

Copyright © 2021 Women’s Legal Education and Action Fund (LEAF)

Published by
Women’s Legal Education and Action Fund (LEAF)
180 Dundas Street West, Suite 1420
Toronto, Ontario, Canada M5G 1C7
www.leaf.ca

LEAF is a national, charitable, non-profit organization, founded in 1985. LEAF works to advance the substantive equality rights of women and girls in Canada through litigation, law reform and public education using the *Canadian Charter of Rights and Freedoms*.

This publication was created as part of LEAF’s Technology-Facilitated Violence (TFV) Project. The TFV Project brings together feminist lawyers and academics to conduct research and produce publications imagining legal responses to TFV against women and gender-diverse people that are informed by equality principles. The project also supports and informs LEAF’s law reform efforts and potential upcoming interventions concerning TFV.

Acknowledgements

Deep gratitude and appreciation go to the many people whose efforts and support made this publication possible.

This report was researched and written by **Cynthia Khoo**, a technology and human rights lawyer and researcher. Cynthia holds an LL.M. (Concentration in Law and Technology) from the University of Ottawa, where she worked on cases as junior counsel at the Samuelson-Glushko Canadian Internet Policy and Public Interest Clinic (CIPPIC). Her paper on platform liability for emergent systemic harm to historically marginalized groups received the inaugural Ian R. Kerr Robotnik Memorial Award for the Best Paper by an Emerging Scholar at We Robot 2020. She has managed a sole practice law firm, Tekhnos Law, and obtained her J.D. from the University of Victoria. Her work and expertise span across key areas of technology and human rights law and policy, including privacy and surveillance, equality and freedom from discrimination, online censorship and freedom of expression, intermediary liability, algorithmic decision-making, and technology-facilitated violence.

The report was overseen and coordinated by **Rosel Kim**, Staff Lawyer at LEAF; **Pam Hrick**, Executive Director and General Counsel at LEAF; and **Megan Stephens**, former Executive Director and General Counsel at LEAF.

Sincere thanks are further extended to **Pam Hrick** and **Rosel Kim** for contributing substantive revisions, editing, and copyediting of the report.

This report benefited significantly from the insights and expertise of the following reviewers (in alphabetical order by last name): **Moira Aikenhead**, **Jane Bailey**, **Karen Bellehumeur**, **Nicole Biros-Bolton**, **Suzie Dunn**, **Lex Gill**, **Nicola Henry**, **Pam Hrick**, **Tamir Israel**, **Jo-Ann Kolmes**, **Rosel Kim**, **Emily Laidlaw**, **Brenda McPhail**, **Jill Presser**, **Molly Reynolds**, **Megan Stephens**, and **Jennifer Tomaszewski**.

Thank you as well to **Donald Jackson**, **Julie Mouris**, **Kat Owens**, **Cee Strauss**, **Rosel Kim**, and **Pam Hrick**, for much-appreciated assistance with the citations for this report.

Special thanks to the **LEAF TFV Advisory Committee**: **Moira Aikenhead**, **Jane Bailey**, **Karen Bellehumeur**, **Suzie Dunn**, **Gillian Hnatiw**, **Pam Hrick**, **Nathalie Léger**, **Raine Lilliefeldt**, **Molly Reynolds**, **Hadiya Roderique**, and **Karen Segal**.

Thank you to **Kristyn Watterworth**, who designed the cover page of this report.

This report was funded by the generous support of: Canadian Bar Association’s Law for the Future Fund; the Pilot Fund for Gender Equality (which is supported by a collaboration between Community Foundations of Canada and the Equality Fund, with support from the Government of Canada); and the J. Armand Bombardier Foundation.



Executive Summary

Digital platforms have enhanced and expanded the ways in which we interact and share information with one another. They have also simultaneously provided new mechanisms for those who might seek to engage in abusive conduct to inflict harm on targeted groups and individuals—particularly from historically marginalized and systemically oppressed communities.

This report examines the role of digital platforms in the proliferation of technology-facilitated gender-based violence, abuse, and harassment (abbreviated as ‘TFGBV’). It also examines whether and how digital platforms—such as Facebook, YouTube, and Twitter—should be held accountable for TFGBV through regulation or the imposition of liability under Canadian law.

