

FREE EXPRESSION V. SAFE SPACE: BILL C-63'S BALANCING ACT FOR A BAND-AID SOLUTION

Dana Ann Lucas*

ABSTRACT

This Article explores Bill C-63, Canada's proposed Online Harms Act, examining how it aims to create a safer online environment while balancing fundamental freedoms, particularly the freedom of expression guaranteed under section 2(b) of the Canadian Charter of Rights and Freedoms. The Article argues that while certain aspects of the Bill will likely be upheld as justifiable limitations on the freedom of expression, other provisions may constitute unjustifiable infringements. The first part of the Article will provide an overview of social media, its evolution, and the prevalent harms it facilitates, including: cyberbullying, sexual exploitation, and the spread of extremist ideology. Second, the Article will analyze Bill C-63's key components, including its regulatory framework overseen by the Digital Safety Commission of Canada and the duties it imposes on social media operators. The Article proposes that the Bill's four duties—to take action, protect children, make harmful content inaccessible, and keep the necessary records—are likely justifiable under section 2(b) of the Charter, given the strong social interest in protecting vulnerable groups and preventing harm, aligning with the principles of a free and democratic society. However, the Article raises concerns about the Bill's provisions that amend existing statutes, particularly regarding the expanded definition of hate

* LL.M., 2025, Georgetown University Law Center (Technology Law and Policy); J.D., 2023, University of Saskatchewan; B.A., 2023, University of Saskatchewan; Strategic Space Law Certificate, 2023, McGill University. The author is a practicing attorney with experience and an educational background in intellectual property law, technology law, international trade law, space law, corporate law, and artificial intelligence governance.

speech and confidentiality provisions, finding that such amendments may be overly broad, subjectively applied, and could stifle legitimate expression, especially from marginalized groups. The Article concludes that while government regulation is necessary to combat online harms, it alone is insufficient. As such, meaningful change requires addressing the underlying business models of social media platforms that incentivize the spread of harmful content for profit. Until this is done, online harms will continue to propagate.

ABSTRACT	25
INTRODUCTION	27
II.SOCIAL MEDIA AND ONLINE HARMS IN CANADA	28
A. <i>The History of the Rise of Social Media</i>	28
B. <i>Harms Facilitated Through Social Media Platforms</i>	32
III.BILL C-63: CANADA’S ONLINE HARMS ACT.....	37
A. <i>Proposed Duties of Social Media Operators</i>	38
B. <i>Amendments to the Code</i>	42
C. <i>Amendments to Civil Statutes</i>	45
IV.THE CANADIAN CHARTER OF RIGHTS AND FREEDOMS AND SECTION 2(B)	49
V. WAYS BILL C-63 MAY ENGAGE SECTION 2(B) OF THE CHARTER	52
A. <i>Mandatory Content Moderation and Removal</i>	53
B. <i>Compelled Speech</i>	57
C. <i>Hate Speech</i>	59
D. <i>The Open Court Principle and Issues of Confidentiality</i> 61	
VI.THE RIGHT PATH FORWARD? APPLAUDING THE RIGHT EFFORTS IN THE WRONG DIRECTION	63
VII.CONCLUSION	68

INTRODUCTION

The widespread infiltration of social media in contemporary society maintains a pervasive presence, shaping everyday behaviors, thoughts,¹ and social interactions. Various social media platforms have transformed the way individuals communicate and interact with each other, connecting people and facilitating the exchange of information on a global scale.² That said, this mass adoption has exposed population-wide vulnerabilities, leading to severe online and offline harms.³ In response to these exposures, the Canadian federal government introduced Bill C-63, also known as the Online Harms Act, to foster a safer and more secure online environment.⁴ Despite seeking to establish a framework of duties and obligations for social media networks that operate in Canada, the Bill has caused

1. See Martin Korte, *The Impact of Digital Revolution on Human Brain and Behavior: Where Do We Stand?*, 22 DIALOGUES IN CLINICAL NEUROSCIENCE 101, 101 (2020) (“Neurological consequences” due to social media may include “internet/gaming addiction, language development, and processing of emotional signals.”).

2. See Linda C. Ashar, *Social Media Impact: How Social Media Sites Affects Society*, BUS. & MGMT. BLOG AM. PUB. UNIV. (May 2, 2024), <https://www.apu.apus.edu/area-of-study/business-and-management/resources/how-social-media-sites-affect-society/> [<https://perma.cc/V8UL-5N5M>] (discussing the advent of social media and its general effects on society).

3. See Brian Fishman, *Dual-Use Regulation: Managing Hate and Terrorism Online Before and After Section 230 Reform*, BROOKINGS INST. (Mar. 14, 2023), <https://www.brookings.edu/articles/dual-use-regulation-managing-hate-and-terrorism-online-before-and-after-section-230-reform/> [<https://perma.cc/H5TT-A7WR>] (discussing in part the long-standing use of social media by extremist groups, including terrorists and white supremacists, to communicate and spread their propaganda on an international scale); see also Darcy Hango, *Online Harms Faced by Youth and Young Adults: The Prevalence and Nature of Cybervictimization*, STATISTICS CANADA – INSIGHTS ON CANADIAN SOCIETY, 1, 2 (Feb. 21, 2023), <https://www150.statcan.gc.ca/n1/en/pub/75-006-x/2023001/article/00003-eng.pdf?st=HotG34F8> (reporting that 25% of Canadian teenagers experienced online bullying in the previous year).

4. See DEP’T. OF JUST., BILL C-63: AN ACT TO ENACT THE ONLINE HARMS ACT, TO AMEND THE CRIMINAL CODE, THE CANADIAN HUMAN RIGHTS ACT AND AN ACT RESPECTING THE MANDATORY REPORTING OF INTERNET CHILD PORNOGRAPHY BY PERSONS WHO PROVIDE AN INTERNET SERVICE AND TO MAKE CONSEQUENTIAL AND RELATED AMENDMENTS TO OTHER ACTS, <https://www.justice.gc.ca/eng/csj-sjc/pl/Charter-charte/c63.html> [<https://perma.cc/FBJ6-GZZH>] (last modified June 4, 2024) (Can.) [hereinafter DEP’T. OF JUST. (Can.)] (“Bill C-63 contains a variety of measures to address a range of harmful content online as well as hate speech and hate crimes both online and offline.”).

public scrutiny by introducing new legislation and amending existing statutes that could undermine the freedom of expression guaranteed under section 2(b) of the Canadian Charter of Rights and Freedoms (the “Charter”).⁵

As such, this Article will first provide a brief overview of social media networks, followed by a discussion of prevalent harms exacerbated through consistent social media use. Secondly, Bill C-63’s components will be introduced, accompanied by the scope of the Charter, with a particular emphasis on section 2(b). Next, the focus will shift to the intersection of Bill C-63 and the freedom of expression. Lastly, the Article will close with a review, drawing information from legal and non-legal sources, and determining which provisions of the Bill are likely to violate the Charter beyond a justifiable and reasonable limit. Seeking to contribute to the discussions on the nature of the freedom of expression in the social media sphere, the impact of government regulation in light of current engagement-priority business models will be addressed. In summary, this Article aims to weigh the protection of Canadians from online harms against the need to safeguard the freedom of expression, ultimately underscoring that despite government oversight being necessary for standard-setting and deterrence, it alone remains insufficient to address the root cause of the business models that social media platforms operate on.

II. SOCIAL MEDIA AND ONLINE HARMS IN CANADA

A. *The History of the Rise of Social Media*

Today, social media platforms can be broadly referred to as web-based applications that promote the creation and exchange of user-generated content.⁶ More specifically, the common conception of social media platforms can be broken down into

5. Canadian Charter of Rights and Freedoms, Part I of the Constitution Act, 1982, being Schedule B to the Canada Act, 1982, c. 11 (U.K.) [hereinafter Charter of Rights]; see Safiyya Ahmad, *What’s in Bill C-63, and why we are alarmed*, B.C. CIV. LIBERTIES ASS’N (Sep. 30, 2024), <https://bccla.org/2024/09/whats-in-bill-c-63-why-are-we-alarmed/> (arguing that “Bill-C-63 will create draconian rules, cause chilling effects on free expression meaning that people may fear speaking out even when they have a constitutional right to do so.”).

