September 24, 2021

Department of Canadian Heritage ("PCH")

Re: Google Submission to Canadian Government's Proposal to Address Online Harms

I. Executive Summary

We appreciate the opportunity to provide comments on the Department of Canadian Heritage's proposals to address certain categories of harmful content online. These are important issues that require thoughtful input from a variety of stakeholders, including online service providers, the Canadian government, civil society, and others.

We are supportive of the government's efforts to find ways to protect Canadians from online harms; however, we are concerned that some aspects of the current proposal could be vulnerable to abuse and may have unintended negative impacts on Canadians' access to valuable information and services, privacy and freedom of expression, and the Canadian economy. We have summarised these concerns below and provide further detail in the rest of our submission.

- **The types of providers and services that are in and out of scope must be clearly identified, recognising the distinct nature of different types of services and user interactivity, differing abilities to moderate content, and the impact on access to information**.
    - We agree with the government's efforts to exclude certain types of services from the definition of Online Communication Service Provider (OCSP) (e.g., private communication services, telecommunications services), and we encourage it to make these exceptions more clear to avoid creating ambiguity about the types of services it considers in scope of the proposed framework.

- **Obligations must be limited to illegal content to avoid spurring the unnecessary removal of lawful, legitimate content.**
    - We believe it is critical that content regulated by the proposed framework be precisely defined and limited to illegal content in order to avoid undermining access to information, limiting freedom of expression, restricting the exchange of ideas and viewpoints that are necessary in a democratic society, and creating a legal framework that could be used to censor political speech in the future. The government should take care to ensure that their proposal does not risk creating different legal standards for online and offline environments, making

legal expression offline illegal to share online. In addition, the government should avoid creating a system that drives OCSPs to adopt a "take down first, ask questions later (or never)" approach. Therefore, we urge the government to be extremely clear and precise when defining the prohibited categories and to give due consideration to the time-pressured circumstances in which OCSPs will be expected to apply these definitions to large volumes of content. Furthermore, we believe that it is essential that the government hew to existing definitions for illegal content under Canadian law in order to avoid restricting lawful expression and potentially undermining the legal validity of the framework.

- **In order for illegal content to be removed expeditiously, formal legal complaint systems must be distinct from systems to address community guidelines violations. Rigid 24-hour deadlines for taking action against reported content do not allow providers to carefully assess the relevant law and context and would be counterproductive.**
    - We agree that OCSPs should act promptly to remove illegal content when they become aware of it. However, it is critical that any legal obligations for content removal account for the nuance that is often required for these reviews and determinations, the potential for user error, the need to triage particularly egregious content, and the sheer volume of content and complaints that OCSPs need to process daily. Therefore, we urge the government to establish a flexible process for addressing illegal content that allows OCSPs to adequately evaluate removals requests within a reasonable timeframe (e.g., "without undue delay," or "expeditiously"). Moreover, any content removal legal obligations should be separate and not displace the voluntary "flagging" systems for legal, but harmful content that many OCSPs have created to address the unique needs of their products and services. We also encourage the government to clarify that OCSPs are empowered to take action against users who abuse flagging or legal notice systems and encourage the government to consider other safeguards that could be built into the framework to further deter misuse and abuse of flagging systems.

- **Mandatory obligations to proactively monitor and identify content across the entire service are disproportionate and will result in the blocking of legitimate content.**
    - While automated systems can be a vital tool for detecting and blocking potentially harmful content at scale, such systems often struggle with the application of nuanced, context-dependent definitions for prohibited content. Therefore, mandating that OCSPs use automated systems to proactively monitor and block content would likely lead to the blocking of large amounts of

legitimate content and undermine Canadians' access to valuable information. We strongly encourage the government to clarify that the use of automated systems for proactive monitoring and blocking of content is not required and should be used in conjunction with human review. This would not preclude OCSPs from taking measures on their own initiative, where appropriate and where technologically feasible.

- **Requirements to disclose user data to law enforcement agencies must be accompanied by due process safeguards to prevent the risk of unwarranted government surveillance and of encroaching on users' privacy rights.**
    - We understand the legitimate needs of law enforcement, and we are supportive of OCSPs making voluntary reports to law enforcement regarding illegal content and assisting law enforcement with judicially authorized production requests. However, we are concerned that some of the proposed framework's reporting obligations may undermine due process and privacy protections, as well as directly conflict with legal obligations applicable to OCSPs in other jurisdictions. We encourage the government to reconsider the law enforcement reporting provisions and include appropriate statutory protections for privacy and due process.

- **The obligation to include demographic data in regular reports to the DSC is impractical and may undermine user privacy.**
    - If the demographic reporting requirement were included in the framework, OCSPs would effectively be forced to start collecting additional sensitive data about Canadian users, contrary to user privacy interests and data minimization principles. It would also create an ongoing privacy risk for Canadians by forcing OCSPs to indefinitely retain detailed demographic data about all of their Canadian users, some of whom could be harmed if their sensitive demographic data were to become public as a result of a data breach. Given the significant risks associated with the mandatory collection of demographic data, we urge the government to remove the demographic data reporting obligation from the proposed framework.

- **Regulatory oversight and enforcement should focus on systemic failures rather than individual cases of non-compliance so as to avoid stifling access to information, free expression, and innovation.**
    - We recognize the need for appropriate sanctions for noncompliance with the law. However, we are concerned that the government's expansive enforcement powers and the open-ended nature of the framework's penalty provision will create enormous legal risk for OCSPs. For example, these provisions could

result in OCSPs being subject to financial penalties up to 5% of global revenue for mistakes they make with respect to individual pieces of content -- even when acting in good faith and under robust compliance procedures. Given the vast amount of content that OCSPs process, the nuanced consideration that is often required to identify prohibited content, and the short deadline for addressing flagged content, it is a near certainty that OCSPs will not be able to achieve perfect compliance with the law with respect to each piece of content. These risks will effectively force OCSPs to err on the side of blocking more content than reasonably required and thereby undermine users' ability to share legitimate content and express themselves. Therefore, we urge the government to clarify and expand the due diligence defence and consider an alternative penalty framework that focuses on systemic compliance with the law.

- **To avoid the unnecessary blocking or removal of lawful, legitimate content, financial and criminal penalties must be applied reasonably and proportionately.**
    - As discussed above, the risk of severe penalties for OCSPs that operate in good faith may pressure OCSPs into adopting imprecise and overly restrictive content moderation strategies that will deny Canadians a full opportunity to share and view legitimate content. In addition, we believe that associating penalties with an OCSP's gross global revenue results in penalties that are disconnected from the OCSP's activities in Canada and further disconnected to the reality of their potential presence in the Canadian marketplace. In order to avoid these risks, we urge the government to provide strong safeguards in the legislation that will assure that monetary penalties are imposed in a reasonable and proportionate manner.