The consideration of these issues begins with a review of the substance and nature of TFGBV that commonly occurs on digital platforms, as well as examples of platform content moderation models. This is followed by an explanation of the current Canadian landscape concerning platform liability for TFGBV and a review of platform liability regimes that exist in other jurisdictions around the world. The report then grapples with some critical issues and legal complexities associated with holding platforms liable for user conduct. It concludes by making 14 recommendations for federal legal reform and complementary actions to address TFGBV in Canada through the lens of digital platform liability and accountability.

Technology-Facilitated Gender-Based Violence, Abuse, and Harassment (TFGBV)

TFGBV refers to a spectrum of activities and behaviours that involve technology as a central aspect of perpetuating violence, abuse, or harassment against (both cis and trans) women and girls. This term also captures those who hold intersecting marginalized identities, such as 2SLGBTQQIA, Black, Indigenous, and racialized women; women with disabilities; and women who are socioeconomically disadvantaged.

Activities that fall under the umbrella of TFGBV include:

- doxing;
- hate speech;
- threats and intimidation;
- trolling;
- voyeurism;
- impersonation;
- spying and monitoring through account hacking or interception of private communications;
- online mobbing;
- coordinated flagging campaigns;
- sexual exploitation resulting from online luring;
- defamation;
- non-consensual distribution of intimate images (NCDII);
- image-based abuse (including both deepfakes and shallow fakes);
- sextortion; and
- stalking.

These activities may be referred to more generally as aspects or examples of ‘online violence’, ‘online abuse’, or ‘online harassment’. The terms are not necessarily interchangeable, and depend on context.

TFGBV relegates women and girls to secondary status online and in the world. They are rendered unable to freely and fully participate in society and prevented from enjoying true or equal protection of their human rights and fundamental freedoms. The most common response to facing online abuse and harassment is that women reduce their online activities, avoid certain social media platforms or conversations, withdraw from expressing their views, or self-censor if they continue to engage online. This curtails their ability to participate in the contemporary public sphere, including engaging in activism and advocacy, influencing public opinion, or mobilizing social, cultural, or political change. The current state of affairs amounts to a systemic democratic failure and must be addressed as such.

Nearly all TFGBV on digital platforms is committed through online expression, whether through speech, images, videos, or other multimedia. Whether or not a specific instance of TFGBV is illegal depends on whether it meets the definition of a pre-existing criminal offence or cause of civil action. For example, acts of TFGBV that constitute invasion of privacy, impersonation, defamation, criminal harassment, threats of violence, interception of private communications, stalking, recording or surveilling someone without consent (where they have a reasonable expectation of privacy), or NCDII are all already civil and/or criminal offences in Canada.

Instances of TFGBV that fall short of attracting legal liability might be considered ‘just’ speech or expression. However, expression-based TFGBV can be as or more damaging to women and girls and impact their lives in ways that go far beyond the screen. This may include, for example, ‘everyday’ online harassment amounting to social persecution; violent threats that fall short of the legal threshold for criminal liability; trolling; creating and disseminating non-sexualized deepfakes; and online mobbing. One question in Canadian law is whether digital platforms can be regulated to address this level of expression-based TFGBV by their users, whether through regulatory obligations or imposed liability. Answering that question is a complicated exercise that raises questions around intermediary liability, constitutional limitations, proportionality, and the particular dynamics of digital platforms and their role in society. A principled focus on the right to equality—intersectional, substantive equality—can help navigate such questions in setting a path forward to meaningfully address TFGBV.

Role of Digital Platforms in TFGBV

Online platforms such as social media networks, discussion forums, search engines, and video sharing websites have become central venues of our personal and professional lives. It thus comes as no surprise that online platforms are also central sites of TFGBV, which has often been exacerbated by the actions (or inaction) of the platforms themselves. For example:

- Facebook has allowed pages glorifying intimate-partner violence to stand, while removing images of women breastfeeding;
- Twitter has been quick to suspend users who are targets of online abuse, while frequently ignoring the activity of abusive users;
- YouTube’s recommendation algorithms have turned it into an efficient right-wing radicalization machine; and
- Google Search has provided top-ranked search results that reflect racist sexual objectification of Black women and girls.

