<u>Submission on the Canadian Government's proposed approach to address harmful content online</u>

The Chinese Canadian National Council for Social Justice (CCNC-SJ) is grateful for the opportunity to provide submissions on the government's technical paper on addressing harmful content online (the "Technical Paper").

CCNC-SJ has a 40-year history of advocating for the rights and equal treatment of all in Canada, with a focus on advocating for the interests of the Asian Canadian community and has been particularly outspoken on the issue of online hate. We are pleased to see that many of the recommendations, made in 2020 to the Minister of Canadian Heritage Minister Steven Guilbeault, were incorporated into the Technical Paper.

As Mr. Trudeau's government continues its mandate, we urge them to remember that the threat of online hate is just as urgent today as it was a week ago. CCNC-SJ supports the Technical Paper in principle and commends its drafters on undertaking an ambitious effort to address online hate and social media regulations.

However, the real work must begin now. It is incumbent on the government to make good on its promises and implement an effective and comprehensive scheme to address online hate and social media.

Timing and Implementation

As the COVID-19 Pandemic has taught us, social media can quickly and effectively be used by proponents of hate to disseminate harmful content and misinformation on a massive scale. To date, in collaboration with Professor Ishtiaque Ahmed of UofT, CCNC-SJ has collected over 3,000 anti-Chinese tweets that contain stigmatizing themes or messages, many of which are still available on Twitter.

No timeline or budget has been established

While CCNC-SJ supports the creation of the three regulatory entities under referred to in the Technical Paper (the Digital Safety Commissioner of Canada, the Digital Council of Canada, and the Advisory Board), we are concerned that no timelines or budgets have been established to enforce these undoubtedly complicated and resource-intensive bodies. There have been implications that it could take months or even years for these bodies to be established. In the meantime, our communities cannot afford to wait, as misinformation and discrimination run amuck on mainstream social media, causing irreparable harm to the reputation and mental health of racialized communities.

CCNC-SJ recommends that:

- A formal bill immediately is tabled, so that the process for implementing new social media and online hate laws can commence with the inclusion of recommendations made here, and by other community experts.
- The bill must provide flexibility in implementation to accommodate changes in technology and understanding of the spread and appeal of online hate and misinformation.
- A fixed budget and timeline must be provided within the next 3 months to ensure efficacy and proper operation of the regulatory bodies that will be created.

The advisory board's establishment and role is not transparent

CCNC-SJ acknowledges the acceptance of recommendations to establish an advisory board as part of the regulatory scheme, as well as the criteria that will be used to select individual board members. However, the lack of discretion and transparency for the Minister to select the 7 members of the board is concerning, given that members are appointed by the Minister "at pleasure".

It is unclear who the advisory board is accountable to in their role of overseeing other regulatory bodies, whether they must make reports, and what kind of enforcement or otherwise powers they may have. First and foremost, it is crucial that the individuals who make up the board have the interests of diverse groups in mind and therefore the appointment of the board must be transparent.

CCNC-SJ recommends that:

- The election of individuals to the advisory board be a public process with clear criteria.
- The role, mandate and powers of the advisory board are clearly articulated and publicly disseminated.

Weaknesses in the proposal

The process of removing content is a burden on victims

CCNC-SJ emphasizes that it is not the responsibility of the victim, and those reporting, to remove harmful content. The Technical Paper requires that victims or other interest groups be subjected to potentially lengthy and arduous quasi-judicial processes to have the content removed, adding to their burden, and extending trauma.

As noted above, the COVID-19 Pandemic has demonstrated how quickly harmful content can spread. Moreover, reports and surveys have indicated that victims are not likely to make reports for fear of incurring legal consequences because of distrust in law enforcement and the risk of retaliation and repeating trauma. The balance between the mental health and safety of individuals and the right to freedom of speech is fraught. If content creators and publishers feel so strongly that their content is a lawful exercise of free speech, then the burden should be on them to demonstrate its validity. Instead, it

is taken on by victims who suffer significantly, and are more likely to be deterred from making reports in the first place.

CCNC-SJ recommends that: content identified as harmful or hate speech be removed automatically and that the burden of an appeals process to assess content as legal or not be the responsibility of content creators and/or publishers.