II. Google and YouTube's approach to content moderation

At Google, our mission is to organise the world's information and make it universally accessible and useful. We build tools to benefit society, and that have been a force for creativity, learning and access to information. They have enabled economic growth, boosted skills and opportunity, and fostered a thriving society. Google's products alone support $1.7 billion CAD annually in incremental exports for Canadian businesses and are equivalent to 1.1% of GDP or supporting 240,000 local jobs. [1] In 2020, Oxford Economics found that YouTube's creative ecosystem contributed approximately $923 million to Canada's GDP and supported more than 34, 000 Canadian jobs.[2] In addition, YouTube has helped Canadian creators of all kinds, both

[1] Public First: Google Canada Economic Impact Report 2019.
[2] Oxford Economics: From Opportunity to Impact: Assessing the Economic, Societal, and Cultural Impact of YouTube in Canada.

amateur and professional, reach a global audience. In fact, Canadian creators see 90 percent of their views come from outside Canada's borders.

While we believe the Internet has an immensely positive impact on society, we also recognise that there can be a troubling side of open platforms, and that in some cases bad actors have exploited this openness. We understand the sensitivity and importance of these areas and have devoted careful attention to developing an approach that limits harm while protecting users' ability to express themselves online. We have not waited for legislation to act in tackling illegal or lawful, but potentially harmful content; we have developed our own guidelines and taken action. We have implemented extensive efforts to help prevent and address harmful and unlawful content across our services, including by working appropriately with government, law enforcement, and other stakeholders in Canada and around the world.

Our approach for moderating content and providing our users with access to high-quality information centres on four complementary levers:

- **Remove**: We comply with legal obligations requiring the removal of unlawful content with clearly defined processes for users and governments to submit legal complaints about our products. In addition, we set responsible and clear rules for each of our products and services and take action against content and behaviours that infringe on them.

- **Raise**: We elevate high-quality content and authoritative sources where it matters most.

- **Reduce**: We reduce the spread of potentially harmful information where we feature or recommend content.

- **Reward**: We set a high standard of quality and reliability for publishers and content creators who would like to monetize or advertise their content.

Our strategy for tackling illegal and potentially harmful content is tailored to each of our platforms. We have processes by which governments and individuals can request removal of illegal content, including reporting violations of country-specific laws, such as those related to anti-terrorism, obscenity, or hate speech. Legal removals processes require detailed, specific information about the nature of the potentially illegal content. We review these requests closely to determine if content should be removed because it violates a law or our community guidelines and policies.

In addition, for each product, we have a specific set of rules and guidelines that are suitable for the type of platform, how it is used, and the risk of harm associated with it. For example, on

YouTube these approaches range from clear community guidelines, with mechanisms to report content that violates them, to increasingly effective artificial intelligence (AI) and machine learning that can facilitate removal of harmful content before a single human user has been able to access it. In April 2021 we introduced a new metric, called Violative View Rate, as part of our quarterly transparency reporting. This metric estimates that the proportion of views of YouTube videos that violate our Community Guidelines has fallen from c. 0.7% in Q4 2017 to c. 0.19-21% in Q2 2021. We calculate this metric using a rigorous statistical methodology, which has just been reviewed and validated by MIT Professor Arnold Barnett.[3]

Our goal is to achieve both accuracy and scale in our work. That's why we have people and technology working together - and we invest heavily in both. We now have over 20,000 people across Google and YouTube dedicated to keeping our users safe from policy development to review and enforcement. This includes reviewers who work around the world across all time zones, speak many different languages, and are highly skilled. On YouTube, for example, reviewers evaluate flagged videos against all of our Community Guidelines and policies, regardless of why the video was originally flagged

While we have made tremendous progress in developing automated systems to detect harmful and illegal content, machine learning and other technologies are still in development. In some instances, automated proactive measures cannot properly take the context of content into account. Machine learning models are not yet consistently good at understanding contextual differences between content that otherwise looks very similar. As a result, automatically removing content is not necessarily the correct decision in every circumstance. In addition, recent research has also shown that even small changes to images can fool computer vision systems into missing what is obvious to human reviewers. Proactive measures are improving all the time, but they should only be deployed carefully, and when judged effective by individual companies.

We continue to invest in developing and improving the policies, products, tools, processes, and teams that handle content moderation across our platforms and are committed to providing trustworthy, useful information that meets our users needs and protects them from harm.

III. Covered Entities - The types of providers and services that are in and out of scope must be clearly identified, recognising the distinct nature of different types of services and user interactivity, differing abilities to moderate content, and the impact on access to information.

[3] Arnold Barnett, YouTube's Violative View Rate Methodology: A Statistical Analysis (2021), available at https://storage.googleapis.com/transparencyreport/youtube/YouTube%27s%20VVR%20Methodology%20-%20A%20Statistical%20Assessment%20-%20Arnold%20Barnett.pdf.

We agree with the government's efforts to exclude certain types of services from the definition of OCSP (e.g., private communication services, telecommunications services), and we encourage it to make these exceptions more clear to avoid creating ambiguity about the scope of the proposed framework.

Because the proposed framework could require OCSPs to view and monitor certain user content, the definition of OCSP should expressly exclude services (and parts of services) where such access or monitoring is technically infeasible, would be highly intrusive to user privacy, may unreasonably limit access to high-quality information online, or harm free expression and creativity. In particular, we believe it is important that the following types of services be more clearly excluded from the definition of OCSP:

A. Cloud storage providers

Cloud providers are limited in what they can do to address illegal content stored at the direction of their customers or their customers' users, given the technical architecture of their services, privacy protections, and the contractual obligations they hold towards their customers' data. Factually and contractually, such providers do not have the requisite authority and control over content, such that they should have responsibility for removing specific content from a third party's service. Our understanding is that the technical paper's statement that "[the OCSP definition] should not include a person who…hosts or caches the content or information about the location of the content, by reason only that another person uses their services to provide an OCS"[4] would prevent many cloud storage providers from qualifying as OCSPs, and we urge the government to make that point clear in legislative text.

