

The Role of Private Online Spaces in Platform Governance

Sam Andrey, Toronto Metropolitan University

Defining the scope of platform governance has become a pressing policy challenge for countries around the world, as the spread of illegal and harmful content continues to be exacerbated and amplified on both public and private platforms. In the absence of meaningful regulation in most jurisdictions, the crucial task of balancing the right to free expression and the mitigation of real and growing harms from ill-intended actors has fallen to a small number of companies who have largely [consolidated and privatized](#) online discourse. These companies are almost all based outside of Canada and have ad-based business models that can [discourage taking meaningful action](#) to reduce certain harms.

Canadians report relatively [frequent exposure](#) to hate speech and harassment on public and private online platforms, with rates higher for racialized people, those who have a disability, and those who identify as LGBTQ2S+. Public safety actors and victim groups are [calling](#) for stronger accountability for removal of illegal online content, including incitement of violence and suicide, terrorist content, sexual exploitation and identity fraud. There are also [growing concerns](#) about the online spread of conspiratorial misinformation and its contribution to polarization, radicalization and undermining of Canada's democratic processes, both through organic reach and foreign influence operations. But private spaces present particular problems, which this brief explores.

Beginning in 2021, Canadian Heritage embarked on [various consultations](#) on the potential design of a regulatory framework to address online safety concerns — including roundtables, a citizens' assembly, and an expert advisory group. This process has given the country a chance to learn from other existing policies. For example, [critics](#) of Germany's NetzDG regulation have noted the law's lack of specificity in guidance for platforms, providing too much discretion and leaving room for over-compliance without sufficient oversight. The delay has provided the space for Canada to shift from an exclusive focus on a 24-hour harmful content takedown approach to a more adaptable regulatory model that emphasizes platforms' [duty to act responsibly](#) and transparently review and mitigate systemic risks to users, with similar models being advanced in Australia, the EU, and UK. Sharing learnings and coordination across democratic states striving to responsibly address harmful content and protect freedom of expression will be crucial to maintaining public support and exerting sufficient pressure on large platforms that would be difficult for Canada to achieve alone, particularly with respect to changes that directly impact business models.

A particularly contentious element of platform governance is deciding which platforms and services will be subject to regulatory oversight. The Government of Canada previously [expressed](#) that it intends to exclude from regulation “services that enable persons to engage only in private communications.” Making a distinction between public and private communications in the regulation of speech is of course not new. For example, the [Criminal Code](#) makes it an indictable offence to communicate

statements that willfully promote hatred against an identifiable group “*other than in private conversation.*”

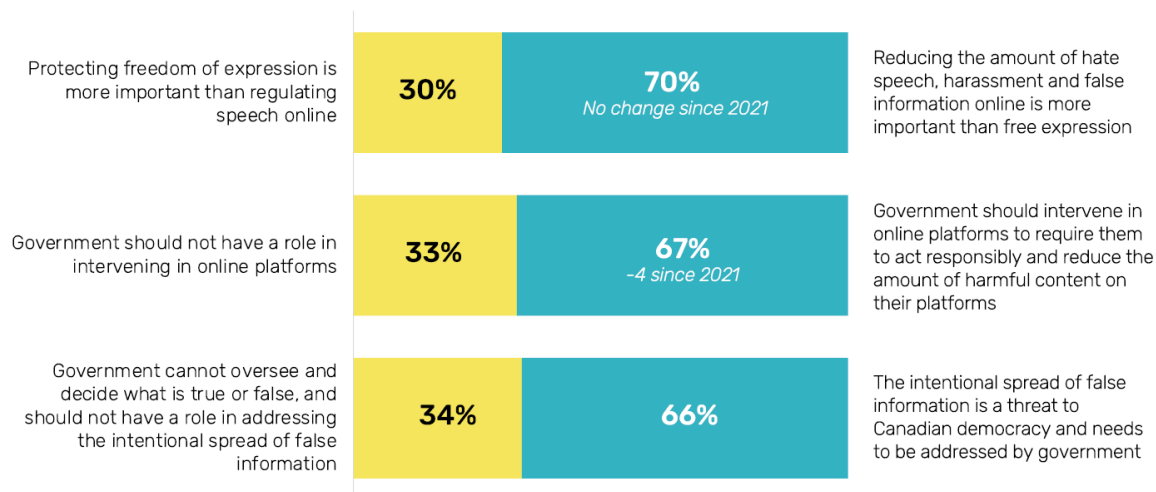
How to make such a distinction in closed online spaces that blur boundaries between private and public is a complex challenge. There are legitimate concerns about the proliferation of illegal content on private online platforms. For example, 26% of Canadians in a [2020 survey](#) reported receiving messages containing hate speech at least monthly on private messaging platforms, with rates higher among people of colour. An [estimated 70%](#) of reports of child sexual abuse on Facebook are through private messages on Messenger or Instagram. After the U.S. Capitol riots and the truckers’ convoy in Canada, concerns were raised about the [role of private groups and messaging in seeding](#) and coordinating the events.