One particular type of platform that warrants specific attention in the context of TFGBV is the category of platforms that seem deliberately designed to encourage and profit from such abuse. These might be termed ‘**purpose-built platforms**’, as opposed to ‘platforms of general application’ such as Facebook and Twitter, which may include (copious) TFGBV, but do not appear to exist exclusively to cater to TFGBV. Examples of ‘purpose-built platforms’ are ‘The Dirty’ and platforms dedicated to sharing NCDII. What would be a balanced and proportionate liability framework for platforms of general application would likely not suffice to address TFGBV where such expression and conduct constitute the core business model or central service or commodity of a purpose-built platform.

The total constellation of TFGBV as facilitated by digital platforms, particularly, can be termed **platformed misogyny**, or **platformed TFGBV**, based on Ariadna Matamoros-Fernández’s concept of ‘platformed racism’. The term is used to denote how the characteristic features of digital platforms’ design choices, business models and content moderation policies—including their embedded cultural values and politics—combine with the power of platform governance to shape the ‘platformed’ systemic oppression in question, in a way that makes it distinct from non-platformed manifestations.

Digital platforms share several common features that contribute to such companies’ particular role in amplifying, promoting, escalating, and entrenching TFGBV. These features include:

- platform companies’ advertising-driven business models, which maximize user engagement in a way that favours more outrageous and sensationalized content;
- companies’ prioritization of business growth above all else;
- the sheer ease, efficiency, and affordability of automating and multiplying instances of abuse against a particular group or individual;
- the ability for abusive users to remain anonymous and remote, taking advantage of ‘safety in numbers’ in online mobs or coordinated attacks; and
- the ability of users to game content moderation features and other platform affordances to abusive ends.

As a result of these platform dynamics, gender-based violence, abuse, and harassment is no longer constrained by physical boundaries. The ubiquity of the Internet means that TFGBV can become omnipresent and relentless, infiltrating a victim’s most intimate physical spaces, such as their home or bedroom. Users engaging in TFGBV can also leverage their own and targeted individuals’ online social networks to further the abuse, by recruiting others to knowingly or unwittingly share abusive material, and by contaminating the targeted individuals’ own online spaces and communities. The online permanence of abusive material—which is exceedingly difficult to completely eradicate once shared online—also ensures continued revictimization, resulting in lasting psychological and other damage.

Platform Content Moderation Policies and Practices to Address TFGBV

When it comes to addressing TFGBV, platform content moderation measures have been deficient in both design and application. This has exacerbated harms to users most targeted by TFGBV, including historically marginalized groups. For example, community standards on digital platforms have included exceptions to rules prohibiting hateful or harmful speech. This has created major loopholes for demonstrably hateful or harmful content to proliferate. Flagging and reporting mechanisms rely on

users using them accurately and in good faith, but have often been gamed to further the abuse such mechanisms are meant to address.

Human reviewers are generally underpaid third-party contractors working in traumatizing conditions who have only seconds to determine whether a given post should be left up, taken down, or escalated. Automated content moderation is rife with further errors, which have resulted in the removal of content, including posts that constitute parody and satire; innocuous images mistaken for nudity; and content related to 2SLGBTQIA issues or sex education. Algorithmic tweaking and downranking, fact-checking, and labelling have been applied weakly, inconsistently, and highly selectively, for the most part.

In addition, platform companies have consistently demonstrated significant degrees of selective attentiveness, contradiction, and hypocrisy in both the development and application of their content moderation policies and practices. Those targeted by TFGBV or otherwise familiar with the issue have continually reported that major platform companies ignore individual requests for help and largely neglect the broader issue of TFGBV. At the same time, they continue to support and build features that contribute to optimizing their platforms for abuse. Experts have also identified overarching systemic issues with digital platforms' content moderation policies and practices, including:

- selective reliance on the rhetoric and strength of the United States' cultural norms around 'freedom of expression' to justify inaction;
- undue reactivity and sensitivity to public opinion and political influence in content moderation decisions; and
- conflicts of interest that result in platform companies prioritizing business growth and maintaining good relations with the political right over effectively addressing TFGBV.