For example, customer data may be encrypted in a manner that allows only the customer to access the data and the cloud storage provider may be contractually prohibited from accessing it. In addition, cloud services are also regularly used by government institutions, research organizations, civil society groups and universities. Placing this category of services in-scope of the definition of OCSP would require monitoring the content of such organizations. Finally, many cloud storage services, including those that directly serve consumers, generally do not make the content they store accessible or searchable to the general public. The absence of general public access and search features inherently limits the potential reach of content that is stored by cloud storage services.

Subjecting cloud services to the proposed framework would raise significant user privacy and business confidentiality concerns, among other harms. For example, the main purpose of many of these services is to allow individual consumers to store personal content. Although some users may use cloud storage services to share content with others (e.g., by sharing a link to a stored file with a limited set of other users), such sharing is often more akin to a private

[4] Technical paper, Module 1(A), 4.

communication (which are expressly exempted from the proposal) than to the widespread public distribution of content that is possible on social media services. Though some OCSPs carry out automated hash-matching of media in cloud storage, what is called for in the framework involves much more than this automated analysis. As a result, we urge the government to clarify that all cloud service providers are also excluded from the definition of OCSP. Short of that, any obligations that are placed on cloud service providers should account for the constraints on their ability to access and monitor user content.

B. Search engines

We agree that "[the OCSP definition] should not include a person who indicates the existence or location of content,"[5] including search engines. Search engines play a critical role in organizing information and making it accessible to the public. They are indexes of the web at large and consist of the automatic and intermediate storage of information hosted by third parties. Given the immense volume of information that search engines process (e.g., hundreds of trillions of pages), it would be impossible for them to substantively evaluate the nature of the content they index while continuing to operate at their current scale. The content, even if it could be evaluated, would remain available on the website where it is hosted. As a result, the law has importantly ensured that responsibility rests with the platforms and webhosts that have control over the content and can determine whether it is available to the public in Canada. To ensure that search engines can continue to provide accurate and up-to-date access to the vast amount of information available on the Internet, we recommend that they continue to be expressly excluded from the definition of OCSP.

IV. Content in Scope - Obligations must be limited to clearly defined categories of illegal content to avoid spurring the unnecessary removal of lawful, legitimate content.

A. Overbroad definitions of regulated content may limit freedom of expression and lead to over-removal of lawful content

We applaud the government's overall goal of combating the spread of harmful content online. At the same time, we also believe it is critical that content regulated by the proposed framework be precisely defined and limited to illegal content in order to avoid creating a framework that spurs the over-removal of content, undermines access to information, limits freedom of expression, restricts the exchange of ideas and viewpoints that is necessary in a democratic society, and could be used to censor political speech in the future. We are concerned that the expansive and subjective content definitions proposed in the framework will make it difficult for OCSPs seeking to comply in good faith to make accurate decisions

[5] Technical paper, Module 1(A), 4.

promptly (especially considering the proposed 24-hour deadline for addressing user-flagged content, for which we separately express additional concerns below).

For example, the technical paper states that "[t]he concept of terrorist content, should refer to content that actively encourages terrorism and which is likely to result in terrorism."[6] The application of this brief definition can require considerable analysis, as it requires OCSPs to consider: (1) whether the content relates to "terrorism," a term that has a fairly broad and complex definition under Canadian law; (2) whether the content is meant to "actively encourage" terrorism; and (3) whether it is "likely to result in terrorism." Faced with the pressure of having to proactively monitor vast amounts of information for prohibited content and to quickly remove and/or report prohibited content, many OCSPs will not be able to give these questions the thoughtful consideration they require. Instead, they will most likely resort to blocking/removing any content that has a remote possibility of qualifying as prohibited content, resulting in the suppression and censorship of lawful expression (potentially including content that is intended to educate and inform the public about terrorism).

Consider, for example, a livestream of a political rally about climate change that is filmed by a bystander and uploaded to a social media website. While the majority of the speakers at the rally argue for activism through peaceful means, one speech raises the idea of destroying fossil fuel infrastructure in order to combat climate change. If committed, such an act could constitute "terrorist activity" under Canadian law. Therefore, an OCSP could potentially conclude that the video of speech "actively encourages terrorism" and, if the speaker is deemed to be persuasive, is "likely to result in terrorism."[7] Even though the bystander simply meant to raise awareness of the climate change rally and had no intention of promoting the illegal activity advocated by the one speaker, the social media site may conclude that it is required to remove the entire video and report the content and bystander to authorities. The framework also does not take into account other important considerations. For example, it does not answer the question of whether the analysis changes if the video is uploaded by a journalist. This is just one example of the proposed framework's broad definitions of prohibited content that could effectively force OCSPs to try to make nuanced decisions about the intent and impact of content at an unprecedented and infeasible scale.

Another concern is how an OCSP should handle human rights matters. For example, many individuals in Syria documented war crimes and uploaded the videos to social media sites. These videos were initially flagged by automated systems. Preserving them, however, was important for international prosecutors, human rights organizations, and Syrian citizens who

[6] Technical paper, Module 1(A), 8.
[7] The content could also fall into the similarly broad category of "content that incites violence," which is defined as "content that actively encourages or threatens violence and which is likely to result in violence."

aimed to hold the perpetrators accountable. Because YouTube and other social media platforms were able to retain these types of videos, they have been used as evidence in criminal cases that have resulted in convictions.

B. Limiting the framework to categories of content that are illegal under existing Canadian law will provide clear rules and expectations for OCSPs and users

The definitions of content in scope should be directly tied to and limited to content that has been found by Canadian courts to be unlawful after a thorough review through a *Charter*-informed lens. Not doing so will likely result in legislation being found unconstitutional and will have a chilling effect on lawful expression. For example, the technical paper proposes that content related to child sexual abuse be extended to include "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 (e.g., screen shots of videos that do not include the criminal activity but refer to it obliquely; up-to-date photos of adults who were exploited/ abused as children being posted in the context of their exploitation and abuse as children)."

Most Canadians are familiar with the tragic stories of two young Canadian women who were both victims of sexualized cyberbullying and child pornography offences. After their tragic deaths, their parents bravely took on significant roles of public information and advocacy for victims, telling the stories of their daughters in order to bring about significant, positive change in our communities, our schools and in legislatures. Not surprisingly, they made extensive use of social media to reach young people with their stories of their daughters, spreading messages of respectful relationships and online safety. In one case, the victim created a video on YouTube in which she told her own story of online abuse and the impact it had on her. The definition proposed in the framework could reasonably be interpreted as requiring the removal from OCSPs of content that involves survivors and their families telling their stories for educational purposes. The definition could also be applied to OCSPs who carry public testimony from the Parliamentary committee's study that informed the framework. Ensuring that the definitions do not inadvertently silence these voices online is beneficial and completely aligned with the objective of reducing the prevalence of online material that harms children and child victims.