6. See Ashar, *supra* note 2.

three parts where an internet medium: (1) allows for the registration or creation of an account or a profile to create, view, or share user-generated content; (2) permits one or more users to create content other users can view; and (3) primarily functions as a platform for user interaction with content generated by other users.⁷ Essentially, these platforms are “forms of electronic communication (such as websites for social networking and microblogging) through which users create online communities to share information, ideas, personal messages, and other content (such as videos).”⁸ Setting these platforms apart from traditional forms of communication is interconnectivity.⁹ Many traditional media platforms operate on a one-to-many model.¹⁰ Alternatively, social media permits mutual engagement between users, where an ‘online community’ is fostered through users interacting with user-created content.¹¹ Here, the bi-directional nature of communication is significant.

In its beginning stages, social media platforms were significantly less interactive, dynamic, captivating, and

7. NSF support of research on impacts of social media on human trafficking. 42 U.S.C. § 1862w(a)(2).

8. *Social Media*, MERRIAM-WEBSTER, <https://www.merriam-webster.com/dictionary/social%20media> [<https://perma.cc/H4ML-45GL>] (last visited Sep. 21, 2025).

9. See Alexander S. Gillis, *Traditional Media vs. New Media: Differences, Pros and Cons*, INFORMA TECHTARGET (Dec. 5, 2024), <https://www.techtargget.com/whatis/feature/Traditional-media-vs-new-media-Differences-pros-and-cons>.

10. See *id.* (noting that traditional media is characterized as having a “top-down communication model” and “audiences are mostly restricted to passively receiving content”); see also Ionica Oncioiu et al., *Fundamental Power of Social Media Interactions for Building a Brand and Customer Relations*, 16 J. THEORETICAL APPLIED ELEC. COM. RESEARCH 1702 (2021). Many traditional media platforms operate on a ‘one-to-many’ model, where a single content creator or broadcaster transmits information to a large, passive audience. This refers to a communication structure where content flows in one direction, from a single source (like a TV station or newspaper), to many recipients who primarily consume rather than interact with the content.

11. See Gillis, *supra* note 9 (“One of the defining characteristics of new media is how it enables users to be active participants. Not only can users consume content, but they can create and post their own content.”); see also Oncioiu et al., *supra* note 10 (“In other words, the mutual interaction between those involved in the dialogue is the precondition for a functional dialogue, reflected in the fact that companies mainly use social networks to contact customers and to build and maintain relationships with them.”).

addictive.¹² The term *social media* was first coined in *Matisse*, a Tokyo online media environment, in 1994.¹³ The earliest identified forms emerged in the late twentieth century.¹⁴ Platforms such as *CompuServe*, *Prodigy*, and *America Online* provided users with a new method to connect via email, chat rooms, and bulletin board messaging.¹⁵ Despite being largely text-based with limited multimedia capabilities, online communities fostered digital interactions, laying the foundation for modern-day social media platforms.¹⁶ In 1997,¹⁷ the launch of *SixDegrees.com* marked a short-lived but pivotal moment in the evolution of social media.¹⁸ *SixDegrees.com* established the blueprint for future networking sites, allowing users to create profiles, build connections with friends, and browse pages.¹⁹

The early 2000s introduced an explosion of new social media platforms catering to differing interests and audiences. *Friendster* gained initial popularity by encouraging users to build virtual communities through their real-world friendships; this

12. See Kimberly Young, *Internet Addiction Over the Decade: A Personal Look Back*, WORLD PSYCH. (June 9, 2010), <https://pmc.ncbi.nlm.nih.gov/articles/PMC2911082/> (discussing how the internet has changed, as well as increased recognition and prevalence of internet addiction).

13. Thomas Aichner et al., *Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019*, 24 CYBERPSYCH., BEHAV. & SOC. NETWORKING 215, 215 (2021) ("It was in these early days of the commercial Internet that the first SM platforms were developed and launched. Over time, both the number of SM platforms and the number of active SM users have increased significantly, making it one of the most important applications of the Internet.").

14. *Id.*

15. See Saqib Shah, *The History of Social Networking*, DIGIT. TRENDS MEDIA GRP. (Oct. 21, 2022), <https://web.archive.org/web/20241206194541/https://www.digitaltrends.com/computing/the-history-of-social-networking/>; see also *The Evolution of Social Media: How Did It Begin, and Where Could It Go Next?*, MARYVILLE UNIV. (May 28, 2020), <https://online.maryville.edu/blog/evolution-social-media/>.

16. See MARYVILLE UNIV., *supra* note 15 (explaining that although rudimentary, these platforms enabled "basic online networking.").

17. See Irfan Ahmad, *The History of Social Media [Infographic]*, SOC. MEDIA TODAY (Apr. 27, 2018), <https://www.socialmediatoday.com/news/the-history-of-social-media-infographic-1/522285/> [<https://perma.cc/X6L2-M56R>] (discussing the creation of Six Degrees).

18. See Shah, *supra* note 15 (discussing the decline of Six Degrees.com and how it inspired future social media platforms).

19. *Id.*

implemented a “Circle of Friends”²⁰ model, which made it easier for users to find and connect with real-life acquaintances on the website.²¹ Similarly, *LinkedIn* launched in 2003 to connect users in a career-oriented space through online professional networking.²² Also in 2003, *MySpace* gained traction among young adults, permitting users to personalize profiles with photos, music, and custom layouts.²³ Such user-generated content on a multi-media platform foreshadowed the interactive nature of social media companies today.

In 2004, the launch of *Facebook* led to a significant shift in the social media landscape.²⁴ Originally intended for Harvard University students, the user-friendly interface, innovative features (e.g., like button and newsfeed), and the prioritization of real-life friends and family led to the platform’s global expansion.²⁵ *Facebook* became a dominant force in 2009 and today represents one of the most common social media platforms.²⁶ Driven by a continued evolution of technological advantages, evolving user preferences, algorithmic amplifications, and mobile-first platforms (e.g., *Snapchat*, *Instagram*, and *TikTok*),

20. *Id.*

21. *Id.*

22. See *The History of Social Media (1844-2018)* - #infographic, DIGIT. INFORMATION WORLD (Apr. 26, 2018), <https://www.digitalinformationworld.com/2018/04/infographic-timeline-of-social-networking.html> (discussing history of social media from 1844 to 2018 generally).

23. See Shah, *supra* note 15; see also *Myspace*, ENCYC. BRITANNICA, <https://www.britannica.com/topic/Myspace> (last updated Aug. 30, 2025).

24. See Shah, *supra* note 15 (discussing how certain features of Facebook were widely adopted by other new media companies after its massive success, and how Facebook led other companies to recognize “the power of social networking”).

25. See *id.* (detailing the history of Facebook and explaining that it remains the dominant social media network across the globe); see also Mark Hall, *Facebook*, ENCYC. BRITANNICA (Sep. 5, 2025, 7:13 AM), <https://www.britannica.com/money/Facebook> (discussing the unique features that Facebook pioneered).

26. See Hall, *supra* note 25; see also Shah, *supra* note 15 (explaining that Facebook led the “global social networking pack”); see also Bobbie Johnson, *Why 2009 was Facebook’s year*, THE GUARDIAN (Dec. 28, 2009), <https://www.theguardian.com/technology/2009/dec/28/facebook-users-social-network> (explaining the significant growth Facebook saw in 2009).

new access and engagement measures revolutionized the adoption of social media.²⁷

Today, the scope of these platforms extends further than mere social media services. In less than a generation, social media evolved from a simple digital information exchange to an imperative 21st-century marketing and connectivity tool.²⁸ Currently, over five billion individuals are reported to use social media platforms, with an average user accessing over six social media platforms monthly.²⁹ Social media has become a vast informal network of online spaces, hosting connections for family celebrations, sports clubs, political campaigns, charitable initiatives, support groups, professional networking, information sharing, fundraising, and more.³⁰ Changing the way society functions, social media platforms carry several societal benefits alongside several significant drawbacks.

B. Harms Facilitated Through Social Media Platforms

“[W]oven into the fabric” of contemporary society, those “under [thirty] represent the first generation born into a [population] where internet use was already ubiquitous.”³¹ As such, the pervasiveness of social media platforms has amounted to various forms of severe online harm. For instance, in 2019, 25% of Canadian teenagers reported being subjected to online bullying in the previous year.³² Additionally, online sexual exploitation has become a prevalent issue, often leading to severe and long-lasting

27. See MARYVILLE UNIV., *supra* note 15 (“The invention of the smartphone liberated social media from the desktop and laptop computer. Apple’s first iPhone, launched by Steve Jobs in 2007, helped shift the focus of online community building to mobile. Facebook, Twitter, Snapchat, Instagram, TikTok, and other social media services thrived in the mobile app environment.”).