Canadian Legal Landscape: Platform Liability for TFGBV

There are a number of Canadian laws that could theoretically create platform liability for TFGBV by a platform's user, given the right circumstances. However, many of these have yet to be tested in court. Canada has laws that do the following:

- establish a general intermediary liability regime (e.g., section 22 of the *Act to establish a legal framework for information technology* in Quebec);
- establish platform liability or legal obligations for non-TFGBV user content (e.g., direct liability for 'enabling' copyright infringement and the notice-and-notice regime under the *Copyright Act*, or what is effectively a notice-and-takedown regime that the courts have developed in defamation law);
- address TFGBV in some form but are silent on the role of platforms (e.g., *Criminal Code* offences for NCDII and hate propaganda); and
- address neither TFGBV nor intermediary liability specifically, but are laws of general application that could apply to platform companies as organizations, provided the factual circumstances met the relevant legal test (e.g., statutory human rights law, criminal corporate negligence, or product liability).

There thus appears to be a gap in Canadian law, in that there is no *specific* form of legal liability for platforms with respect to TFGBV. However, some common principles emerge from current intermediary liability law and jurisprudence, which can inform how the law should be extended to address platform liability for TFGBV.

As a starting point, courts have generally been reluctant to hold online intermediaries liable for user expression or conduct, without something more to justify holding one party liable for another party's misconduct. This is particularly true where the intermediary is a 'mere conduit' and simply plays an infrastructural role of connecting third parties to one another. However, Canadian defamation law will hold a platform accountable for a user's speech if the platform had specific knowledge about it but took no action to address it. Canadian copyright law places a legal obligation on platforms to assist potentially injured parties, but will not hold a platform liable for user copyright infringement unless the platform's involvement rises beyond a certain level according to a six-factor test for being an 'enabler'. Overall, the degree of liability rises the more the platform is involved and the more that is at stake for the injured party, up to direct liability where the platform has essentially abandoned its 'intermediary' role in producing content that constitutes a civil or criminal offence.

Even if a platform company is not a party to a legal proceeding and is not liable for the harmful content in question, it may still be required to take certain steps to address the content, including:

- complying with court orders or statutory obligations, such as forwarding a notice to the author;
- removing, deindexing, or disabling access to content; or
- releasing information to help identify an anonymous user engaging in abuse.

These obligations are rooted in ensuring access to justice and practical remedies for victims, in a way that recognizes the realities of the Internet. Providing platforms with explicit protection from liability acknowledges the particular role of platforms in the Internet ecosystem and with respect to specific harms—namely, a dominant and facilitative role which justifies accountability and responsibility for assisting in the remedy, but does not generally warrant imposing liability for the wrongdoing itself.

There are some laws of general application in Canada that could potentially ground platform liability for TFGBV on a systemic or institutional level, based on the platform company's design choices and business decisions. Examples include human rights statutes, in situations where disproportionately exposing historically marginalized groups to TFGBV constitutes discrimination in the provision of goods, services, or facilities. The laws of commercial host liability and product liability may also apply to digital platforms in certain circumstances, based on underlying reasoning of moral hazards and incentives to act against the customer (or user), or the public interest. Additional laws that may apply include criminal corporate negligence or being party to an offence as an organization, though these would only apply to existing criminal offences, such as hate speech, intimidation, threats of violence, or criminal harassment.

Platform Liability Models: Jurisdictional Scan

Several lessons and conclusions may be drawn from the jurisdictional scan of platform liability models for harmful behaviour by users contained in this report.

First, the manner in which section 230 of the *Communications Decency Act* (CDA 230) has been applied in the United States to provide broad immunity to platforms for user content does not strike an

appropriate balance among the considerations relevant to the context of TFGBV, nor does it necessarily reflect what CDA 230 was initially intended to achieve. The breadth of platform protection under CDA 230 derives largely from a long line of overly expansive interpretation by American courts. Danielle Citron and Benjamin Wittes suggest that the core benefits of CDA 230 could have been achieved without causing such significant damage to the lives of women and girls (particularly those who hold multiple intersecting identities) and their ability to fully participate in online spaces.