Given the risk of the suppression of legitimate and lawful expression, we urge the government to be precise in defining the prohibited categories of content, limit the definitions to what Canadian courts have deemed to be unlawful and to account for the fact that OCSPs will be under pressure to review enormous volumes of content and make quick determinations of whether the content falls into a prohibited category.

As mentioned below in Section V, it is also important that illegal content has a separate legal removals system from voluntary systems OCSPs may create for their users to flag lawful content that violates their products' community guidelines. Requiring formal legal notice for removals of illegal content, would ensure that violative content is removed as expeditiously as possible. Formal legal notice would provide OCSPs with details about the illegal nature of the content as well as sufficient information about the identity and location of the individual or entity reporting the content. Additionally, OCSPs have the benefit of evaluating these illegal removals requests against clear legal standards set forth in criminal codes.

V. Obligations for OCSPs - Rigid deadlines for taking action against reported content do not allow providers to carefully assess the relevant law and context.

We agree that OCSPs should act promptly to remove illegal content when they become aware of it. However, any legal obligations for content removal should account for the nuance that is often required for these reviews and determinations, potential for user error, and the sheer volume of content and complaints that OCSPs need to process on a daily basis. The proposed framework's 24-hour deadline for addressing all user-flagged content fails to take these realities into account and should be removed. Additionally, treating all user flags as triggers for a legal takedown obligation (including the running of the 24-hour deadline) will inevitably make the system vulnerable to abuse and lead to the removal of legitimate content.

A. 24-hour deadline

As discussed in section IV, OCSPs may need to engage in a nuanced consideration of context, intent, and impact in order to determine whether content meets the definitions of one of the five categories of prohibited content. Given the potential breadth of the prohibited categories, "grey-area" cases will undoubtedly be common and the 24-hour deadline will not allow sufficient time for thoughtful consideration of the case (as we have noted separately, the definitions for prohibited content should also be tied to existing definitions for illegal content under Canadian law).

The problems associated with an extremely short takedown timeline will only be compounded by the fact that any user flag can trigger the start of the countdown. OCSPs that have millions of users and host billions of pieces of content could easily receive tens or hundreds of thousands of flags per day. For example, over 500 hours of video are uploaded to YouTube every minute. In the second quarter of 2021, users submitted 17,226,571 flags (around 190,000 a day) to YouTube about content that allegedly violated community guidelines. In the face of high volumes of flags, OCSPs would need to rely on automated systems for processing, which, as discussed above, struggle with making nuanced content classification decisions.

In addition, confronted with the short deadline and prospect of extremely high penalties for noncompliance, many OCSPs will choose to prioritize speed over accuracy and automatically block/remove content that is subject to a flag if their automated system concludes there is even a remote possibility that the content is prohibited. As a result, significant amounts of legitimate and lawful expression that was either incorrectly flagged by a user[8] or mischaracterized by an automated system will be removed. While some such content could potentially be reinstated through the proposed framework's mandated appeal process, this would not eliminate the risk that Canadians would be denied access to valuable information online. Some content, for example, may be time-sensitive (e.g., news coverage of a recent event) and the removal of such content during the relevant time period would greatly undermine its value. Other content may not be appealed, in which case the legitimate and lawful expression will remain censored.

This short deadline to address takedown requests also raises considerable issues related to innovation and competition among OCSPs of differing sizes, and has the potential to stifle innovation and growth of Canadian OCSPs. Being able to address takedown requests will require personnel and other resources that are often in short supply within start ups and rapidly growing companies. While large, established companies -- particularly those that already take harmful content seriously -- will have people and processes that will have to be deployed to comply with a Canadian framework, smaller companies simply do not have these resources. This framework will immediately create a disincentive for the creation of Canadian OCSPs and overburdening the resources of smaller companies will compound the incentive to simply take down content that is at all questionable, but perhaps lawful. Furthermore, it is foreseeable that new, emerging OCSPs will simply forgo making their services available to Canadians.

It is worth noting that other democracies have avoided or pushed back against short removal deadlines for content moderation rules in recognition of the practical difficulties associated with the deadlines and their potential negative impact on consumers' right to access information and freedom of expression. For example, Germany's Network Enforcement Law (NetzDG), which includes strict content removal deadlines, only requires a 24-hour turnaround time for *"manifestly unlawful" content* and allows an extension from 24 hours to 7 days for more complex cases, as well as additional time for decisions that require specific legal expert knowledge and are referred to a joint industry body.[9] Similarly, in France, a 2020 bill with a 24-hour removal mandate was struck down by the French Constitutional Council over

[8] In Q2 2021, users submitted 17,226,571 flags to YouTube, and in the same period only 351, 570 videos were removed as a result of user flags.
[9] Network Enforcement Act (Netzdurchsetzunggesetz), Section 3.

concerns about the chilling effect the bill would have on free expression by incentivising intermediaries to remove legal speech in an effort to remain compliant.[10]

As an alternative to the overly brief and rigid 24-hour deadline set out in the proposed framework, we urge the government to consider more reasonable, flexible standards that would still require OCSPs to address reported content with urgency. For example, a more workable standard could be to require OCSPs to address reported content "with all due speed," "without undue delay," or "expeditiously." This would allow the company to carry out appropriate consideration and seek expert guidance, while prioritizing the most important cases. Regulators could also issue guidance or best practices that give a sense of the typical timelines in which OCSPs should generally seek to address reported content. The proposal could also include "stop-the-clock" safeguards that allow OSCPs to pause the countdown to the deadline when they require more information to evaluate the complaint.

B. User-submitted flags

A separate, but related problem with the obligation to address user flags within 24 hours is the fact that user-submitted flags are often inaccurate and can be used as a tool to harass and infringe on the expression of other users. Our experience with the YouTube community guidelines flagging tool illustrates this risk. We receive hundreds of thousands of content flags on a daily basis. While many are good-faith attempts to flag problematic content, large numbers of them represent mere disagreement with views expressed in legitimate content or are inaccurate. These types of user flags are best used as "signals" of potentially policy violative content, rather than definitive statements of violations, and should not be treated as flags that trigger specific legal obligations. It is critical that OCSPs have discretion to review and use such flags in ways that make the most sense to protect their users (e.g., evaluating flags in conjunction with technical signals and other factors to prioritize reviews of flagged content).