28. *Id.* (“In less than a generation, social media has evolved from direct electronic information exchange, to virtual gathering place, to retail platform, to vital 21st-century marketing tool.”).

29. *Id.*; see also *Social Media Usage & Growth Statistics*, BACKLINKO (Feb. 10, 2025), <https://backlinko.com/social-media-users>.

30. See Ashar, *supra* note 2 (“Social media wields cultural influence on fashion and food trends, family and adolescent health issues, world news and local events, political and community action events.”).

31. See Hango, *supra* note 3, at 2.

32. *Id.*

impacts on victims.³³ Offline self-harm has further been found as a co-variant with increased social media use among youth.³⁴ Likewise, terrorist and extremist groups have been found to exploit these platforms to propagate harmful ideologies and incite violence.³⁵ As these harms continue to prevail in society, calls for content regulation to mitigate risks have grown increasingly louder.³⁶

The accessibility and common use of these social media platforms has led to new opportunities for online aggression. A prominent example is the increase in online bullying.³⁷ Online bullying, or cyberbullying, is defined as the use of technology to intentionally threaten, harm, target, or embarrass another.³⁸ Social media's widespread use, anonymity, and persistence of

33. Xavier V. Caddle et al., *Duty to Respond: The Challenges Social Service Providers Face When Charged with Keeping Youth Safe Online*, 7 PROC. ACM ON HUM.-COMPUT. INTERACTION 6:1, 6:8–14 (2023) (“Overall, online sexual risks . . . and cyberbullying were the most salient concern across all social service domains, especially when these experiences crossed the boundary between the digital and physical worlds.”).

34. *Id.* (“Experiencing other forms of victimization was more common among those who were cybervictimized. This includes being stalked and being physically or sexually assaulted, and experiencing other types of unwanted behaviours in public.”).

35. See Alexander Tsesis, *Social Media Accountability for Terrorist Propaganda*, 86 FORDHAM L. REV. 605, 605 (2017) (“Terrorist organizations have found social media websites to be invaluable for disseminating ideology, recruiting terrorists, and planning operations. National and international leaders have repeatedly pointed out the dangers terrorists pose to ordinary people and state institutions.”); see also Fishman, *supra* note 3 (discussing, similarly, terrorist and white supremacist groups and their use of social media).

36. See generally Tyler Hogue et al., *The Boundaries of Internet Speech*, REGUL. REV. (May 18, 2024), <https://www.theregreview.org/2024/05/18/the-boundaries-of-internet-speech/> (discussing differing views on the current state of regulation of hate speech online, as well as efforts from both private and public sectors in the U.S. to address the issue).

37. See Caddle et al., *supra* note 33, at 4, 8, 12 (discussing how cyberbullying is one of the most concerning risks for social service providers because it is among “the most prevalent [of] online risks to which underprivileged youth [are] exposed”).

38. *Cyberbullying*, NEMOURS KID’S HEALTH, <https://kidshealth.org/en/teens/cyberbullying.html> (last updated Aug. 2022) (defining cyberbullying); see also *One in six school-aged children experiences cyberbullying, finds new WHO/Europe study*, WORLD HEALTH ORG. (Mar. 27, 2024), <https://www.who.int/europe/news/item/27-03-2024-one-in-six-school-aged-children-experiences-cyberbullying-finds-new-who-europe-study> (noting that the increase of cyberbullying has been “magnified by the increasing digitalization of young people’s interactions”).

content easily foster breeding grounds for online bullying.³⁹ Perpetrators exploit social media platforms to spread hurtful content, harmful rumors, and outright threats⁴⁰—all of which severely harm the victims' mental and physical health.⁴¹ Cyberbullying can lead to intense feelings of helplessness, despair, anxiety, and depression in young victims.⁴² Moreover, social service providers face challenges in addressing this online risk as they mainly rely on self-reporting and tend to lack sufficient understanding of platform operations.⁴³

Studies have also indicated a strong correlation between social media use and self-harm. Specifically, adolescents with a history of non-suicidal self-injury ("NSSI") tend to be more active on social media compared to their peers with no such history.⁴⁴ Online interactions, particularly those involving NSSI communications, negatively impact the well-being of vulnerable youth.⁴⁵ Social networking platforms "contribute to increased exposure to and engagement in self-harm behavior, as users tend to emulate the self-injurious behavior of others online, adopt self-injurious practices from self-harm videos, or are encouraged and acclaimed by others."⁴⁶ Social media normalizes these behaviors, and greater use of social media is associated with

39. See NEMOURS KID'S HEALTH, *supra* note 38 (explaining that cyberbullying can be "particularly damaging" as it is usually anonymous and difficult to control because the content can be spread to many people and victims can be harassed every time they check their phone or computer).

40. See Caddle et al., *supra* note 33, at 13.

41. *Id.* at 9.

42. See Charisse Nixon, *Current Perspectives: The Impact of Cyberbullying on Adolescent Health*, *Adolescent Health*, MED. THERAPEUTICS 143, 144 (2014) ("results indicate that there is a significant relationship between cybervictimization and depression among adolescents... 93% of cybervictims reported negative effects, with the majority of victims reporting feelings of sadness, hopelessness, and powerlessness.").

43. *Id.* at 152.

44. See Aksha M. Memon et al., *The Role of Online Social Networking on Deliberate Self-Harm and Suicidality in Adolescents: A Systematized Review of Literature*, 60 J. OF PSYCHIATRY 384, 385 (2018) ("Adolescents with a history of NSSI are more active on social media than adolescents with no NSSI history. Adolescents are actively engaged in online video sharing, and these web-based interactions may involve communication regarding NSSI, which could adversely affect their well-being and behavior.").

45. *Id.*

46. *Id.* at 19.

increased suicidal ideation, psychological distress, and unmet mental health support needs in youth and young adults.⁴⁷

General advantages to social media platforms (e.g., user open access and quick transmission of information⁴⁸) become grave concerns regarding sexual exploitation. In 2012, Canadian teenager Amanda Todd committed suicide after being sexually exploited and harassed on *Facebook* by a Dutch predator.⁴⁹ Before her death, she posted a viral video about her online sexual exploitation.⁵⁰ Importantly, Amanda Todd's experience is not unique, as thousands of websites are dedicated solely to sharing non-consensual sexual content, particularly targeting women and children.⁵¹ Of the nearly 85,000 deepfakes identified to circulate online, 96% of them are non-consensual, sexually explicit content.⁵² Social media platforms have been found to facilitate this exploitation as the easy-to-use platform interfaces on mobile devices enable predatory adults to groom children and other vulnerable groups by text and live streaming.⁵³ As such, online sexual exploitation through social media has only increased throughout the years.

Extremism, hate speech, and terrorist activities via social networking platforms tend to thrive due to the ease of information

47. See Candice Biernesser et al., *Social Media Use and Deliberate Self-Harm Among Youth: A Systematized Narrative Review*, 116 CHILD YOUTH SERV. REV. 1, 7 (2020) ("Daily use of social networking sites (e.g. Facebook, Twitter, etc.) of two hours or more was observed in 25% of high school students in Ontario and was associated with a five-fold increased odds of suicidal ideation, compared to youth who used social media for under two hours daily").

48. See Ashar, *supra* note 2.

49. See Edward Butts, *Amanda Todd Case*, THE CANADIAN ENCYCL. (Oct. 25, 2024), <https://thecanadianencyclopedia.ca/en/article/amanda-todd-case>.

50. *Id.* (explaining how her YouTube video used flashcards to tell her story of cyberbullying, online sexual orientation, abuse, and distress).

51. See Jennifer Laffier & Aalyia Rehman, *Deepfakes and Harm to Women*, 3 DIGIT. LIFE LEARN. 1, 7 (2023) ("more than 3000 websites are dedicated to the submission of non-consensual sexual media content... [compared to men] images of women are more commonly viewed, commented on, and shared."); see also INTERNET WATCH FOUND., *Annual Report 2022* 1, 75 (2023) (observing there were 5,416 second-level domains hosting child sexual abuse material in 2022).

