with global ramifications

Content posted to the Internet does not exist within a single national jurisdiction. Posts are available globally, and there is no easy way for platforms to justify that some national laws should apply to a given piece of content, but not others.

It is a deeply unreasonable expectation of this consultation's proposals that platforms will separately consider the nuance of law around expression within each jurisdiction they function in, for each piece of content, and individually mark content to be removed in only some jurisdictions.

As with other legal patchworks, a much more likely longer-term outcome is that platforms will take the broadest interpretation of Canada's laws on content takedowns, and combine it with broad interpretation of similar law in other jurisdictions, to create a single global standard for their moderation that universally protects them from legal threat.

This amalgamated standard would be systematically biased against freedom of online speech. The product would not be a product of the thoughtful weighing of the expressive rights of users versus removing illegal content that exists in any given democratic legal system, but rather a kind of race to the bottom for restrictions on user speech. Any overly broad law in any jurisdiction that poses a credible legal or financial threat to platforms would have the potential to become universalized, and limit expression across the global Internet.

---

[44] Phoenix CS Andrews (2021). "Social Media Futures: What is Brigading?" *Tony Blair Institute for Global Change*.
https://institute.global/policy/social-media-futures-what-brigading
[45] Bruce Livesey (2017). "Spies in our midst: RCMP and CSIS snoop on green activists". *National Observer* May 5 2017.
https://www.nationalobserver.com/2017/05/05/news/spies-our-midst-rcmp-and-csis-snoop-green-activists

There's one alternative, and it is even worse: Canada doing its part to usher in the so-called 'Splinternet'.[46] In a Splinternet model, large parts of the Internet are fenced off by restrictive national legislation controlling what comes in or out. This is not only what Canada is proposing, but an incredibly dangerous precedent for Canada to encourage, in terms of what it means for regulation by other governments of online content. Currently, major platforms are amongst the primary bulwarks against the emergence of splinternets, as their independent content standards make restrictive national censorship by governments more difficult. If Canada is successful in forcing many platforms to tailor their content systems to a specific government model, but only for Canadian users or within Canadian IP addresses, we will not only be building a shallow Splinternet of our own. But we'll also be furthering the legitimacy of much more restrictive Splinternets elsewhere.

This is only furthered by the proposals to directly tie these content regulations to law enforcement reporting requirements, something that could lead to the direct persecution of millions of Internet users globally who currently use online platforms as one of the few areas they are able to express themselves.

C. What a better discussion of online harms might look like

i) Clearly separate illegal content from online harms

The government's continued pattern of conflating discussions of illegal content and other problems with legal speech online is deeply concerning. Throughout this proposal, the distinction is blurred, with the term 'harmful content' used as a stand-in for illegal content.

Outside of this proposal, Minister Guilbeault has spoken of problems of online civility, misinformation, and rude language directed at politicians as types of harmful online content that the government is concerned with addressing.

It is not, and can never be the government's role to police online civility or factualness. That's not a power that is safe for any government to have, or that people in Canada will tolerate.

The power to criminalize and remove speech from the Internet, either directly or functionally by foreseeable consequences of your legislation, must be handled extraordinarily carefully. Any new regulation must be restricted to illegal speech, with careful attention to whether it is disproportionately leading to removal of legal speech, as we've argued above.

Moving forward from this consultation, it is critical that the government be extremely clear about when it is speaking about illegal content, plainly falling within the five forms of illegal content described in the consultation paper, and when it is speaking of other issues on the Internet.

---

[46] Jeff John Roberts (2019). "The Splinternet is Growing" *Fortune* May 29, 2019.
https://fortune.com/2019/05/29/splinternet-online-censorship/

We've flagged many concerns with the new proposals put forward in this consultation for treating this illegal content, and believe they should be largely rejected due to their enormous splash consequences for lawful expression. However, we recognize the government's legitimate interest in enforcing the existing laws around illegal content.

