

Background

Online hate is pervasive. It is a barrier to the free expression and democratic participation for Canadians who are part of protected groups. It is often criminal, but has not been sufficiently addressed by large online platforms or the criminal justice system.

Online rhetoric contributes to offline and in-person harassment, violence, and terrorist attacks. Since the start of COVID 19 and the corresponding and fallacious anti-Chinese rhetoric, nearly 1 in 3 Chinese Canadians have reported being physically attacked¹. Online incel (anti-women) and anti-Muslim ideology have been responsible for several terrorist attacks in Canada, claiming over 17 lives.^{2 3 4}

Canadian law promises protections to groups identified by colour, race, religion, national or ethnic origin, age, sex, sexual orientation, gender identity or expression, or mental or physical disability.⁵

Efforts by large platforms (eg. Facebook) to curb online hate have been insufficient⁶, with organizations like ours unable to notice a significant difference. In many cases here and internationally, their inaction has caused incalculable harm. They are capable of, and must, harness their algorithms to detect and remove hate speech and the users who spread it.

The government is considering measures to address online hate where platforms have demonstrated themselves to be negligent. First and foremost, we believe any solution must include fines significant enough that there is a financial impetus for platforms to behave responsibly as opposed to serve as a pay-to-play system.

The government must also consider the “anything-goes” platforms which attract neo-Nazis and other bigots and serve as a focal point of criminal hate propaganda and hate organizing. Incel forums, for example, are full of calls to rape and murder women and glorify terrorist attacks. These kinds of international platforms and their operators must not be overlooked.

¹<http://angusreid.org/racism-chinese-canadians-covid19/>

²<https://www.cp24.com/news/i-am-scared-even-for-my-granddaughter-for-my-daughter-van-attack-victim-says-of-incels-1.5189760>

³<https://www.cp24.com/news/message-parlour-stabbing-was-act-of-incel-terrorism-rcmp-1.4945387>

⁴<https://www.cbc.ca/news/canada/ottawa/quebec-city-mosque-shooting-1.5445427>

⁵<https://laws-lois.justice.gc.ca/eng/acts/h-6/page-1.html>

⁶<https://www.isdglobal.org/wp-content/uploads/2020/12/20201218-ISDG-NewsGuard-QAnon-and-Facebook.pdf>

Key Recommendations

1. The government forms an independent body with a mandate that includes enforcing and auditing online hate regulations, updating definitions of prohibited content, and hearing significant appeals.
2. Online hate regulations carry financial penalties significant enough to compel pro-social behaviour from social platforms; the government considers criminal penalties for significant negligence or wilful harm.
3. The government supports the freedom of expression and democratic participation of protected groups by taking strong action against online hate; that freedom of expression is enshrined in the process through strong reporting, transparency, and accountability measures taken by the independent body.

We propose the following principles in regulating large platforms to reduce online hate:

- A good solution today, not a more-perfect solution in two years. This will likely become a politically charged issue. We suggest an incremental and adaptive approach, starting with measures targeting the most egregious forms of online hate.
- A focus on targeted harassment, threats, incitement towards or celebrations of violence, obvious examples of racism and other forms of bigotry towards protected groups enshrined in the 'hallmarks of hate' endorsed by the Supreme Court⁷, and the denial or minimization of Canada's genocides of Indigenous peoples.
- A pro-active solution leveraging algorithms and machine learning to detect and prevent hate messages. There is simply too much hate content for humans to review.
- Adequate safeguards for freedom of speech such as an appeals process for the re-upload of content that is removed by the system. Especially to protect advocacy organizations, who face targeted reporting campaigns carried out by the hate groups they are exposing and criticizing.
- External reporting and audits on platforms' performances. Platforms have already demonstrated they will not responsibly self-enforce or self-regulate.
- More severe penalties, up to and including criminal charges, for operators of those platforms who will not comply. Consider that there are many internationally hosted platforms *designed* to be a platform for what in Canada is criminal hate speech.
- A means of blocking websites sharing the most egregious examples of hate, including the promotion of violence and terrorism, from reaching Canadian users.

