24 September 2021

Michele Austin
Head | Public Policy

Twitter Canada
% 901 King Street West
Toronto, Ontario, M5V 3H5

(613) 290-7870
maustin@twitter.com
@_MicheleAustin

CONFIDENTIAL

Department of Canadian Heritage
25 Eddy St
Gatineau QC K1A 0S5
By email: pch.icn-dci.pch@canada.ca

RE: The Canadian government's proposed approach to address harmful content online

To Whom It May Concern:

On behalf of Twitter, thank you for the opportunity to respond to the Government of Canada's proposed approach to regulating online content.

Online safety is a shared responsibility. Digital service providers as well as governments, private citizens and network service providers play an important role in protecting their communities from harmful content online.

We create rules to keep people safe on Twitter and promote healthy conversations. Our rules are continuously evolving to reflect the realities of the conditions in which we operate.

Under our rules, Twitter currently takes action on all categories of content listed in this consultation (terrorist content; content that incites violence; hate; non consensual sharing of intimate images; and child sexual exploitation content). The five categories are also currently actionable under existing Canadian criminal and civil law. Each of the categories of content listed in this consultation is the subject of an offence under the *Criminal Code of Canada*. The *Criminal Code* prohibits publishing and distributing non-consensual sexual images[1] and child sexual exploitation[2], promoting hate propaganda[3], instructing or counselling a person to commit a terrorism offence[4], and communicating statements that incite violence[5]. The *Mandatory Reporting Act* requires reporting of online child sexual exploitation[6]. The Canadian common and civil law regimes also provide recourse and remedies to those who have suffered harm from these kinds of activities.

Any changes proposed by this consultation should be mirrored by amendments to the *Criminal Code of Canada* and the *Mandatory Reporting Act*.

Twitter would like to emphasize that online content regulation requires a proportionate approach to balance protections from harm, on one hand, against the fundamental right to freedom of expression under the *Canadian Charter of Rights and Freedoms* and against the right to procedural fairness and privacy, on the other. This is a fine balance, and requires a tailored and constantly evolving approach.

When the right balance is struck, companies and regulators alike have clearly delineated responsibilities regarding protections for users' rights, and a shared commitment to foster a diverse public conversation consistent with community expectations within a free and democratic society like Canada. We welcome the opportunity to comment on how to achieve that balance.

As we continue to develop and review Twitter's rules in response to changing behaviors and challenges with serving the public conversation, we understand the importance of considering a global perspective and thinking about how policies may impact different communities and cultures equally. Since 2019, we've prioritized feedback from the public, external experts, and our own teams to inform the continued development of our policies.

Further to our comments on the proposal, Twitter is calling for:

- Consideration of a much wider range of interventions to deliver online safety than proposed, such as renewed emphasis on media literacy and education; greater user control over and choice between algorithms; and the importance of open standards.
- Recognizing personal choice and affording the ability to do nothing. As the work of the Canadian Media Ecosystem Observatory[7] has illustrated with regard to political content, actioning some content can cause it to spread not just on its own terms, but through other channels such as traditional media in their coverage of the actioned content. Once this content is amplified "out in the wild" it can take on a life of its own, where individuals may come to believe it based on their personal beliefs rather than whether or not it is true. Sometimes the best course of action is to do nothing.
- A sustained role for the public to engage in the development of this proposal, including through social media itself. The timing and approach to the public feedback process has discouraged input and analysis from a broad range of stakeholders with diverse and valuable perspectives. The government has not released any data to accompany these proposals. An approach such as that of the United Kingdom which published a White Paper two years before a draft bill with an extended time period for comment is encouraged.

- In order to be effective, this proposal needs to address the offline, real world components of radicalization, extremism and political campaigns.
- Consideration of cost. It is our sincere hope you release the revenue and costing estimates of this proposal publicly so they can be reviewed by experts in the field.

This submission will address key issues outlined in the discussion guide and technical paper released by Canadian Heritage.

Please don't hesitate to contact me if you have any questions about these recommendations.

Sincerely,

Michele Austin
Manager | Public Policy (US & Canada)
Twitter Inc.

* details follow on next page

[1] Criminal Code (R.S.C., 1985, c. C-46), section 162.1
[2] Code section 163.1
[3] Code sections 318 and 319
[4] Code sections 83.21 and 83.22
[5] Code section 319(1)
[6] An Act respecting the mandatory reporting of Internet child pornography by persons who provide an Internet service (S.C. 2011, c. 4)
[7] https://mediaecosystemobservatory.com/press-release-canadian-election-misinformation-project-launch

ISSUE: PROACTIVE MONITORING AND "FLAGGING HARMFUL" CONTENT

Twitter's view is the framework proposed (beginning in module 1B of the technical paper) for proactive monitoring of content sacrifices freedom of expression to the creation of a government run system of surveillance of anyone who uses Twitter.

Even the most basic procedural fairness requirements you might expect from a government-run system such as notice or warning are absent from this proposal. The requirement to "share" information at the request of the Crown is also deeply troubling.

Twitter is committed to respecting the human rights of our users, in line with the expectations articulated in the UN Guiding Principles on Business and Human Rights. We have looked to internationally recognized human rights standards to guide our approach to content policy and enforcement, including those related to the protection of freedom of expression, privacy, security, non-discrimination, and to ensuring due process.

These rights are also enshrined in the *Canadian Charter of Rights and Freedoms.*

We value these approaches and standards in guiding how we navigate instances where rights may be in tension with one another. Each of our Twitter rules is designed to address specific harms on the platform. We try to ensure that content moderation actions we take are both necessary and proportionate to addressing such harms. We welcome further public discussion on how to ensure that regulatory frameworks are designed to prevent harm and reinforce broad equality rights as well as other fundamental human rights.