Toxic and harmful behaviour clearly exists on the Internet outside these forms of illegal content. But the government's appropriate role in contributing to addressing these issues is not the blunt enforcement of mass content removal, without context or accountability.

ii) Addressing the knowledge gap around harmful online content

Calls for further research can be read as a call to do nothing. But content moderation at the scale online platforms deal with has existed for barely 10 years, and is still very poorly understood. As content moderation scholar Evelyn Douek writes, it is "striking how much we do not know about online speech… we are only at the very beginning of the process of determining what works, outside of the take-down/leave-up paradigm."[47]

Further research isn't just necessary: it is the single most important thing we need to do. No sound policy can be designed without much more information on how people actually respond to different levels and types of content moderation.

There are two enormous gaps in our understanding of both illegal content, and lawful but harmful content and behaviour online – and our government could very productively contribute to both.

The first is a data gap; despite years of pressure, platforms resist requests for them to share data they hold on how their platforms are impacting their users. Data is provided to researchers looking to understand content moderation grudgingly, often incomplete, and withdrawn at the slightest sign of controversy or bad press.[48] Platform users are given obscure, misleading or incomplete accounts of what data platforms hold on them, and how or why their content has been promoted or moderated. As a result, we rely far too much on occasional leaks from platform whistleblowers to understand how platforms are affecting us, both collectively and individually.

Documenting the impact of social media spaces and algorithms on us is much too important to be restricted to internal platform reports, as platforms have a vested interest in burying or minimizing findings that are bad for business. As such, OpenMedia endorses legislating detailed transparency requirements for all major online platforms on how they're moderating content,

---

[47] Evelyn Douek (2021). "Governing Online Speech: From "Posts-as-Trumps" to Proportionality and Probability" *Columbia Law Review* Vol 121 No.3 page 819 (April 2021)

[48] Taylor Hatmaker (2021). "Facebook cuts off NYU researcher access, prompting rebuke from lawmakers" *Techcrunch* August 4 2021.
https://techcrunch.com/2021/08/04/facebook-ad-observatory-nyu-researchers/

including how much and what type of content is removed, appended with fact-checking labels, or consciously downranked.

We recognize some language supporting increased transparency in the consultation paper, but it cannot be provided just to a closed Canadian regulator.[49] Detailed transparency reports must be made available to all users of a platform, and the opportunity to study data and audit algorithms made available to qualified academic researchers, not only government.[50] As Douek writes, the goal should be "to expose to public scrutiny the decision-making process already taking place, so that it can be subject to public argumentation, contestation, and disruption."[51]

The data gap has fed a research gap on questions that are essential to making good decisions moving forward on how to support user expression online while limiting damaging outcomes. There is an overwhelming need for more research on how users are interacting with each other in legal but negative ways, including having negative or toxic interactions, spreading misinformation, and making use of or being failed by content takedown mechanisms. Innovative ideas for approaches that could better balance user expression with mitigating potential harms abound, including making it easier for users to block or hide certain types of posts, warning labels, small nudges to read articles before sharing, or demonetizing certain types of content around important and sensational issues.[52] [53]

We are not recommending any of these approaches; more research is needed to determine whether they're effective to their purpose, and what their side consequences could be. We're pointing to them as examples of areas where more research could reveal rights-protective solutions to some online problems.

Support for research on content moderation and partnerships with platforms is referred to in a single vague mention in the consultation's technical paper.[54] Yet this is a key area that the government could make a meaningful difference to with further attention and support.

Recommendation: Mandate detailed, open and public transparent reporting on how content moderation practices are applied to online platforms.

---

[49] Department of Canadian Heritage (2021). *Technical Paper* Para. 14.
https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/technical-paper.html
[50] Nicolas Suzor, Sarah West, and Jillian York. *"*What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation" *International Journal of Communication 13(1526-1543).*
https://ijoc.org/index.php/ijoc/article/view/9736/2610
[51] Evelyn Douek (2021). "Governing Online Speech: From "Posts-as-Trumps" to Proportionality and Probability" *Columbia Law Review* Vol 121 No.3 page 819-820 (April 2021)
[52]   Evelyn Douek (2021). "Governing Online Speech: From "Posts-as-Trumps" to Proportionality and Probability" *Columbia Law Review* Vol 121 No.3 page 826 (April 2021)
[53] Ben Kaiser, Jonathan Mayer, J. Nathan Matias (2021). "Warnings that Work: Combating Misinformation Without Deplatforming." *Lawfare Blog.*
https://www.lawfareblog.com/warnings-work-combating-misinformation-without-deplatforming
[54] Department of Canadian Heritage (2021). *Technical Paper* Para. 35b.
https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/technical-paper.html

Recommendation: Explore legal requirements to mandate online platforms share content moderation and user engagement data with independent research teams. Ensure any requirements show due consideration for platform user privacy.