Second, the experiences of the *Allow States and Victims to Fight Online Sex Trafficking Act* and *Stop Enabling Sex Traffickers Act* (FOSTA-SESTA) in the United States demonstrate that it is imperative to listen to the vulnerable and marginalized populations who will be the most impacted by any proposed legislation. Lawmakers must consider the advice and insights of directly impacted communities in assessing whether a new law will disproportionately harm vulnerable and marginalized Internet users, including by driving them off of the Internet, rather than meaningfully address TFGBV itself. Experts in TFGBV and platform accountability have emphasized that any proposed platform liability regimes, as well as platforms' own policies and practices in content moderation, must be victim/survivor-centred and trauma-aware.

Third, Germany's experience with *Netzwerkdurchsetzungsgesetz* (NetzDG) suggests that relying on industry self-regulation is insufficient to meaningfully address TFGBV, for countries other than the United States where major American platform companies operate. The law's primary result appeared to be galvanizing greater enforcement of platforms' own community standards, whereas researchers found that a mass spike in wrongful takedowns did not happen, though the law continues to raise concerns. There was broad consensus that it was exceedingly difficult to determine whether the law effectively achieved its goal of reducing hate speech and other objectionable content, due to lack of meaningful data—highlighting that implementing a platform liability law may be of limited use without setting up a way to effectively evaluate its impact.

Fourth, jurisdictions such as the United Kingdom and European Union, through the *Online Harms White Paper* (and related documents) and proposed *Digital Services Act* (DSA) respectively, have begun to incorporate explicit recognition of the harms that digital platforms cause *systemically*. This is reflected in both regimes' tiered approaches, which place greater obligations on platforms beyond a certain size and influence; on the UK providing for 'super-complaints' to address systemic issues; and on the DSA's requirement that Very Large Online Platforms regularly assess and respond to systemic risk flowing from use of their services. Laws that address TFGBV and similar content on a systemic, not individualized, level are particularly important given the systemic nature of TFGBV as a pillar of structural discrimination and systemic oppression.

Fifth, digital rights advocates and those primarily concerned with generic freedom of expression (as opposed to the freedom of expression of those who are driven offline or forced to self-censor by TFGBV) have consistently raised valid concerns about platform liability regimes generally. Such concerns primarily focus on transparency and oversight; due process and appeal mechanisms; definitional clarity; and safeguards to mitigate wrongful takedowns or overbroad legislation. These concerns should be taken into account and given due weight, with measures to address them incorporated into any legal and policy reforms. Simultaneously, reforms must still apply an intersectional feminist analysis and focus on upholding substantive equality.

Constitutional and Critical Analysis of Platform Liability for TFGBV

Canadian constitutional and human rights law has repeatedly recognized the necessity and justifiability of limiting free expression in order to uphold equality rights and protect historically marginalized groups. The rights to equality and freedom from discrimination are as fundamental as freedom of expression, and equally protected under the *Canadian Charter of Rights and Freedoms*. This legal understanding must govern legislative reforms to address TFGBV, including through platform regulation and platform liability.

Multiple decisions from the Supreme Court of Canada, including *R v Taylor*, *R v Keegstra*, and *Saskatchewan Human Rights Commission v Whatcott*, in addition to *Lemire v Canada (Human Rights Commission)* at the Federal Court of Appeal, have affirmed the constitutionality of criminal and statutory human rights laws prohibiting hate speech, including hate speech published and distributed online. Restricting expression-based abuse that directly targets and silences marginalized communities, or results in their members self-censoring, directly promotes and protects both freedom of expression *and* equality.

The platform liability context can be distinguished from circumstances that gave rise to much of the leading hate speech jurisprudence. This is due to the position of the intermediary, which is typically at least one step removed from the actual speaker or publisher at the centre of most case law in this area. The all-important layer of users whose expression is facilitated by online platforms must not be ignored in the equation, and precedents cannot necessarily be applied directly from speaker (or publisher) to platform. Establishing a platform liability regime will require considering issues such as:

- the risks of overbroad removal of legitimate, beneficial, or legal content;
- issues arising from potential privatized governance of user speech and public discourse; and
- how a platform liability framework would account for the wide range of platform companies, which vary widely by size, nature, purpose, audience, business model, and content, among other relevant factors.