52. See Laffier & Rehman, *supra* note 51, at 8.

53. See Ashar, *supra* note 2 (detailing how predatory adults can use social platforms to groom and exploit children through online communications, live streaming, and sexting).

dissemination and communication. The advent of social media platforms transformed how terrorists operate; firstly, it created a space for proactivity by terror supporters, and secondly, it removed the difficulties of geographical separation.⁵⁴ Offering extremist ideologies a global reach, social media platforms aid terrorist and extremist organizations in radicalizing and recruiting members.⁵⁵ Prior, traditional terror web forums relied on the potential convert to take the first step by seeking out the forum.⁵⁶ However, these platforms now enable terrorist organizations to proactively identify and target vulnerable individuals through strategically deployed and widespread outreach.⁵⁷ For instance, investigators discovered that one of the Boston Marathon terrorists was partially radicalized through frequent online videos shared on social media by a prominent al-Qaeda propagandist.⁵⁸ Terrorist groups exploit social media networks to disseminate messaging, incite violence, and cause public fear.⁵⁹

In 2015, a class action lawsuit filed against *Facebook* claimed that the platform permitted terrorists to incite violent attacks against Israeli citizens.⁶⁰ The lawsuit alleged that requests for the removal of these pages were denied despite violating *Facebook's*

54. See Anna Elisabeth Jayne Goodman, *When You Give a Terrorist a Twitter: Holding Social Media Companies Liable for Their Support of Terrorism*, 46 PEPP. L. REV. 147, 154 (2019).

55. *Id.* at 158.

56. *Id.* at 154.

57. *Id.* at 154–57.

58. See Scott Shane, *The Enduring Influence of Anwar al-Awlaki in the Age of the Islamic State*, CTC SENTINAL 15, 16, 18 (July 2016) (describing how Anwar al-Awlaki, a U.S.-born al-Qaeda propagandist, influenced the Boston Marathon bombers through his online lectures and writings); see also Tsesis, *supra* note 35, at 617–18 (indicating the scope and nature of ISIS's social media presence, including statements calling for emulation of terrorist acts – like the Boston Bombing).

59. See generally U.N. OFFICE ON DRUGS & CRIME, *Handbook on Children Recruited and Exploited by Terrorist and Violent Extremist Groups: The Role of the Justice System*, 10, 13–14 (2017), https://www.unodc.org/documents/justice-and-prison-reform/Child-Victims/Handbook_on_Children_Recruited_and_Exploited_by_Terrorist_and_Violent_Extremist_Groups_the_Role_of_the_Justice_System.E.pdf.

60. See *Israel-based rights group files lawsuit against Facebook for inciting violence*, BUS. STANDARD (Oct. 30, 2015), <https://perma.cc/49WT-P5DF> (discussing the lawsuit and depictions on various sites on Facebook).

community decency standards.⁶¹ Designed to curate personalized experiences per user, social media platforms inadvertently increase online radicalization by amplifying extremist content.⁶² Further, private messaging features and encrypted messaging applications enable terrorist organizations to communicate and coordinate, avoiding detection.⁶³ Obscuring the activities of terrorists, the procedural requirements, cost and time to break encryption technology, and the public policy debates over privacy rights versus security needs complicate safeguarding society from extremism, violence, and terrorism.

Evidently, social media platforms foster several online harms to society prevalent in Canada. This is done through exacerbating various forms of trauma (e.g., cyberbullying, sexual exploitation, and the sharing of intimate content without consent). Moreover, it can aid the promotion of extremist ideology, terrorism, violence, and hate speech. As such, the Canadian government has not been blind to these severe consequences. New legislation has been proposed that would create active duties, greater mitigation measures, and further regulation of the social media platforms that operate within Canadian borders.

III. BILL C-63: CANADA'S ONLINE HARMS ACT

On February 26, 2024, Bill C-63, the Online Harms Act, was introduced in the House of Commons.⁶⁴ The Bill imposes heightened accountability and transparency requirements on social media service providers across the country. Further, the Bill outlines several active duties, specifically the duty to take action, to protect children, to make harmful content inaccessible, and to keep the necessary records.⁶⁵ The Bill aims to reduce harmful

61. See Paul Goodman, *Facebook Sued After Richard Lakin Killed by Hamas Terrorists*, NBC News (July 17, 2016), <https://www.nbcnews.com/news/world/facebook-sued-after-richard-lakin-killed-hamas-terrorists-n609881>.

62. See U.N. OFFICE ON DRUGS & CRIME, *supra* note 59, at 13. In particular, one page, called "Stab Israelis", posted graphic images and called for violence. See Tsesis, *supra* note 35, at 609.

63. *Id.*

64. Online Harms Act, Bill C-63, 44th Parliament (2024) (Can.) (unenacted) [hereinafter Bill C-63].

65. *Id.* at Preamble, Part 4, §§ 54, 64–67, 72.

content on social media platforms by requiring social media operators to establish measures for reducing the risk of specific types of harmful content.⁶⁶ The Bill divides the targeted harmful content into seven categories, namely content that (1) sexually victimizes children or revictimizes a survivor, (2) is employed to bully children, (3) induces children to harm themselves, (4) incites violence, (5) incites violent extremism or terrorism, (6) foments hatred, and (7) is intimate communication without consent.⁶⁷ To enforce the framework, a Digital Safety Commission would be created to support users and victims.⁶⁸ Further, amendments to existing statutes, including the Criminal Code (the “Code”), the Canadian Human Rights Act (“CHRA”), and An Act Respecting the Mandatory Reporting of Internet Child Pornography by Persons who Provide an Internet Service, would be put in place.⁶⁹ Monetary penalties would also be imposed as an additional compliance mechanism.⁷⁰ As such, several changes to the current online legal regulation sector are introduced through Bill C-63.

A. Proposed Duties of Social Media Operators

Bill C-63 defines a social media service as “a website or application that is accessible in Canada, the primary purpose of which is to facilitate interprovincial or international online communication among users of the website or application by enabling them to access and share content.”⁷¹ Rather than attacking all social media services, Bill C-63 targets those that meet specific criteria, primarily related to user base size.⁷² As such, a “regulated service”⁷³ under the Bill refers to social media services

66. *Id.* § 55.

67. *Id.* § 2.

68. See DEPT. OF JUST. (Can.), *supra* note 4 (discussing how Bill C-63 would establish the Digital Safety Commission of Canada and the purpose of penalties imposed by the commission.).

69. *Id.*

70. *Id.* (discussing the purpose of the penalties in Bill C-63 would be to encourage compliance rather than for punishment.).

71. See Bill C-63, *supra* note 64, § 2(1) (defining “social media service”).

72. See *id.* § 3(1).

73. *Id.*

with a user base equal to or greater than a “significant number”⁷⁴ of users. The amount designated as the applicable threshold to constitute a significant number of users will vary based on the type of social media service, as determined by the regulations once the Bill is enacted.⁷⁵ Further, social media services with a smaller user base than the defined significant number can still be covered if they are specifically designated to be covered by the regulations.⁷⁶ Once enacted, the Governor in Council will be empowered to establish the different types of social media services,⁷⁷ determine the significant number for each type,⁷⁸ and outline the method for calculating user number bases.⁷⁹ That said, Bill C-63 expressly includes adult content services, as well as live streaming services.⁸⁰ More specifically, the Bill’s application is limited to public social media services and explicitly excludes private messaging features within a regulated service.⁸¹ As such, the Bill seemingly centers on the dynamics and potential harms associated with large-scale public communication, as opposed to smaller services and private messaging platforms.

To oversee the rights and duties of the covered social media services, the Bill proposes the establishment of two entities, namely the creation of the Digital Safety Commission of Canada (“DSCC”) and a digital safety ombudsman’s office.⁸² The DSCC will be tasked with overseeing and enforcing the regulatory framework of the Bill.⁸³ Its responsibilities would include handling user complaints, ensuring the removal of certain harmful content, developing new safety standards, and offering

74. *Id.*

75. *Id.*

76. *Id.*

77. See Mandy Lau, *An Overview of Canada’s Online Harms Act*, TECH POL’Y PRESS (Mar. 12, 2024), <https://www.techpolicy.press/an-overview-of-canadas-online-harms-act/> [<https://perma.cc/XNU8-RJDJ>].

78. See Bill C-63, *supra* note 64, § 2.

79. See *id.* § 5(1).

80. See *id.* § 6(1).

81. *Id.* at § 6. (defining a social media service as one that allows users to communicate content publicly, thereby excluding services that lack a public element).