Recommendation: Canada should be a global leader in funding research that seeks to better understand the patterns and drivers of both illegal content, and legal but potentially harmful user behaviour and content online.

iii) Empowering internet users; not Big Tech

As Sue Gardner rightly points out, the government's current 'attack' on Big Tech targets symptoms of problems with the modern Internet, not the cause.[55] Most Internet users feel they have very little control over what they see, control, and are able to protect themselves from online.

This is facilitated by a world in which Internet users are data products for online platforms, not communities they are meaningfully accountable to. Online platforms largely make a living buying and selling access to our data, while keeping us on their platform for as long as they can. Illegal content is rarely welcomed by mainstream platforms, but emotionally upsetting and polarizing content that drives high user interaction can be harder for them to turn down. Many Internet users are frustrated by knowing they are being played by algorithms in this way, yet recognize they have little meaningful power to change their online experience.

The government seems to have succumbed to the tempting but deeply misguided approach of stepping in and attempting to assume the role of arbiter of what's good and bad on the Internet. It won't work for many reasons, including that many world governments are currently grappling with the same temptation, and they disagree on what ought to be considered good and bad. But as we've documented in this response, a failed attempt to exercise that enormous governing power could do a great deal of damage to people's speech and experience online before playing itself out.

Smart government regulation should focus on empowering Internet users to retake control of their own respective online experiences, and effectively pressure Big Tech platforms.

First, it needs to be made much easier for Internet users to leave a platform, taking all their personal data with them, without severing their ties with friends and family left on the platform. This means bringing back a version of the last parliament's *Bill C-11, An Act to enact the Consumer Privacy Protection Act and the Personal Information and Data Protection Tribunal Act*

---

[55] Sue Gardner (2021). "The crackdown on 'Big Tech' targets symptoms rather than the disease itself" *Globe and Mail* May 21, 2021.
https://www.theglobeandmail.com/opinion/article-the-crackdown-on-big-tech-targets-symptoms-rather-than-the-disease/

and to make related and consequential amendments to other Acts, and patching its many holes to make it strong and effective legislation, such that users have a strong and easily actionable right to access, modify, delete, and transfer their data held by any company.

It is striking that despite Bill C-11's introduction in November 2020, promising major reform of our privacy rights in the private sector, the government made no effort to actually pass the Bill, let alone fix the many loopholes and areas that needed tightening identified by privacy experts. We hope to see that change in our next Parliament.

If effective user control of our data was combined with strong transparency and research requirements, platforms would find themselves with a user base that can easily leave a platform they're dissatisfied with. This would allow users themselves to effectively pressure platforms to reform themselves if their content moderation or privacy standards are not adequate.

Second, a hard, data- and research-driven look needs to be taken at whether an advertising and 'time spent on platform' business model is compatible with the needs of healthy democratic discourse. This business model demands engagement above all else, and that includes a lot of deeply negative engagement. Without addressing the underlying business models and incentives, the problems the government aims to tackle here will remain fundamentally unsolved.

Throughout, the government must consider whether their approach is encouraging a reduction of major platform power, or reinforces and depending on it. A recurring theme in scholarly discussion of the power of Big Tech is the need to avoid regulatory 'lock-in' of their power and prominence.[56] [57]
Expensive and complex regulatory obligations make it difficult for new online platforms to compete with the handful of platforms that dominate our Internet today. That's why they can be surprisingly popular with some of the largest entrenched platforms.[58]

But careful government legislation could erode that dominance. Some online platform dominance comes from making good products, but much of it comes from translating early leads in the market into runaway network effects. The more people use a given platform, the more valuable being on that platform becomes. Over time, online platforms have converted their platforms into so-called 'walled gardens' – trapping many users who do not necessarily approve of their practices or want to be on their service.