Our experience with Germany's NetzDG law provides similar evidence about the inaccuracy of user flags even in the context of a legal complaint system. Our current NetzDG transparency report shows that more than 84% of content reported under the NetzDG was determined not to violate our Community Guidelines or the criminal statutes referred to in NetzDG and was therefore not removed or blocked.[11]

---

[10] Decision n° 2020-801 DC of June 18, 2020, available at https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm.
[11] Removals under the Network Enforcement Law, available at: https://transparencyreport.google.com/netzdg/youtube?hl=en.

Given the high risk of inaccurate user flags, we urge the government to consider alternative approaches, such as requiring users to submit a legal complaint. If users were to submit such a complaint, they would be required to provide the legal grounds for the removal, their identification, and precise location. This standard has been successfully implemented in regulations across the globe, including most recently in France. Adding more specificity to the user reporting process would not only increase the likelihood that users will report actionable content but also provide us with the information we need to evaluate the content fairly and quickly.

We encourage the government to consider permitting OCSPs to require that users provide detailed information about the nature of their report-- if they are claiming that the content is prohibited under Canadian law. For example, a formal report pursuant to the Canadian framework should require the user to:

- identify themselves;
- clearly identify the content at issue by URL, video timestamp, or other unique identifier.
- state the law and basis of the legal claim (e.g., explain why the content meets the definition of one of the prohibited categories of content); and
- attest to the good faith and validity of the claim.

The government and OCSPs could collaborate to provide guidance and educational resources in order to help users understand the nature of the law and complete these requests. However, we believe that it is important to maintain a distinction between complaints that trigger significant legal requirements and the simple 'click to flag' buttons that are used for community guideline violations that may not have legal implications. Requiring users to go through additional steps to submit a legal complaint would highlight the significance of the action and potentially deter abuse of the system. Reducing the number of incorrect or abusive complaints submitted pursuant to the legal reporting requirement will also enable OCSPs to spend more time on legitimate complaints and help them block prohibited content in a timely manner.

Alternatively, notice could be limited to removal requests submitted by certain trusted organizations. For example, YouTube has developed a Trusted Flagger program to help provide robust tools for individuals, government agencies, and non-governmental organizations (NGOs) that are particularly effective at notifying YouTube of content that violates our Community Guidelines. The program provides these partners with training, a bulk-flagging tool, and a channel for ongoing discussion and feedback about YouTube's approach to various content areas. The program is part of a network of more than 180 academics, government partners, and NGOs that bring valuable expertise to our enforcement systems. For instance, to help address violent extremism, these partners include the International Centre for the Study of Radicalization at King's College London, the Institute for Strategic Dialogue, the Centre for Israel and Jewish Affairs, the National Council for Canadian Muslims and government agencies

focused on counterterrorism. Because their flags have a higher action rate than the average user, we prioritize them for review.

Lastly, we encourage the government to clarify that OCSPs are empowered to take action against users who abuse flagging or legal notice systems. For example, under flagging systems that platforms have voluntarily established, platforms have the ability to ban users who repeatedly make false reports. The framework should provide OCSPs with a safe harbour from liability for actions they take to ban or otherwise penalize users who misuse any legally mandated flagging system. In addition, we encourage the government to consider other safeguards that could be built into the framework in order to further deter misuse and abuse of flagging systems.

VI. Obligations for OCSPs - Mandatory obligations to proactively monitor and identify content across the entire service are disproportionate and could result in the blocking of legitimate content.

We are supportive of OCSPs voluntarily implementing robust systems to identify and address harmful content. We are concerned, however, about the potential negative consequences of the proposed framework's broad requirement that OCSPs "take all reasonable measures, which can include the use of automated systems, to identify harmful content that is communicated on its OCS and that is accessible to persons in Canada, and to make that harmful content inaccessible to persons in Canada."[12] Specifically, we are concerned that some could seek to interpret this language as a mandatory obligation to implement automated systems that proactively monitor and block prohibited content. As discussed in more detail below, doing so would create a series of significant negative ripple effects for Canadian users. For example, given that it is often difficult for automated systems to determine whether content falls into highly context-dependent categories (e.g., the five prohibited categories of content under the framework), the required use of such systems would likely result in the over-blocking of content and Canadians losing access to valuable content and information. Additionally, bad actors may seek to exploit weaknesses in these automated systems in order to intentionally censor legitimate content (e.g., political speech, speech by minority groups). To avoid these negative outcomes, we encourage the government to clarify that no part of the framework mandates the proactive monitoring and filtering of content.

While breakthroughs in machine learning and other technology used to monitor and identify potentially harmful content are impressive, the technology is still evolving and is less accurate for more nuanced or context-dependent content. For example, automated systems that are trained to recognize certain images or patterns of text that may be associated with categories of prohibited content (e.g., terrorist content, hate speech) may mistake news coverage,

[12] Technical paper, Module 1(B), 10.

documentaries, educational material, and academic research of these subjects as prohibited content because they contain some of the same images and text.

Consider a video of military conflict. In one context, the footage might be documentary evidence of atrocities in areas that journalists have great difficulty accessing. In another context, the footage could be promotional material for an illegal organisation (e.g., a terrorist organisation). And in another, important political speech by marginalized populations. In the same vein, the exact same iconic and horrifying images of historic genocide are used by those who want to advocate for justice and tolerance, on one hand, and those who advocate for violence and further genocide, on the other hand. Between these two poles are those who aspire to report on historic events in an objective manner. Computers cannot yet distinguish this key context. Even a highly trained reviewer could have a hard time telling the difference, and machines are even more limited.

Similarly, while automated systems can make it easier to prevent known violative content from being re-uploaded, they have limitations here as well. For example, on YouTube, we use digital hash technology to catch copies of known violative content before it is available to view. For some content, like child sexual abuse images and terrorist recruitment videos, we contribute to shared industry databases of hashes to increase the volume of content our machines can catch at upload. This technology generally works well when *exact* copies of, for example, the same terrorist propaganda video are re-uploaded. In contrast, an automated tool may have difficulty detecting the same video if it has been subject to minor alterations.

The accuracy limitations of automated systems can also be seen in data we maintain about appeals on YouTube. From April-June 2021, we received 217,446 requests for appeal, an increase from the previous quarter; of those, 52,696 videos were reinstated.[13] During the onset of the COVID-19 outbreak, there was an increase in successful appeals which may have been attributable to an increased deployment of machine learning to tackle challenging content during that period, and thus reinforces the view that machine automation simply cannot replace human judgment which requires time for proper analysis and deliberation.