82. See *id.* § 29.

83. See *id.* § 11.

educational resources to mitigate risks.⁸⁴ The implementation of a new digital safety ombudsman's office further aims to safeguard the public interest in addressing systemic online safety issues.⁸⁵ The ombudsman would be appointed to serve at the pleasure of the Governor in Council for a renewable term of no more than 5 years.⁸⁶ Their duties would involve consulting with users, issuing calls for contributions on specific issues, and highlighting significant or widespread online harm concerns to the relevant entities.⁸⁷ Following the establishment of these oversight bodies, the Bill also outlines specific responsibilities for social media service operators to ensure compliance with the new regulatory framework and protect users from online harm.⁸⁸

The Bill places obligations on the "operators" of a regulated service, defined as the person or entity running the service.⁸⁹ In doing so, Part I of the Bill outlines four fundamental duties for operators.⁹⁰ First, operators will have the duty to act responsibly.⁹¹ This requires operators to assess the risk of harmful content exposure that their platforms present and adopt the needed measures to reduce such risk.⁹² Further, they will need to provide users with guidelines, tools, and processes to block users and flag harmful content.⁹³ They will also need a resource service to receive and provide advice for user complaints, including a designated contact person for this purpose.⁹⁴ Moreover, operators will be required to label content they

84. *See id.* § 10–11.

85. Bill C-63, *supra* note 64, at § 31.

86. *See id.* § 30.

87. *See* Charlotte Simard-Zakaib et al., *Bill C-63: The Government of Canada Intends to Establish a New Responsibility Regime for Social Media Operators and Direct Content Distributors*, MCCARTHY TÉTRAULT (Mar. 26, 2024), <https://www.mccarthy.ca/en/insights/publications/bill-c-63-government-canada-intends-establish-new-responsibility-regime-social-media-operators-and-direct-content-distributors>.

88. *See* Bill C-63, *supra* note 64, § 54.

89. *Id.* § 2(1) (defining "operator" and explaining the creation of two new bodies to oversee implementing Bill C-63).

90. *Id.* §§ 54, 64, 68(4), 72.

91. *Id.* §§ 54, 55.

92. *Id.* § 55(1).

93. *Id.* § 57-59.

94. Bill C-63, *supra* note 64, § 61.

reasonably believe is disseminated by bots, where the content garnered prominence through such automated amplification.⁹⁵ Second, operators will have a duty to protect children.⁹⁶ Operators will be required to implement age-appropriate design features (e.g., warning labels or parent-control default settings).⁹⁷ The service design features must be in line with the measures and guidelines established by the DSCC once the Bill is enacted.⁹⁸ In addition to outlining proactive duties, the Bill also imposes reactive duties.

Sharing a common goal to reduce user exposure to harmful content, the duty to restrict access to certain content is similar to the aforementioned duties. As the third duty, the Bill requires operators to make certain content inaccessible.⁹⁹ Operators must block their users' access to two types of content considered the most harmful under the Bill; specifically, the Bill outlines a duty to make child sex abuse material and non-consensually distributed intimate images, including deepfakes of a sexual nature, inaccessible. This block of access must occur within 24 hours or any other period prescribed by the regulation.¹⁰⁰ The fourth duty imposed through Bill C-63 is the duty to keep records.¹⁰¹ Operators would be required to keep all records, including information and data, that are necessary to determine whether they are complying with their duties under the Bill.¹⁰² Importantly, all of these duties aim to foster a safer and more responsible online environment, focusing on potential harms arising from social media use and the best means to mitigate such harms.

To reinforce and strengthen compliance with the Bill, penalties are directly linked to the operator's duties. Specifically, failure to comply with the Bill can result in a maximum administrative monetary penalty for operators of not more than

95. *Id.* § 60.

96. *Id.* § 64.

97. *Id.* § 65.

98. *Id.* § 66.

99. *Id.* § 67.

100. Bill C-63, *supra* note 64, §§ 68(1), 68(5).

101. *Id.* § 72.

102. *Id.*

6% of the gross global revenue of the operator deemed to have violated the Bill or \$10 million Canadian dollars (CAD), whichever is greater.¹⁰³ Further, operators convicted of an offence will be liable on conviction on indictment to a fine no more than either 8% of their gross global revenue or CAD \$25 million (whichever is greater); or on summary conviction to a fine not exceeding 7% of their gross global revenue or CAD \$20 million (whichever is greater).¹⁰⁴ Likewise, the DSCC has the discretion to order an operator to publish a public notice detailing a violation of the Bill or an undertaking entered with the DSCC.¹⁰⁵ These penalties and compliance mechanisms are designed to ensure that operators take their responsibilities under the Bill seriously, with the financial and publishing consequences underscoring the Bill's commitment to holding businesses accountable for protecting users from online harms.

B. Amendments to the Code

Part II of Bill C-63 brings significant changes to the Code, which includes a new definition of hatred, a new hate crime, and stronger sentences for existing hate propaganda offences. These changes are targeted at the users of social media platforms and not the operators.¹⁰⁶ Currently, there is no explicit definition of hatred in the Code.¹⁰⁷ Rather, this has been historically inferred through case law and judicial interpretation as extreme emotional reactions that go beyond mere dislike or criticism.¹⁰⁸ Bill C-63 would introduce a new explicit definition for hatred regarding hate propaganda offences per sections 318 and 319 of

103. See DEPT. OF JUS. (Can.), *supra* note 4 (explaining that the Commission is allowed to impose a penalty in an amount up to a maximum of 6% of an operator's global revenues or \$10 million, whichever is higher).

104. See Bill C-63, *supra* note 64, § 120.

105. *Id.* § 94(1).

106. See *id.* § 14 (amending sections 318-319 of the Criminal Code to include hate propaganda offences); see also *id.* at § 15 (creating a new hate crime offence under section 320.1001).

107. There is no definition for the term "hatred" in the Canadian Criminal Code. See Criminal Code, R.S.C. 1985, c. C-46 (Can.) [hereinafter Canada Criminal Code].

108. See Mark J. Freiman, *Hate Speech and The Reasonable Supreme Court of Canada*, 63 SUPREME COURT L.R. 299, 300 (2013) (discussing the judicial interpretation of hatred that the Supreme Court used to have in Canada).

the Code (i.e., advocating genocide per section 318 and promotion of hatred per section 319).¹⁰⁹ This would broadly describe hatred as “detestation or vilification”¹¹⁰ that extends beyond mere humiliation, hurt, or offence.¹¹¹ Notably, while the current interpretations could extend to online platforms, they do not explicitly outline or identify the concept of digital hate speech.¹¹² As such, the new definition would outright ensure the inclusion of online hate speech.¹¹³

The new stand-alone criminal offence titled “offence motivated by hatred”¹¹⁴ under the Bill would be based on “race, national or ethnic origin, language, color, religion, sex, age, mental or physical disability, sexual orientation or gender identity or expression.”¹¹⁵ Before this, the hate propaganda offences under the Code focused on criminalizing hate speech and the incitement of hatred, particularly through speech or publications that advocate violence or discrimination against identifiable groups.¹¹⁶ This new offence under Bill C-63 would expand the Code and allow application to any criminal act (e.g., assault or vandalism) where hatred is the motive.¹¹⁷ Applying more broadly than the current (and quite infrequently applied) crimes of hate propagation,¹¹⁸ the Bill expands the scope of the

109. See Canada Criminal Code, *supra* note 107, §§ 318, 319.

110. Bill C-63, *supra* note 64, § 14(4).

111. *Id.* § 14(5).

112. See generally Canada Criminal Code, *supra* note 107 (showing that the Code makes no mention of the concept of digital hate speech).

113. Bill C-63, *supra* note 64, at iii (“Part 3 amends the Canadian Human Rights Act to provide that it is a discriminatory practice to communicate or cause to be communicated hate speech by means of the Internet or any other means of telecommunication in a context in which the hate speech is likely to foment detestation or vilification of an individual or group of individuals on the basis of a prohibited ground of discrimination.”).

114. *Id.* § 11.

115. Lau, *supra* note 77.

116. See Bill C-63, *supra* note 64, § 319.

117. *Id.* (“320.1001 (1) Everyone who commits an offence under this Act or any other Act of Parliament, if the commission of the offence is motivated by hatred based on race, national or ethnic origin, language, colour, religion, sex, age, mental or physical disability, sexual orientation, or gender identity or expression, is guilty of an indictable offence and liable to imprisonment for life.”).