---

[56] Cory Doctorow (2021). "Competitive Compatibility: Let's Fix the Internet, Not the Tech Giants" *Communications of the ACM*, Vol 64. No. 10 (October 2021).
https://cacm.acm.org/magazines/2021/10/255710-competitive-compatibility/fulltext
[57] Evelyn Douek (2021). "Governing Online Speech: From "Posts-as-Trumps" to Proportionality and Probability" *Columbia Law Review* Vol 121 No.3 page 829-830 (April 2021)
[58] Amanda Macias (2020). "Facebook CEO Mark Zuckerberg calls for more regulation of online content" *CNBC* Feb 15 2020.
https://www.cnbc.com/2020/02/15/facebook-ceo-zuckerberg-calls-for-more-government-regulation-online-content.html

There are many innovative ideas currently being explored about how content moderation could be done better on a less Big Tech centered web. Recent proposals have included separating content moderation from platform responsibility as an independent form of 'middleware'; developing competitive content moderation protocols that individual users can adopt to provide the type of protection they want on the web; encouraging 'trusted flaggers' systems, in which users whose reports of illegal content are consistently valid have expedited processing time; and instituting a duty of care on platforms for their users, less focused on case by case outcomes and more systematically evaluative of how they approach their overall responsibility to user safety and wellbeing.[59][60][61][62]

We are not endorsing any of these approaches; more research is needed to evaluate their potential effects. But they should at least be considered in a more appropriately open and thoughtful future consultation from the government.

Recommendation: Empower Internet users against Big Tech. Give them the rights they need to leave platforms they don't like, and they can hold platforms accountable themselves to moderate content responsibly.

Recommendation: Many lawful but harmful online behaviour and user experiences are driven by an ad-centric business model that works to keep users on a platform at any cost. Solicit ideas for regulatory remedies that would discourage the proliferation of this model.

Recommendation: Ensure that any new regulation discourages the centralization and concentration of online platforms.

D. Conclusion

This response is not wholly comprehensive of OpenMedia's concerns with the government's apparent intended direction for our Internet; it is only the beginning of that conversation.

We are a community that is immensely passionate about the tremendous liberating power of the open Internet. That does not make us enemies of all ideas for regulating it, as we trust we've made clear.

---

[59] Francis Fukuyama (2021). "Making the Internet Safe for Democracy." *Journal of Democracy* Vol 32(2):37-44.
https://www.journalofdemocracy.org/articles/making-the-internet-safe-for-democracy/
[60] Mike Masnick (2019). "Protocols, not Platforms: A Technological Approach to Free Speech." *Knight First Amendment Institute at Columbia University.*
https://knightcolumbia.org/content/protocols-not-platforms-a-technological-approach-to-free-speech
[61] Darryl Carmichael and Emily Laidlaw (2021). *The Federal Government's proposal to Address Online Harms: Explanation and Critique*
https://ablawg.ca/2021/09/13/the-federal-governments-proposal-to-address-online-harms-explanation-and-critique/
[62] Daphne Keller (2020). "Systemic Duties of Care and Intermediary Liability." Blog entry:
http://cyberlaw.stanford.edu/blog/2020/05/systemic-duties-care-and-intermediary-liability

We whole-heartedly reject the very nature of this consultation and manner in which it has been held, in addition to the specific proposals currently being proposed. As we've laid out, this proposal will overwhelmingly censor lawful speech more than illegal content, produce an unprecedented surveillance funnel of lawful speech to law enforcement, and hurt marginalized communities far more than it will help them.

But we've also highlighted many measures our government could take now that would genuinely take on and roll back the power of major online platforms, while contributing to a healthier and less hateful Internet. These include strong data ownership, research and transparency reporting changes that would make platforms far more accountable to their users, and highlighting some interesting ideas for more innovative content moderation models that would make a more genuine future public consultation on these issues far more fruitful.

We'll close with a fundamental question: what do Canadian Internet users actually want for our Internet? How would they like to see their rights defended, and their content moderated?

It is unclear that our government wants to find out, preferring to use single answers to generic poll questions to justify their current intentions. That needs to change.

Due to the nature of this consultation, and the short timeline, this submission only represents a fraction of our community's concerns and perspectives. But we hope they have helped to highlight just how damaging this current proposal is to not only the internet in Canada, but globally.

We represent some of the most concerned and engaged people in Canada on these issues, and will continue to make ourselves heard by our elected officials and representatives, whether the government provides appropriate formal opportunities to do so, or not.