In addition to potentially blocking legitimate speech, mandatory proactive monitoring requirements may also stifle innovation and competition in the OCS industry in Canada. Building and implementing automated systems to monitor content can entail substantial costs and engineering, legal, and trust and safety resources. Small companies and startups may be deterred from entering the OCS market in Canada if they are unable to bear these costs.

[13] Google Transparency Report, available at: https://transparencyreport.google.com/youtube-policy/appeals?hl=en&total_removed_videos=period:2020Q1;exclude_automated:all&lu=total_videos_reinstated&total_videos_reinstated=period:2019Q4.

Given the limitations of automated systems and risks associated with their use, several other countries and organizations have taken a strong stance against general content monitoring obligations. For example, the EU's e-Commerce Directive[14] and proposed Digital Services Act[15] contain express prohibitions on mandating "general monitoring." The EU Commission stated that requiring monitoring "could disproportionately limit users' freedom of expression and freedom to receive information, and could burden service providers excessively and thus unduly interfere with their freedom to conduct a business. The prohibition also limits incentives for online surveillance and has positive implications for the protection of personal data and privacy."[16] A 2018 UN report on freedom of expression also stated that "[s]tates and intergovernmental organisations should refrain from establishing laws or arrangements that would require the 'proactive' monitoring or filtering of content, which is both inconsistent with the right to privacy and likely to amount to pre-publication censorship."[17] Similarly, several organisations dedicated to promoting and protecting fundamental rights and freedoms in the digital environment have stated that "general monitoring would undermine free expression and privacy by imposing ongoing and indiscriminate control of online content with mandatory use of technical filtering tools."[18]

We urge the government to clarify that the "reasonable measures" that are required by Module 1(B) of the proposal do not include mandatory proactive monitoring and filtering of content. Such a clarification would help avoid the problems discussed above and better align the framework with international norms.

VII. Notification to Law Enforcement - Requirements to disclose user data to law enforcement agencies must be accompanied by due process safeguards to prevent the risk of unwarranted government surveillance and of encroaching on users' privacy rights.

[14] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, Article 15.
[15] 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, Article 7.
[16] Proposal for a Regulation of the European Parliament and the Council on a on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020PC0825&rid=2.
[17] "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression," United Nations (2018), available at https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/35.
[18] Letter to Members of the Telecommunications Council, Executive Vice-President Vestager, and Commissioner Breton (June 4, 2020), available at https://cdt.org/wp-content/uploads/2020/06/Telecommunications-Council-Joint-Letter.pdf.

We are supportive of OCSPs making voluntary reports to law enforcement regarding illegal content and assisting law enforcement with judicially authorized production requests;[19] however, we are concerned that some of the proposed framework's reporting obligations may undermine due process and privacy protections and conflict with OCSPs' other legal obligations.

The general reporting provisions in the proposed framework would require OCSPs to:

- notify the RCMP in circumstances where the OCSP has reasonable grounds to suspect that content falling within the five (5) categories of regulated harmful content reflects an imminent risk of serious harm to any person or to property, as may be prescribed through regulations established by the Governor in Council; or
- [as an alternative] report prescribed information in respect of prescribed criminal offences falling within the five (5) categories of regulated harmful content to prescribed law enforcement officers or agencies, as may be prescribed through regulations established by the Governor in Council. Under this provision, OCSPs would also be required to report information respecting terrorist content and content that incites violence that will be made inaccessible in accordance with this legislation to the Canadian Security Intelligence Service (CSIS) in a manner that conforms to Governor in Council regulations relating to the threshold, timing, format and any other requirements for such reports.[20]

Both of these proposed approaches notably do not call for this data sharing process to be overseen by an independent judicial authority, as is required to comply with Section 8 of the *Charter*: Canadian courts have held that law enforcement must have a court-approved production order to obtain such user data from OCSPs. Additionally, the proposal gives the Governor in Council discretion to specify the information that must be included in the notifications or reports.[21]

Absent further clarification in the legislation, we are concerned that these provisions may require OCSPs to regularly provide extensive amounts of user data to law enforcement authorities. Given the breadth of the definitions for the five categories of prohibited content and risk of heavy penalties for noncompliance, many OCSPs may feel pressured to report any content that could *potentially* fall into the prohibited categories. In addition to flooding law enforcement entities with many unhelpful reports about non-prohibited content, this regular

[19] Google receives law enforcement requests for data from all over the world, and we have a dedicated team that responds to them around the clock, every day of the year. We also work to streamline the process for governments to obtain digital evidence. For example, our Law Enforcement Request System (LERS) allows a verified law enforcement agent to securely submit a legal request for user data, view the status of the submitted request, and download the response submitted by Google.
[20] Technical paper, Module 1(B), 20.
[21] Technical paper, Module 1(B), 20.

flow of large volumes of user data from private companies to law enforcement organizations without user knowledge would violate consumer expectations about privacy and government surveillance in a democratic country. The establishment of such a reporting system may also restrict political speech and free expression, as users may be hesitant to publish legitimate content that relates to prohibited content (e.g., a documentary about terrorism) if they know that it may lead to their information being reported to law enforcement.

While it is possible that some of the privacy impacts of the reporting obligations could be mitigated by limiting the contents of the law enforcement report to information about the content itself (which will in some cases be publicly available), serious privacy risks would remain. For example, many OCSPs provide users with the ability to limit the audience of content they post. In cases where a user has shared content with a handful of people, the content is arguably more akin to a private communication than publicly available information. Private communications are expressly exempted from the framework, and we encourage the government to ensure that similar communications are given similar treatment under the framework.

The Supreme Court of Canada has observed that anonymity is one of the key elements of constitutionally-protected privacy, and this is "particularly important in the context of Internet usage.[22] The provision of identifying information to law enforcement could also potentially affect the user's *Charter* section 8 rights related to unreasonable search and seizure. Currently, Canadian law enforcement agencies are only able to obtain information about an Internet user who has posted content online if they prove to a judge, under oath, that "there are reasonable grounds to believe that an offence has been or will be committed under this or any other Act of Parliament". The judge is then tasked with determining whether the public interest in the police acquiring this information outweighs the privacy and other public interests at stake. The proposed framework for notification to law enforcement removes this judicial check, which has been developed in order to balance the critical constitutionally protected interests at stake. In essence, it replaces a cornerstone of our legal system, the impartial judge, with a private sector entity that has been structurally incentivised to over-report.