118. See Jason Proctor, *The difficulty of prosecuting hate in Canada*, CBC NEWS (Sep. 7, 2025, 1:34 PM), <https://www.cbc.ca/news/canada/british-columbia/racists-attacks->

offence beyond mere avenues of verbal speech to address hate-driven crimes in both offline and online contexts.¹¹⁹ As such, the maximum penalties for hate propaganda offences would be increased to life imprisonment from the previous five-year maximum under the current Code.¹²⁰ Setting a new legal foundation for hate-driven actions, these are not the only legal tools the Bill implements to prevent online harms.

The Bill will also create a new peace bond to be issued if there are grounds to believe an individual will commit either hate propaganda or a hate crime offence in the future.¹²¹ Under Section 810.2 of the Code, a peace bond is used as a preventive measure to protect individuals or the public when someone is believed to be at risk of committing a criminal offence (e.g., violent acts or threats).¹²² The person subject to the peace bond is required to keep the peace, be of good behavior, and may be required to abide by specific conditions.¹²³ Under current provisions, the maximum duration for a peace bond is twelve months, with further

court-hate-crimes-1.5604912 (stating that hate crimes are very underreported, and only about a third of hate crimes that are reported are solved).

119. See Simard-Zakaib et al., *supra* note 87 (outlining requirements for social media operators and distribution services).

120. See DEP'T. OF JUS. (Can.), *supra* note 4 ("The Bill would amend the *Criminal Code* to increase the maximum punishments for the four hate propaganda offences in sections 318 and 319 of the *Criminal Code*. It would raise the maximum sentence for the offence of advocating or promoting genocide against an identifiable group in section 318, which is an indictable offence, from five years to a maximum sentence of imprisonment for life. It would raise the maximum sentence for the hate propaganda offences in section 319, when prosecuted as indictable offences, from two to not more than five years of imprisonment. There would be no change to the maximum punishments imposed for these offences when prosecuted as summary conviction offences. The offences in section 319 are inciting hatred against an identifiable group in a public place where it is likely to lead to a breach of the peace; wilfully promoting hatred against an identifiable group; and wilfully promoting antisemitism by condoning, denying or downplaying the Holocaust.").

121. *Id.* (detailing how the Bill would allow a judge to impose an order requiring a person to keep the peace and be of good behavior if the judge is satisfied by the evidence that there are reasonable grounds to fear that the person will commit a hate propaganda offence).

122. See Canada Criminal Code, *supra* note 107, § 810.2.

123. *Id.* ("If the provincial court judge before whom the parties appear is satisfied by the evidence adduced that the informant has reasonable grounds for the fear, the judge may order that the defendant enter into a recognizance to keep the peace and be of good behaviour for a period of not more than 12 months.").

extensions if necessary.¹²⁴ Modeled after the existing peace bond provisions, the new provision would allow a judge to impose an order that requires a person to keep the peace and be of good behavior if there are reasonable grounds to believe they will specifically commit hate propaganda or hate crime offences.¹²⁵ The peace bond could be imposed for up to one year, or two years if the defendant has a prior conviction for a hate crime offence, and could include reasonable and necessary conditions such as prohibitions on communication, location restrictions, curfews, and drug or alcohol prohibitions.¹²⁶ If the conditions of the peace bond are violated, a defendant could face imprisonment for up to four years.¹²⁷ The Bill also gives youth courts exclusive jurisdiction to impose peace bonds on young persons.¹²⁸ As such, the new provision would include a hearing process and Attorney General consent before the order, allowing a process for appealing the order as well.¹²⁹ Mimicking the current provisions, the new provision extends the applicability of peace bonds, broadening them to include hate-motivated offences.

C. Amendments to Civil Statutes

Part III of Bill C-63 would amend the CHRA and reinstate the “communication of hate speech”¹³⁰ as a discriminatory practice. The current CHRA does not address hate speech in the same way

124. *Id.*

125. See DEP’T. OF JUS. (Can.), *supra* note 4 (“This provision would be modelled on existing peace bond provisions in the *Criminal Code*.”).

126. *Id.* (detailing how the new authority provides a judge that may impose any reasonable condition as part of the peace bond in order to secure the defendant’s good behaviour and prevent any offences from happening).

127. *Id.* (referencing Section 811 breach of recognizance offence, which if prosecuted can impose a punishment of imprisonment for up to four years).

128. *Id.* (amending the Youth Criminal Justice Act to give youth justice courts the exclusive jurisdiction to impose a new peace bond, which is a jurisdiction that youth courts already have for other peace bonds).

129. *Id.* (“A person subject to an order could appeal the order and could at any time seek to vary the conditions. In addition, applying for the peace bond would require the consent of the Attorney General[.]”).

130. Bill C-63, *supra* note 64, § 13(1).

that Bill C-63 proposes to.¹³¹ The existing legal framework deals with discrimination based on prohibited grounds outlined in the Charter, as discussed in detail later.¹³² However, hate speech is not directly framed as a discriminatory practice under the current CHRA, relying instead on the Criminal Code provisions for addressing harmful expression.¹³³ Following this, Bill C-63 would introduce a clear prohibition on hate speech under the CHRA, defining it as communications that express detestation or vilification, or those that are likely to incite hatred or violence, against an individual or group based on protected grounds (e.g., race or religion).¹³⁴ Under these amendments, the new prohibition specifically targets online hate speech or speech communicated through telecommunications by the user themselves, rather than the operators of the platforms.¹³⁵

The amendments will provide individuals with the right to bring a related complaint before the Canadian Human Rights Commission (the "Commission") and authorize the Commission to assess penalties.¹³⁶ Under Bill C-63, higher penalties for aggravated offenders are introduced, including up to CAD \$20,000 in compensation to victims.¹³⁷ Likewise, if public bodies (e.g., federal departments, federal agencies, or public institutions) fail to adhere to the rules set out under the new amendments in the

131. See Bill C-63, *supra* note 64, at iii (showing that one of the aims of Bill C-63 is to amend the CHRA to add hate speech as a discriminatory practice); see also Canadian Human Rights Act, R.S.C. 1985, c. H-6 (Can.) [hereinafter Canadian Human Rights Act] (showing that the CHRA does not include a provision addressing hate speech as a discriminatory practice).

132. Canadian Human Rights Act, *supra* note 131, § 3(1).

133. Hate speech is not framed as a discriminatory practice under the Canadian Human Rights Act. See generally *id.*

134. See Bill C-63, *supra* note 64, § 13(1), *modifying* Canadian Human Rights Act, *supra* note 131, § 13.

135. See DEP'T. OF JUS. (Can.), *supra* note 4 ("Part 3 of the Bill would amend the *Canadian Human Rights Act* to complement the regulation of social media services by allowing recourse to the Canadian Human Rights commission against individual users who post hate speech on those services and elsewhere online.").

136. *Id.* ("The Bill would empower the Canadian Human Rights Commission to accept complaints alleging this discriminatory practice and the Canadian Human Rights Tribunal to adjudicate complaints and order remedies.").

137. *Id.* ("In general, the available remedies would [include]...an order to pay compensation of not more than \$20,000 to any victim personally identified in the communication that constituted the discriminatory practice...").

CHRA, they could be subject to a fine of up to CAD \$50,000.¹³⁸ Notably, while aggravating factors could influence the amount of compensation, the current provisions do not as strongly promote such punitive measures.¹³⁹ Further, the Bill grants the Commission and the Canadian Human Rights Tribunal (“CHRT”) the power to issue confidentiality orders to protect complainants, witnesses, or victims from retaliation.¹⁴⁰ This measure is not in the existing CHRA, which currently has no specific confidentiality protections for those involved in hate speech complaints.¹⁴¹ As such, the amendments try to increase accountability for repeat offenders and serious discriminatory practices under the CHRA,¹⁴² hoping to better encourage complainants to come forward.

Lastly, Part IV of Bill C-63 would expand the types of internet services regulated under An Act Respecting the Mandatory Reporting of Internet Child Pornography by Persons who Provide an Internet Service to include online platforms and apps, increase evidentiary preservation from the previous 21 days to one year, and require the gathering of transmission data.¹⁴³ Currently, the Act mainly applies to traditional internet service providers (“ISPs”), rather than a broader range of online platforms.¹⁴⁴ Many

138. See Bill C-63, *supra* note 64, § 121(2).

139. See DEPT. OF JUS. (Can.), *supra* note 4 (“...Part 2 of the Bill would amend the *Criminal Code* to create a new peace bond to help prevent and deter hate crimes and hate propaganda offences from being committed, increase the maximum punishments for hate propaganda offences, and create a new general hate crime offence.”).