Still, the reasoning and principles supporting the constitutionality of Canadian hate speech prohibitions remain highly relevant in the context of TFGBV. Law and context combine to justify legal reforms that would impose some degree of legal obligation or liability on digital platforms for TFGBV by a user. The most effective legal reforms would account for the distinct role of digital platforms in the Internet ecosystem, and as differentiated from the direct perpetrator of TFGBV, while simultaneously recognizing that digital platforms do play a facilitative role—and sometimes more—in the devastating and widespread perpetuation of TFGBV.

Guiding Priorities and Recommendations for Federal Action

This report provides 14 recommendations for federal action, including legislative reform. These recommendations are based on six guiding priorities that emerged from the research and analysis conducted in this report and should govern efforts to address TFGBV in Canadian law. These priorities are:

- recognizing a need for legal reform to address TFGBV, including through platform regulation;

- recognizing that Canadian constitutional law justifies imposing proportionate limits on freedom of expression in order to uphold and protect the rights to equality and freedom from discrimination, and also to give full effect to the core values underlying freedom of expression;
- guaranteeing that legal reforms that address TFGBV build in victim/survivor-centered, trauma-informed, and intersectional feminist perspectives;
- ensuring expedient, practical, and accessible remedies for those targeted by TFGBV;
- providing due process mechanisms to users who wish to contest platforms' content moderation decisions (whether a decision to leave up or take down content); and
- requiring transparency from platform companies regarding their content moderation policies and decisions, as well as the outcomes of such policies and decisions concerning TFGBV.

Recommendations for Federal Action

A. Centering Human Rights, Substantive Equality, and Intersectionality

1. Apply a principled human rights-based approach to platform regulation and platform liability, including giving full effect to the rights to equality and freedom from discrimination.
2. Ensure that legislation addressing TFGBV integrates substantive equality considerations and guards against exploitation by members of dominant social groups to silence expression by members of historically marginalized groups.
3. When pursuing legislative or other means of addressing TFGBV, consult substantively with and take into account the perspectives and lived experience of victims, survivors, and those broadly impacted by TFGBV.

B. Legislative Reforms

4. Establish a centralized expert regulator for TFGBV specifically, with a dual mandate: a) to provide legal remedies and support to individuals impacted by TFGBV on digital platforms, including regulatory and enforcement powers; and b) to develop research on TFGBV and provide training and education to the public, relevant stakeholders, and professionals.
5. Enact one or more versions of the current 'enabler' provision in subsections 27(2.3) and 27(2.4) of the *Copyright Act*, adapted to specifically address different forms of TFGBV, including 'purpose-built' platforms.
6. Enact a law that allows for victims/survivors of TFGBV to obtain immediate removal of certain clearly defined kinds of content from a platform *without* a court order, such as NCDII.
7. Ensure that legislation to address TFGBV focuses solely on TFGBV (including intersectional considerations)—do not dilute, compromise, or jeopardize the constitutionality of such legislation by 'bundling' TFGBV with other issues that the government may wish to also address through platform regulation.

C. Legal Obligations for Platform Companies

8. Require platform companies to provide to users *and non-users* clearly visible, easily accessible, plain-language complaint and abuse reporting mechanisms to expediently address and remedy instances of TFGBV.
9. For ‘purpose-built’, ‘enabling’, or otherwise TFGBV-dedicated platforms, and where a clearly delineated threshold of harm is met, provide that an order to remove specific content on one platform will automatically apply to any of that platform’s parent, subsidiary, or sibling platform companies where the same content also appears.
10. Require platform companies to undergo independent audits (which could be conducted by the new TFGBV agency) and publish comprehensive annual transparency reports.
11. When determining legal obligations for digital platforms, account for the fact that platforms vary dramatically in size, nature, purpose, business model (including non-profit), extent of intermediary role, and user base.

D. Research, Education, and Training

12. Fund, make widely available, and mandate (where appropriate) education resources and training programs in TFGBV, which include information on how to support those who are subjected to TFGBV.
13. Fund frontline support workers and community-based organizations working to end, and supporting victims/survivors of, gender-based violence, abuse, and harassment, specifically to enhance their internal expertise, resources, and capacity to support those impacted by TFGBV (which often accompanies gender-based violence and abuse).
14. Fund further empirical, interdisciplinary, and law and policy research by TFGBV scholars, other TFGBV experts, and community-based organizations on TFGBV and the impacts of emerging technologies on those subjected to TFGBV.