The reporting obligations may also conflict directly with legal obligations applicable to OCSPs in other jurisdictions. For example, the personal data of users/customers in the European Economic Area (EEA) and Switzerland is subject to the protections of the EU's General Data Protection Regulation (GDPR). If an EU data subject posts content that triggers the law enforcement reporting requirements, an OCSP that is subject to the GDPR may be unable to share the personal data of that user with Canadian law enforcement organizations due to the limitation of Canada's adequacy decision to commercial organizations. The disclosure would either be simply unlawful or may risk violating the applicable EU laws. Such a dilemma would

[22] *R. v. Spencer*, 2014 SCC 43, at para 45 <https://canlii.ca/t/g7dzn#par45>.

force the OCSP to choose between the risk of significant penalties for noncompliance with the proposed framework and the risk of significant fines for violations of the GDPR if an OCSP were to disclose the personal data of European users. It may also risk substantial damages payable to the affected individual, as a mandatory disclosure in Canada would not be a defence to a claim in the EEA. It is well established within customary international law and Canadian domestic law that a legal requirement in Canada that would cause an offence under another country's law offends sovereignty, comity and international norms. The Canadian framework for online harms needs to take this into account, particularly where the other jurisdiction is closely allied with Canada.

Given the risks associated with the current reporting requirements, we urge the government to include appropriate statutory protections for privacy and due process. Potential revisions could include narrowing the scope of the reporting requirements and/or prescribing the specific information that must be included in a report instead of leaving that issue to the discretion of the Governor in Council.

VIII. Reports to the Digital Safety Commissioner - The obligation to include demographic data in regular reports to the DSC is impractical and may undermine user privacy.

While we agree that it is important for the government to examine the impact of online harms on different demographic populations, we believe that the mandated inclusion of demographic data in OCSP reports to the DSC is unlikely to yield accurate or helpful information and may undermine user privacy by forcing OCSPs to collect sensitive demographic data when it is otherwise not necessary. Currently, the proposed framework provides that OCSPs must generate and provide reports on a scheduled basis to the DSC on Canada-specific data that includes, among other things, "information on their (a) notifications to the Royal Canadian Mounted Policy (RCMP) or (b) reporting to law enforcement" and, for such notifications, "anonymized and disaggregated information about the kinds of demographics implicated."[23]

Many platforms would simply be unable to comply with this requirement under their current data collection practices. OCSPs often do not collect detailed demographic data about their users because: (a) it is frequently not necessary in order to provide services to users; and (b) such data can be sensitive personal information and is subject to additional legal protections in many jurisdictions around the world, including Canada. For example, Canadian privacy laws follow the "data minimisation principle" and require organizations to only collect personal information where it is reasonably necessary to perform the services being delivered. Therefore, if the demographic reporting requirement is included in the framework, OCSPs would effectively be forced to start collecting this sensitive data about Canadian users. Forced

[23] Technical paper, Module 1(B), 14.

collection of this data runs contrary to user privacy interests and conflicts with general norms regarding data minimization. It would also create an ongoing privacy risk for Canadians by forcing OCSPs to indefinitely retain detailed demographic data about all of their Canadian users, some of whom could be harmed if their sensitive demographic data were to become public as a result of a data breach.

It is also important to note that this blanket collection of demographic data may yield inaccurate information. For most OCSPs, the only practical way to collect demographic data will be through user self-reporting. Where users are forced to provide this data, they may choose to report inaccurate data in order to protect their privacy or signal their resistance to this unwanted mandate. Additionally, if it becomes widely known that the government relies on this data in order to understand the impact of online harms on different demographic populations, bad actors may intentionally report false demographic data in an attempt to undermine this goal (e.g., a malicious user may self-report membership in a marginalized group before posting hate speech about that group).

Given the significant risks associated with the mandatory collection of demographic data, we urge the government to remove the demographic data reporting obligation from the proposed framework.

IX. A New Regulatory Scheme - Regulatory oversight and enforcement should focus on systemic failures rather than individual cases of non-compliance so as to avoid stifling free expression and innovation.

A. Focus on systemic noncompliance

We recognize the need for appropriate sanctions for noncompliance with the law. However, we are concerned that the expansive powers granted to the Digital Safety Commission and the open-ended nature of the proposed framework's penalty provision will result in OCSPs being subject to significant financial penalties for mistakes they make with respect to individual pieces of content, even when acting in good faith and under robust compliance procedures. This will have negative consequences for freedom of expression and innovation in the OCS industry.

Under the current proposal, the Digital Safety Commissioner is given the power to "require an OCSP to do any act or thing, or refrain from doing anything necessary to ensure compliance with any obligations imposed on the OCSP by or under the Act."[24] Additionally, it provides that administrative monetary penalties may be imposed on an OCSP for failure to comply with such

[24] Technical paper, Module 1(D), 80.

an order, or for "[a]ny other violations of the Act or regulations."[25] Given the vast amount of content that OCSPs process, the nuanced consideration that is often required to identify prohibited content, and the short amount of time that OCSPs have to evaluate flagged content, it is a near certainty that OCSPs will not be able to achieve perfect compliance with the law with respect to each piece of content.

Therefore, it is possible that an OCSP that has acted in good faith and implemented robust procedures to comply with the law could nonetheless be subject to a significant penalty if, for example, it fails to report certain prohibited content to law enforcement because of an oversight by its automated systems. OCSPs that act in good faith could also be held liable for failure to adhere to inaccessibility orders.[26] For example, an OCSP that takes all reasonable steps to comply with an order to block content could fail to comply with its obligations if a user uploads a slightly altered version of the prohibited content that evades the OCSP's automated systems.

These risks will effectively force OCSPs to err on the side of blocking more content than reasonably required (e.g., adjusting their automated systems so they are overly sensitive in detecting content that *may* be prohibited) and thereby undermine users' ability to share legitimate content and express themselves. It may also stifle innovation and deter new entrants in the OCS space, as the cost of providing such services will incorporate a high risk of significant regulatory penalties.

Although the framework does include a due diligence defence[27] that could potentially prevent a good faith OCSP from being subject to penalties, the scope and requirements to qualify for that defence are currently unclear. The common law due diligence defence has generally been developed and refined in connection with strict liability regulatory offences (such as pollution and motor vehicle offences) that are very different from the present context, which inherently requires the exercise of judgement and investigations into the context in which content was created or posted. The framework should articulate a defence of due diligence that takes account of the complexity of interpreting expression and anticipating harm within an enormous quantity of material. Additionally, the OCSP would still bear the cost of defending itself in a proceeding and raising that defence.