140. Bill C-63, *supra* note 64, § 36(2), *modifying* Canadian Human Rights Act, *supra* note 131, § 36(2).

141. See *generally* Canadian Human Rights Act, *supra* note 131 (containing the current Canadian Human Rights Act and its protections, which do not include any specific confidentiality protections).

142. Bill C-63, *supra* note 64, § 41.

143. Cf. Bill C-63, *supra* note 64, § 45-46 (“Internet service includes a service (a) providing Internet access; (b) providing Internet content hosting, regardless of the originator of the content or the manner by which the content is made accessible; or (c) facilitating interpersonal communication over the Internet, including a service providing electronic mail.”) with S. 2011, c. 4, § 1 (“Internet Service means a service providing Internet access, Internet content hosting or electronic mail.”).

144. Cf. Bill C-63, *supra* note 64, § 45-46, with S. 2011, c. 4, § 1 (the current definition of Internet Service only includes services providing Internet access, Internet content hosting, or electronic mail, while the amended definition would also include services facilitating interpersonal communications over the Internet).

widely-used services, such as social media platforms, messaging apps, and other digital services, are outside the scope of regulation.¹⁴⁵ Resultingly, harmful content could circulate on these platforms without necessarily being flagged under the Act. As such, Bill C-63 would expand coverage to these non-traditional forms of media. Further, the Act currently does not require service providers to gather and retain transmission data for reports, only mandating reports when “reasonable grounds” of child pornography exist.¹⁴⁶ Under the new amendments, ISPs would be mandated to include transmission data (e.g., telephone number, internet protocol address, port number, date, and time) in their reports when they encounter what is “manifestly child pornography”.¹⁴⁷ The Bill does clarify that ISPs are obligated only to report data when they encounter child pornography, not to search for it.¹⁴⁸ Failure to comply could result in civil penalties, including monetary fines.¹⁴⁹ However, no criminal charges are proposed for non-compliance, with administrative and civil remedies serving as the only compliance mechanisms.¹⁵⁰

In summary, Bill C-63 aims to cultivate a safer online environment by addressing harmful content through a multifaceted approach. The Bill introduces a new regulatory framework overseen by the DSCC and a digital safety ombudsman to address user complaints, ensure the removal of harmful content, and develop safety standards. Primarily focusing on large-scale social media services, the Bill imposes four key duties on their operators. In addition to regulating social media services, Bill C-63 amends the Code to include a broader definition of hatred that encompasses online and offline hate speech, introduces a new hate crime offence, and increases

145. See DEP'T. OF JUS. (CAN.), *supra* note 4 (“The Act also would not apply to social media services that do not enable users to communicate content to the public and would not require operators to address any harmful content that may be disseminated via any private messaging features of their regulated services.”).

146. *Id.*

147. *Id.*

148. *Id.*

149. *Id.*

150. *Id.* (explaining that the amended penalties for violators would only include administrative and civil penalties, and would be intended to promote compliance with the Act rather than inflicting punishment).

penalties for hate propaganda offences. Likewise, the CHRA is also amended to address hate speech as a discriminatory practice, empowering the Commission to impose penalties and ensuring confidentiality protections. Lastly, Bill C-63 expands the scope of An Act Respecting the Mandatory Reporting of Internet Child Pornography by Persons who Provide an Internet Service to encompass online platforms and apps, extending the evidentiary preservation period, and mandating the collection of transmission data. Importantly, these measures aim to balance online safety with individual rights and freedoms, hoping to create a more accountable and secure online landscape in Canada.

IV. THE CANADIAN CHARTER OF RIGHTS AND FREEDOMS AND SECTION 2(B)

As part of the Constitution of Canada, the Charter guarantees certain legal, political, and equality rights to all Canadian citizens.¹⁵¹ Enacted in 1982, the Charter is a cornerstone of Canadian law and society, safeguarding fundamental freedoms such as thought, belief, opinion, and expression.¹⁵² As Walker writes, “[a]ll laws in Canada must comply with the Charter and are interpreted by Canadian courts in a way that is consistent with the supremacy of the Canadian Constitution.”¹⁵³ That said, these rights are not absolute and can be subject to “reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society.”¹⁵⁴ The Charter serves as a framework for balancing individual rights with societal interests, ensuring that laws are consistent with Canada’s commitment to fundamental freedoms and democratic values.¹⁵⁵ As such, there persists a complex interplay between legislative efforts and the Charter when addressing broader social issues. The Charter has shaped

151. See Charter of Rights, *supra* note 5, §§ 3, 7-15.

152. *Id.* § 2(b).

153. Julian Walker, *Hate Speech and Freedom of Expression: Legal Boundaries in Canada*, No. 2018-25-E LIBRARY OF PARLIAMENT 1, 2 (2018), <https://lop.parl.ca/staticfiles/PublicWebsite/Home/ResearchPublications/BackgroundPapers/PDF/2018-25-E.pdf>.

154. Charter of Rights, *supra* note 5, § 1.

155. Walker, *supra* note 153, at 3 (“[C]ourts must use a balancing test to weigh the objectives and actions of the government or other institution against the interests of an individual claiming that a Charter right has been violated.”).

both legal debates and decisions in Canada for years. For example, in *Saskatchewan (Human Rights Commission) v. Whatcott*,¹⁵⁶ the Supreme Court of Canada ("SCC") provided the definition of hatred for Charter purposes. As such, the courts and the Charter will heavily influence the drafts and future enactment of Bill C-63.

The Charter protects several freedoms that may be engaged by virtue of Bill C-63. This includes section 7 (protection of an individual's life, liberty, and security), section 8 (protection against unreasonable search or seizure), section 12 (protection against any cruel and unusual treatment or punishment), and section 15(1) (the right of equality).¹⁵⁷ Importantly, the focus of this Article centers around section 2(b), the right to the freedom of expression.¹⁵⁸ Thus, while imperative to note the additional freedoms engaged by Bill C-63, they will not be discussed in further detail for the purpose of this Article. Under section 2(b) of the Charter, every individual is guaranteed the right to freedom of thought, belief, opinion, and expression; this includes the freedom of the press and other media of communication.¹⁵⁹ Section 2(b) has been broadly interpreted to encompass any activity or communication, excluding violence or threats of violence, that conveys or attempts to convey meaning.¹⁶⁰ The freedom from being compelled to express oneself is also protected.¹⁶¹ Moreover, section 2(b) also encompasses the "open court principle"¹⁶² whereby the public has the right to access information concerning the court and any adjudicative proceedings. As such, this Charter right is afforded to all Canadian citizens.¹⁶³

Freedom of speech has also been declared a fundamental freedom and human right in the Canadian Bill of Rights.¹⁶⁴

156. *Saskatchewan (H.R.C.) v. Whatcott*, [2013] 1 S.C.R. 467 (Can.).

157. Charter of Rights, *supra* note 5, §§ 7, 8, 12, 15(1).

158. *Id.* § 2(b).

159. *Id.*

160. DEPT. OF JUST. (Can.), *supra* note 4.

161. *Id.*

162. *Id.*

163. Charter of Rights, *supra* note 5, § 2.

164. Canadian Bill of Rights, S.C. 1960, c. 44 §1(d).

Various rights are set out in this federal law, including the freedom of religion and freedom of the press.¹⁶⁵ Though not technically part of Canada's Constitution, the SCC has described it as quasi-constitutional and therefore mandated that other laws be interpreted consistently.¹⁶⁶ Importantly, the SCC has recognized that the freedom of expression per section 2(b) is not absolute. Restrictions on it have been upheld when they run contrary to the spirit of the Charter (e.g., hate speech, which works to restrict the rights of others).¹⁶⁷ In general, certain limitations can be placed on Charter guarantees. Specifically, section 1 of the Charter itself recognizes that all rights and freedoms guaranteed under it are subject to "reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society."¹⁶⁸ This means that once an infringement has been established, the courts are to decide whether the violation by the government or respective institution to which the Charter applies can be considered justified.¹⁶⁹ In doing so, courts use a balancing test that weighs the objectives and actions of the government or relevant body against the interests of the individual claiming the violation of their Charter right.¹⁷⁰ Per section 52 of the Constitution Act, the law (or part of the law) found to be unconstitutional must be struck down.¹⁷¹ In the

165. *Id.* §§ 1(c), 1(f).