The third-party reporting process must be clearly spelled out

Canadian Heritage has indicated that reports can be filed both by victims, as well as organizations representing them, even if such organizations have not been fully retained or authorized to represent them. While this will benefit victims by allowing their interests to be pursued with minimal risk, the process for such representation must be specified in the legislation. Adequate support, guidance, and funding must also be provided to third-party organizations who will have to undertake these undoubtedly daunting and resource-intensive advocacy actions.

CCNC-SJ recommends that: various procedural requirements for reporting and third-party reporting be clearly spelled out:

- What the process is and how it is different when third-party organizations report and represent victims.
- Whether victims who are the subject of the third-party reporting process can participate or discontinue a report.
- Whether funding will be provided to community organizations that provide support to or represent victims in removal processes and if so, what the quantum and requirements for receiving funding will be.

No uniform reporting system has been suggested

While section 12(a) of the Technical Paper sets out that reporting systems must implement "accessible and easy-to-use flagging mechanisms for harmful content", there are no specifics provided. This will be detrimental to regulators, researchers, and public interest groups who will require uniform data to assess the efficacy of reporting mechanisms, and track trends in online hate and misinformation. It will also be detrimental to users, who may be dissuaded from filing reports due to inconsistent or difficult to access reporting mechanisms.

Indeed, a review of the efficacy of Germany's NETZDG law (one of the world's first and most comprehensive attempts to regulate social media and online hate) revealed that failures to standardize reporting mechanisms and data led to inconsistencies in reports and reported numbers. Comparisons between data sets across platforms are also difficult to analyze. For example, Facebook and Twitter's data focused on the number of complaints (where there could be multiple complaints about a single piece of content), whereas Google provided the number of content items, making it difficult to assess the actual quantity of hateful content across platforms. Facebook also received 100 times less reports

than Youtube and Twitter despite its wide usage, likely because the reporting function was difficult to find, and required multiple clicks to access.[1]

While CCNC-SJ commends the inclusion of section 14 of the Technical Paper that will require scheduled, detailed reports on various considerations related to online hate and how it is monetized, we repeatedly stress the importance of uniformity in reporting requirements to ensure transparency and accurate comparisons.

CCNC-SJ recommends that: a uniform reporting function must be designed and mandated for social media platforms, specifying:

- Location of the reporting function.
- A standardized form and specifications as to what data is collected in the report and provided to regulators.
- Number of clicks necessary to complete the report.

Enforcement measures aren't stringent enough

CCNC-SJ believes that the enforcement measures proposed for repeated non-compliance with orders to remove child exploitation and terrorist content should be extended to hate speech and misinformation. This would mean that social media platforms may be at risk of being blocked in whole or in part by internet service providers should they repeatedly fail to comply with orders to remove hate speech and misinformation.

Hate speech has no place in Canada, and the NETZDG law has demonstrated that social media companies can act much quicker in removing content than they claim. In Germany, Facebook hired several thousand more content moderators and can provide detailed country-wide reports.[2] We also learned from the NETZDG law that social media companies are more than happy to pay fines to underreport hate speech, even when the fines are in the range of 2 million Euros.[3]

It is important to recognize that any financial penalty, even up to significant fines of $20-25 million or 4-5% of global revenues for offences by social media companies, may ultimately result in a pay-to-play system, where social media companies consider fines as simply part of the cost of doing business. While the fines set out in the Technical Paper are admirable, they pale in comparison to social media companies' potential earnings. The anti-vax industry for example, is estimated to generate Facebook $1

---

[1] https://www.ivir.nl/publicaties/download/NetzDG_Tworek_Leerssen_April_2019.pdf

[2]https://static1.squarespace.com/static/5ea874746663b45e14a384a4/t/5fd76fe7e5ac582896424816/1607954415617/MTD_Report_Tenove_Tworek.pdf

[3] https://www.politico.eu/article/germany-fines-facebook-e2-million-for-violating-hate-speech-law/

billion in advertising revenues a year, making even fines in the tens of millions negligible by comparison.[4]

It is also important to note that if the government uses fines, in whole or in part, to deter social media companies, then part of that revenue from the fines should be used to heal and benefit the communities that online hate and misinformation target and harm.