In order to avoid these negative outcomes, we urge the government to clarify and expand the due diligence defence and consider an alternative penalty framework that focuses on systemic compliance with the law. A framework centred around systemic compliance would allow the government to go after wilful noncompliance and the worst offenders while avoiding

[25] Technical paper, Module 1(D), 94.
[26] Id.
[27] Technical paper, Module 1(D), 110.

the creation of perverse incentives that force good-faith OCSPs to adopt overly restrictive content moderation systems that harm consumers and society. For example, transparency requirements in the law can provide regulators with a window into the complaints that the OCSP receives and the processes it has in place to address prohibited content. Where systematic failures are suspected, the regulator can conduct a more thorough investigation and impose penalties as appropriate (e.g., "naming and shaming" the OCSP for its failure to meet its obligations, imposing monetary penalties).

The majority of OCSPs will endeavour to comply with all their legal obligations related to content in scope and will act in good faith in doing so. However, both machines and humans are fallible, particularly when it comes to the inherently subjective exercise of parsing content that requires context in order to determine whether it fits into the five categories of online harms. There will also be a "learning curve" as new requirements are implemented and disseminated through an organization. Any resulting framework should recognize this and require that the regulators first take a remedial approach when dealing with individual complaints and systemic issues that are appropriately addressed through collaboration and cooperation with the regulators.

Under this approach, it will be particularly important for the regulations to clearly describe what constitutes a "systemic failure" in order to provide OCSPs with clarity about their obligations. This definition should consider factors such as the amount of content processed by the OCSP, the amount of prohibited content identified on the OCS, and the success rate in promptly addressing prohibited content.

B. Blocking of content by telecommunications service providers

Another aspect of regulatory power provided under the framework about which we have concern is the Digital Safety Commissioner's power to apply to the Federal Court to seek an order to require Telecommunications Service Providers to implement a blocking or filtering mechanism to prevent access to all or part of a non-compliant OCSP's service in Canada, where that OCSP has repeatedly refused to remove child sexual exploitation and/or terrorist content.

While we agree in principle with the application of this proposal to child sexual exploitation content, its application to terrorist content, which is much more context-dependent, requires carefully crafting the definition of "terrorist content" to ensure that the government cannot use such language to stifle expression that the government does not agree with. It is a slippery slope from this type of blocking for context-dependent content to state-sanctioned Internet censorship, which could have serious consequences for Canadian citizens' freedom of expression and access to information.

X. Incident Response Protocol

We recognize the importance of implementing the Christchurch Call to Eliminate Terrorist and Violent Extremist Content Online. In May 2019, Google and YouTube signed the Christchurch Call to Action. As part of our steps to implement the Call, the Global Internet Forum to Counter Terrorism (GIFCT), of which YouTube is a founding member, developed the Content Incident Protocol (CIP) for industry to respond efficiently to perpetrator-created content after a violent attack. The CIP is a process by which GIFCT member companies quickly become aware of, assess and address potential content circulating online resulting from an offline terrorist or violent extremist event. The CIP sits alongside, and is complementary to, national and multinational crisis response protocols.

Since the attack in Christchurch, GIFCT member companies have developed, refined and tested the CIP through workshops with Europol and the New Zealand Government. To date, we have activated the protocol twice; after the attack on a synagogue in Halle, Germany in October 2019 and following a shooting in Glendale, Arizona in May 2020. In addition, GIFCT members have mechanisms to exchange situational awareness which, since April 2019, we've used over 150 times following terrorist or violent extremist attacks around the world.

The proposed framework allows the DSC to establish a national incident response protocol.[28] We urge the government to ensure that any such national protocol be consistent with the CIP. As noted above, the CIP represents a globally-coordinated approach to implement the Christchurch Call to Action. The introduction of a national approach inconsistent with the CIP risks undermining the effectiveness of the latter, particularly in time-critical situations, as OCSPs would be forced to grapple with multiple competing frameworks.

XI. Penalties - To avoid the unnecessary blocking or removal of lawful, legitimate content, financial and criminal penalties must be applied proportionately.

As discussed above, we are concerned that that proposed framework will create a system that unduly punishes OCSPs that operate in good faith and, as a result, pressures OCSPs into adopting imprecise and overly restrictive content moderation compliance strategies that will deny Canadians a full opportunity to share and view legitimate content. These risks are greatly exacerbated by the size of the penalties that are permissible under the proposed framework.

As drafted, the framework allows for penalties of up to the higher of $25,000,000 or 5% of the OCSP's gross global revenue.[29] Such figures create enormous legal risk for OCSPs, particularly

[28] Technical paper, Module 1(B), 18.
[29] Technical paper, Module 1(D), 119.

if violations can be imposed for noncompliance with respect to individual pieces of content. The threat of these fines may deter established companies from providing OCS services in Canada and discourage startups in the OCS space from launching in Canada.

The range of penalties set out in the framework is disproportionate to the underlying actions sought to be deterred. Associating the penalties with an OCSP's gross global revenue results in penalties that are disconnected from the OCSP's activities in Canada and further disconnected to the reality of their potential presence in the Canadian marketplace. While the factors to be considered in the imposition of any particular penalty will hopefully be connected to the blameworthiness of the conduct, its recklessness and the harm that may have arisen with respect to Canadian residents, pegging penalties to global turnover unnecessarily but inevitably focuses on a company's operations that are wholly disconnected from Canada, and thus from any regulatory impact in Canada.

In addition, the possible imposition of penalties related to the blocking of content that has not been determined by a court of competent jurisdiction to actually be unlawful penalizes OCSPs whose only malfeasance is failing to block access to content that a complainant and a regulator consider to be likely unlawful in Canada.

In order to avoid these risks to free expression and innovation, we urge the government to provide strong safeguards in the legislation that will assure that monetary penalties are imposed in a reasonable and proportionate manner. Although the current text of the framework lists factors that must be considered when determining the amount of a monetary penalty,[30] a clear requirement for proportionality and greater guidance on the application of these factors are needed. Such steps would help the framework better align with international norms regarding content moderation laws.[31]

Thank you for the opportunity to comment on Canada's proposed approach to address harmful content online. We are committed to continuing our efforts to ensure our platforms provide a safe community where our users can thrive and we welcome the opportunity to discuss our recommendations in more detail.

[30] Technical paper, Module 1(D), 107.
[31] See, e.g., "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression," United Nations (2018) (discouraging states from "imposing disproportionate sanctions, whether heavy fines or imprisonment, on Internet intermediaries, given their significant chilling effect on freedom of expression.").