⁷<https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/12876/index.do>

- Language changes quickly, and the language hate actors use online changes frequently. Outside experts and algorithmic and artificial intelligence tools must be leveraged to keep up to date on the current language, definitions, and trends in order to avoid missing incidents in the reporting system.
- Prohibiting designs known to cause harm, for example, recommendation algorithms known to lead to even more extreme content.⁸
- Platforms should be required to give warnings and remove repeat offender accounts on the basis of their content that is reported as violating their terms of service or which breaks the law, rather than just removing the content alone but allowing the account to stay up. Individual users have escaped previous bans of groups or pages, and continue to act as superspreaders of hateful ideologies.⁹
- Platforms must take into account the actions of their users offline and on other platforms. It is common for hate propagandists and organizers to moderate their message so that they can organize on some platforms without facing bans.
- The government seriously consider categorizing social media and engagement-oriented social media (eg. forums) as publishers of content rather than distributors, subjecting them to legal liability, should they not comply with regulations in a fulsome manner.
- The independent body undergo a communications campaign to share with the public the intention of these measures, highlighting the harms of online hate, the free expression benefits of curtailing it, and the independence and accountability of the independent body.

With these goals in mind, we are proposing one possible solution:

- An independent oversight body with expert input sets minimum reporting function standards for hate-motivated content that all social platforms (eg. Facebook, Discord, forums) over a certain size must implement, with expert input.
- All reports of hate violations by algorithm and by user are sent to both the platform and to the oversight body, which can audit their compliance.
- The oversight body has the ultimate say in what violates standards, within the confines of our Supreme Court endorsed “hallmarks of hate” guidelines.
- Large platforms with algorithms are required to leverage those algorithms to detect and proactively remove hate content.

⁸<https://www.hopenothate.org.uk/wp-content/uploads/2020/11/A-Better-Web-Final-.pdf>

⁹<https://www.isdglobal.org/wp-content/uploads/2020/12/20201218-ISDG-NewsGuard-QAnon-and-Facebook.pdf>

- Decisions appealed to the platform may be further appealed to the independent oversight body or dispute resolution body, to be dealt with at a staff level.
- Difficult and precedent setting appeals will be heard by a tribunal of the independent oversight body, which should include at least one judge well-versed in the jurisprudence surrounding criminal hate propaganda in Canada, and an advisor with knowledge of how social platforms operate (eg. algorithms, machine learning/AI).
- If the platform does not remove hate content within a short time period, they face significant and escalating fines which, in turn, provide continuing funding for the independent oversight body. These oversight mechanisms must themselves be accountable to public scrutiny.
- The independent body publishes its reports, audits, definitions, and decisions, and incorporates the feedback of the public and expert organizations

While these recommendations deal specifically with platforms, we also strongly recommend that the government reintroduces a better version of s. 13 of the CHRA or its equivalent with the input of key stakeholders, provides significant resources to the CHRC and CHRT earmarked for s. 13 complaints, and conducts a comprehensive legislative review of the Canadian Human Rights Act, which is long overdue.

Signed,

| | | |
|--|--|--|
| Across Boundaries | City of Toronto - Confronting Anti-Black Racism Unit | Progressive Intercultural Community Services Society |
| Bangladeshi-Canadian Community Services | Gibraltar Leadership Academy | Roots Community Services |
| Canadian Anti-Hate Network | Islamic Foundation of Toronto | Salaam Canada |
| Canadian Arab Federation | JSpace Canada | Social Services Network |
| Canadian Centre for Victims of Torture | Labour Community Services of Metropolitan Toronto Inc. | South Asian Legal Clinic of Ontario |
| Canadian Muslim Vote | Midaynta Community Services | Tamil Canadian Centre for Civic Action |
| Canadian Council of Muslim Women | Mosaic Institute | University of Toronto Mississauga Students' Union |
| Council of Agencies Serving South Asians | National Council of Canadian Muslims | Urban Alliance on Race Relations |
| Chinese and Southeast Asian Legal Clinic | Ontario Council of Agencies Serving Immigrants | World Sikh Organization |
| Chinese Canadian National Council for Social Justice | Ontario Agency for Health Protection and Promotion | YWCA Canada |