CCNC-SJ recommends that:

- Social media companies face the potential penalty of being blocked in whole or in part in Canada for persistent failure to adhere to orders to remove online hate and misinformation.
- Social media companies be subject to higher fines for punishable offences.
- The government commits revenue from fines to support disadvantaged communities, victims, and organizations combating online hate and misinformation.

The role of the police is still unclear and should be handled with caution

The Technical Paper outlines extensive and far-reaching powers of the police to gather information on proponents of online hate and misinformation.

Police involvement in the prosecution and removal of online hate and misinformation must be done with extreme caution. Disadvantaged communities have a tumultuous relationship with law enforcement, particularly concerning hate crimes and hate speech. Current police training and guidelines are underdeveloped to tackle real time incidents of hate and they will also face incredible challenges in prosecuting and handling online hate speech cases.

There must be measures of accountability with clear guidelines, boundaries, and training for the police to play a role in the fight against online hate.

CCNC-SJ recommends that:

- Community organizations are provided with the opportunity and funding to provide oversight.
- Extreme caution is exercised when integrating the police into the regulation of online hate and social media.
- Standardized training on dealing with hate crimes and hate speech be provided to police across Canada.
- Clear guidelines and boundaries are established for the role of police and the extent of their powers.

---

[4] https://252f2edd-1c8b-49f5-9bb2-cb57bb47e4ba.filesusr.com/ugd/f4d9b9_6910f8ab94a241cfa088953dd5e60968.pdf

Research and adaptation

The proposed social media legislation will undoubtedly encounter a litany of social and technological hurdles and inefficacies along the way. The Technical Paper appears to require the provision of accessible, comprehensive data that can better help shape public policy on online hate and social media. Research is crucial and must be conducted continuously with mechanisms put in place to expeditiously adapt regulations based on said research.

Understanding the algorithms

One thing we know for certain is that social media platforms and their algorithms, like Instagram, recommend posts containing misinformation, resulting in hundreds of thousands of likes.[5]

There are robust requirements for public transparency and reporting in the Technical Paper, but notably missing is the requirement for social media to disclose information on their algorithms and how they work. While there is some loose wording that inspectors can examine computer algorithms under section 89 of the Technical Paper, the extent of this power is unclear and doesn't guarantee that inspectors will have the skills and expertise to understand the algorithms without enforcing their own implicit biases. Data on the algorithms must be publicly accessible and comprehensive for research purposes. Social media algorithms are undoubtedly complicated and will require significant resources to understand and analyze.

How regulatory bodies will be able to analyze and ensure full disclosure from social media companies is not clear, which is fatal to any attempts to regulate these companies who will undoubtedly resist disclosure obligations.

If the government is unable to lay out a comprehensive plan as to how it will compel full disclosure of how social media algorithms work, it is incumbent for the government to invest in third-party resources and research that aim to understand these algorithms, like that conducted by CCNC-SJ.

CCNC-SJ recommends that:

- Social media companies are required to disclose how their algorithms work.
- The government provide funding to organizations researching social media algorithms and the spread of online hate and misinformation.

---

[5] https://252f2edd-1c8b-49f5-9bb2-cb57bb47e4ba.filesusr.com/ugd/f4d9b9_9877528dd81b402b948044ab10a989d9.pdf

Content Removals cannot be the only measure of efficacy

Content removals are a problematic measure of success. Though it may be beneficial to see the removal of hate content increase over time, this does not necessarily indicate success in fighting hate speech. Instead, it may indicate an increase in hate speech or incentivization of its creation.[6]

The measure of success of Canada's legislation against online hate and misinformation is likely to change as we come to better understand social media algorithms, the scale of online hate and misinformation, and how people respond to legislation and regulation.

It is crucial that any social media legislation that is tabled provides for regulations that are flexible and have space and mechanisms for amendments based on further research.

CCNC-SJ recommends that:

- More research is conducted and organizations conducting research are provided with adequate funding to perform research on a continued basis.
- The regulations that are a part of the legislation are adaptable and provides space for amendments in policy to reflect further research and findings, as well as better measures of efficacy.

- 30 -

---

[6]https://static1.squarespace.com/static/5ea874746663b45e14a384a4/t/5fd76fe7e5ac582896424816/1607954415617/MTD_Report_Tenove_Tworek.pdf