We support the spirit of the Santa Clara Principles on Transparency and Accountability in Content Moderation in considering how best to obtain meaningful transparency and accountability around government demands for increasingly aggressive moderation of user-generated content on Twitter.

At Twitter, we have identified our own responsibilities and limits. By using Twitter's services, you agree to be bound by our Terms of Service. Further, a user may not use our service for any unlawful purpose or in furtherance of illegal activities

In our continuing effort to make our services available to people everywhere, if we receive a valid and properly scoped request from an authorized entity, it may be necessary to withhold access to certain content in a particular country from time to time. Such withholdings are

limited to the specific jurisdiction that has issued the valid legal demand or where the content has been found to violate local law(s).

At Twitter, transparency is embodied in our open APIs, our information operations archive, and our disclosures in the Twitter Transparency Center. Tens of thousands of researchers access Twitter data we have made available over the past decade via our APIs. Most recently, we have offered a dedicated Covid-19 endpoint to empower public health research, and a new academic platform to encourage cutting edge research using Twitter data. Our archive of state-linked information operations is a unique resource and offers experts, researchers and the public insight into these activities.

In the long term, we believe a greater openness across the industry would be invaluable in delivering the transparency and accountability we all want to see.

Transparency is also vital to protecting freedom of expression. We have a notice policy for withheld content. Upon receipt of requests to withhold content, we promptly notify affected users unless we are prohibited from doing so (e.g., if we receive a court order under seal). When content has been withheld, we also clearly indicate within the product and publish requests to withhold content on Lumen—unless, similar to our practice of notifying users, we are prohibited from doing so.

“Flagging” will be used as a political tactic. As lived during the recent Canadian federal election, a general approach to flagging will result in censorship. Throughout the election campaign, political parties and their officials tried to have content “flagged” as “harmful” in an effort to have it removed from public discourse or score political points. Three of the many examples can be found here, here and here.

Further, individuals who report content should always be offered the option to remain safely anonymous. In some cases, there is a danger the reporter or the victim would be caught up and exposed via any national security investigation or in the sharing of information between governments or law enforcement agencies.

Our position on freedom of expression carries with it a mandate to protect our users’ right to speak freely. While we may need to release information as required by law, we try to notify Twitter users before handing over their information whenever we can so they have a fair chance to fight the request if they so choose.

ISSUE: 24 HOUR TAKEDOWN REQUIREMENTS

Twitter opposes the recommendation of a time limit on “addressing” any content “flagged” by any person in Canada as “harmful” content.

- The proposed time limit does not allow for judicious, thoughtful analysis in a manner that balances the right to freedom of expression in Canada with the right to freedom from discrimination and prejudice.
- According to existing research and analysis, the proposed system has a high probability of negatively impacting marginalized, racialized and intersectional groups. More information from Prof. Suzie Dunn at Dalhousie University can be found here.
- The 24 hour proposal should be abandoned. Content should be addressed as quickly and as possible and within the scope of existing Canadian jurisprudence, terms of service and rules by the online communication service providers.
- Further, any standard applied in the digital world should also be applied in real life. For example, law enforcement should be required to both launch an investigation within 24 hours of “flagging” as well as remove any hateful content - graffiti on a statue for example - that appears within 24 hours across the country.

ISSUE: WEBSITE BLOCKING

The proposal by the government of Canada to allow the Digital Safety Commissioner to block websites is drastic. People around the world have been blocked from accessing Twitter and other services in a similar manner as the one proposed by Canada by multiple authoritarian governments (China, North Korea, and Iran for example) under the false guise of ‘online safety,’ impeding peoples’ rights to access information online.

Further, there are no checks or balances on the commissioner’s authority, such as the requirement of judicial authorization or warnings to service providers. The government should be extremely mindful of setting such a precedent - if Canada wants to be seen as a champion of human rights, a leader in innovation and in net neutrality globally, it must also set the highest standards of clarity, transparency and due process in its own legislation.

Clear guardrails must be put in place, and full assessments of potential unintended consequences should be undertaken before regulatory action is pursued. When this analysis takes place it must be released publicly.

ISSUE: WORKING WITH AND REPORTING TO LAW ENFORCEMENT AND OTHER AGENCIES

Twitter has an excellent working relationship with both the Royal Canadian Mounted Policy (RCMP) and the Canadian Security Intelligence Service (CSIS), which we value greatly.

We also work in partnership with the Canadian government though the Global Internet Forum to Counter Terrorism (GIFCT), the Christchurch Call to Action (CCTA), and the National Center for Missing and Exploited Children (NCMEC).

For example, Twitter already complies with domestic investigations of terrorist content and content that incites violence. Via Canada and Twitter's membership in GIFCT, we jointly announced in July that GIFCT is expanding its taxonomy database to capture terrorist manifestos. In the CCTA's *Crisis Response Work Plan*, Canada and Twitter both agreed to provide investigatory and prosecutory cooperation and trusted information exchanges, given that both are conducted in a manner that is consistent with the rule of law, has strong protections for human rights, and has relevant data protections and privacy regulations in tact.

The Government of Canada should not be using this proposal to grant CSIS or the Crown additional powers outside of those that are clearly identified in the *CSIS Ac*t. In addition, digital service providers are not an extension of Canadian law enforcement organizations.

If Twitter is required to preserve child sexual exploitation data beyond the NCMEC standard, Twitter will need clarity around what is required from the Government of Canada over and above what we currently provide. The feedback we have received from the RCMP is that our reporting is excellent. Industry practices vary widely and some peer companies do not submit the same set of data to NCMEC/the RCMP as Twitter.

Twitter will also need to consult and build out a new retention policy. We do not recommend holding this data indefinitely. Requirements for companies to hold on to personal data longer than necessary goes against best privacy practices and creates more risk of harm in the event of a breach.