Some private platforms, such as those run by Meta, have taken [some steps](#) to address harmful online content within private messaging, such as enabling users to report harmful content to moderators, introducing labels and limits on message forwarding to create friction for messages to go ‘viral’, and encouraging users to verify highly forwarded message content. Other private platforms, such as Signal or Telegram, have designed their platforms with less oversight and moderation, including much larger maximum group sizes (up to 200,000 users in the case of Telegram, compared to 250 on Messenger/Instagram).

Allied jurisdictions have taken a variety of regulatory approaches to the inclusion of private content (e.g., private profiles, groups, channels and direct messages). [Australia’s Online Safety Act](#) enables the eSafety Commissioner to regulate the removal of “cyberbullying” material on all private platforms. The [EU’s Digital Services Act](#) requires only private platforms with significant user reach, such as Messenger, to enable users to report harmful content and have it reviewed, as well as annual transparency reporting requirements, but does not provide independent oversight over content. In an effort to combat child sexual abuse material (CSAM), the EU has also [proposed obligations](#) for platforms to screen private communications to detect related harmful content. The [UK’s proposed approach](#) likewise proposes to enable content scanning of private content for terrorist and child abuse, though excludes emails and SMS/MMS messages. Both of these proposed approaches to screen private communications have been [criticized with fears](#) surrounding the weakening or breaking of end-to-end encrypted messaging, leading to potentially compromised private communications and, in turn, rights to privacy and free expression.

While policy-making that seeks to reasonably balance competing rights should never merely be subject to majority opinion, the political context for action in this space is a critical dynamic. The very nature of large online platforms means that state regulation could affect most Canadians. So it is worth highlighting [evidence from representative public surveys](#) conducted by our team over the last four years (alongside [others](#)) that suggest a significant majority of Canadians distrust online platforms and are supportive of platform governance efforts and the timely removal of illegal online content in

Canada. As an example, when asked to choose between a set of statements on balancing rights, about two-thirds of Canadians indicated preference for intervention (see figure below). While those on the left and centre of the political spectrum have significantly higher average levels of support for intervention, a majority of those on the right of the political spectrum still support intervention.



Source: [Survey of Online Harms in Canada, 2023](#)

When asked specifically about which types of online spaces they thought should be required to remove illegal content like hate speech or the promotion of violence, a significant majority of 87% supported content moderation on public pages/profiles, while smaller majorities supported it for private groups (61%) and private pages/profiles (59%). Support fell to 40% for private messaging.

Canada should take inspiration from the EU’s Digital Services Act and place minimum standards on messaging platforms with significant user reach in Canada, such as reviewing their systemic risks, having user reporting features, and providing transparency reports. Such an approach would still enable harm reduction, promote greater understanding of online harms on these channels to inform future action, and mitigate the risk of an incentive for companies to create more closed platforms as a means of avoiding new content moderation obligations, without imposing content scanning requirements or weakening encryption. Lessons should also be learned from past efforts to regulate and monitor harms in private communications, such as Canada’s Anti-Spam Legislation, National Do Not Call List, and regulations against knowingly sending false electronic messages through Canada’s Competition Act.

Mounting evidence suggests a majority of Canadians are prepared for regulatory action to address the rise in harmful online content. Doing so in a way that is sensitive to the different forms of private online spaces, and respects the unique role of direct messaging, will go a long way in maintaining public support and confidence.