166. *Hogan v. The Queen*, [1975] 2 S.C.R. 574, 583 (Can.) (describing the Canadian Bill of Rights as a quasi-constitutional instrument whose provisions must be construed and applied as if they were constitutional provisions).

167. *Saskatchewan*, [2013] 1 S.C.R. 521 (Can.) (holding that hate speech lies at some distance from the spirit of § 2(b) because it undermines rather than promotes the values of free expression and can silence its targets).

168. *Id.* at 557.

169. Where Charter rights have been infringed on an individual or group, the party who has had their rights infringed may apply for a remedy under § 24(1) of the Charter. This section permits courts to award individualized forms of remedy that are considered "appropriate and just" in the circumstances. Contrastingly, human rights tribunals are limited to making orders that are provided for in their governing legislation, despite their general broad remedial powers. Walker, *supra* note 153, at 3 n.21.

170. *R. v. Keegstra*, [1990] 3 S.C.R. 697, 735 (Can.) (stating that under § 1 judges must engage in a balancing task between the limitation of a fundamental right and the state's conflicting objective, which must be sufficiently important to justify the restriction).

171. The courts have given the government time to rectify any part of a law that is unconstitutional in some cases prior to striking down the law. See *Att'y Gen. of Can. v. Bedford*, [2013] 3 S.C.R. 1101, 1122 (Can.) (demonstrating the Court's ability to rectify

alternative, the law may be found constitutional, justifiably limiting an individual's Charter right.

Numerous laws have placed restrictions on the freedom of expression as part of the law's intended purpose or as an indirect effect. Examples include perjury cases, creating child pornography, publishing election survey results on a federal election day with polls still open, and counselling suicide. All of these have been limited through the designation as criminal offences in the Criminal Code.¹⁷² Further, the promotion of hatred is prohibited in publications through both provincial human rights law and the Criminal Code. The hate promotion offences can be found in sections 318 to 320.1 of the Code. Notably, prosecutions under these sections have been limited, with very little jurisprudence on them.¹⁷³ Nevertheless, the court decisions that do exist remain key for the judicial interpretation of section 2(b) of the Charter.¹⁷⁴ In general, such examples highlight that freedom of expression can be limited in situations where promoting other goals or values is considered of greater social importance.¹⁷⁵ Thus, in understanding the basic tenets of sections 1 and 2(b), the discussion will move towards how Bill C-63 tries to balance protecting these fundamental rights while also safeguarding vulnerable populations from online harms.

V. WAYS BILL C-63 MAY ENGAGE SECTION 2(B) OF THE CHARTER

"What does Justin Trudeau mean when he says the words 'hate speech'? He means the speech he hates. You can assume he will ban all of that."¹⁷⁶ Conservative Leader Pierre Poilievre has

certain provisions of the Criminal Code that are declared unconstitutional through severance of the statutory language).

172. Canada Criminal Code, *supra* note 107, §§ 132, 163.1 (2), 241 (1).

173. Walker, *supra* note 153, at 3—7.

174. Several proposals to amend the hate propaganda provisions have occurred since their inception, however, Parliament rarely has changed them. One instance was when the Anti-Terrorism Act, S.C. 2001, c. 41 was adopted, amending whose section § 10 of the Criminal Code by adding §320.1, which dealt with hate material found on computers and the seizure of such material. Walker, *supra* note 153, at 3 n.24.

175. Walker, *supra* note 153, at 2—3.

176. John Paul Tasker, *Asked About Online Harms Bill, Poilievre Raises Trudeau's Past Use of Blackface*, CBC NEWS (Feb. 21, 2024, 10:44 AM), <https://www.cbc.ca/news/>

proclaimed that his party vehemently opposes the forthcoming legislation.¹⁷⁷ Many opponents of Bill C-63 acknowledge the good intentions behind the provisions. Specifically, the Canadian Civil Liberties Association (“CCLA”) endorses the declared purpose to uphold public safety, protect children, and support marginalized communities.¹⁷⁸ However, many underscore that the Bill needs further rectifications irrespective of the purpose.¹⁷⁹ Largely, the concerns stem from the potential engagement of the Bill’s provisions with section 2(b) of the Charter.¹⁸⁰ As such, the engagement and justification stances are imperative to discuss.

A. Mandatory Content Moderation and Removal

Bill C-63 imposes several duties and obligations on operators. One of them is the duty to act responsibly, compelling operators to mitigate harmful content exposure to users.¹⁸¹ This duty requires implementing mitigation measures such as content identification, user blocking, flagging options, and content removal.¹⁸² The federal government maintains that the duty to act responsibly does not disproportionately limit the freedom of expression.¹⁸³ For example, while the Bill mandates measures to mitigate the risk of users being exposed to harmful content, this obligation provides a non-exhaustive, non-mandatory list;¹⁸⁴ operators can determine the most effective ways to mitigate harm

politics/online-harms-bill-poilievre-trudeau-blackface-1.7120979.

177. *Id.*

178. CANADIAN CIV. LIBERTY ASSOC., *CCLA Urges Substantial Amendments to the Online Harms Act* (Feb. 28, 2024), <https://ccla.org/press-release/ccla-urges-substantial-amendments-to-the-online-harms-act/> (noting that the CCLA endorses the declared purposes of upholding public safety, protecting children, and supporting marginalized communities).

179. *Id.* (warning that, despite endorsing the Bill’s purposes, it contains overbroad violations of expressive freedom, privacy, protest rights, and liberty that must be rectified).

180. DEPT. OF JUST. (Can.), *supra* note 4.

181. *See* Bill C-63, *supra* note 64, § 54.

182. *See id.* § 62(1)(a)(iii); *see also* Lau, *supra* note 77 (discussing the methods social media operators may employ to act in accordance with Bill C-63).

183. *See* DEPT. OF JUST. (Can.), *supra* note 4 (asserting that the Act is principled upon the notion of free expression).

184. *See id.* (noting that the Act would enable operators to determine which mitigation measures best suit their service platform).

on their services independently. Further, the Bill specifies that unreasonable measures that disproportionately limit a user's expression are not required.¹⁸⁵ As such, platforms can tailor the measures to fit their specific needs. Bill C-63 drafters also emphasize that the focus is on content outside the core values of the freedom of expression (e.g., content inciting violence, terrorism, and sexual abuse). With this, the Bill imposes no direct consequences on the user through these duties.¹⁸⁶ Thus, the federal government contends to recognize the importance of preserving free expression while acknowledging the need for platforms to take steps to create a safer online environment through backstop measures targeting specific categories.

Bill C-63 also imposes a duty upon operators to make certain content inaccessible, such as content that sexually victimizes a child or revictimizes a survivor, and intimate content communicated without consent.¹⁸⁷ Operators are required to act within 24 hours of becoming aware of the content to remove and make it inaccessible.¹⁸⁸ The exception to this is if the flagged content by the user is trivial, frivolous, vexatious, or made in bad faith.¹⁸⁹ As such, the federal government asserts that content caught under this duty is severely harmful and outside the scope of the freedom of expression.¹⁹⁰ Specifically, this duty only applies to three categories of content (that is, sexualizing a child, revictimizing a survivor, and intimate content communicated without consent), all of which inflict serious psychological and physical harm that undermine the principles of free expression,¹⁹¹

185. *Id.* ("...[T]he Act provides that operators are not required to implement measures that unreasonably or disproportionately limit users' expression on their services.").

186. *Id.*

187. Bill C-63, *supra* note 64, § 67(1)(a).

188. See DEP'T. OF JUST. (Can.), *supra* note 4 (asserting that operators are required to make flagged content inaccessible within 24 hours of flagging unless the flag is "trivial, frivolous, vexatious or made in bad faith, or duplicative of another flag.").

189. *Id.* ("When content is flagged, operators would need to make it inaccessible on a preliminary basis within 24 hours (or a different period that could be set out in regulation) unless the flag is trivial, frivolous, vexatious or made in bad faith, or duplicative of another flag.").

190. *Id.*

191. See *id.* (Content that includes the sexual victimization of children or survivors of sexual misconduct lies from the core guarantees of free expression. This content will receive less protection under the Charter).

Beyond the direct harm to victims, these content types actively discourage online participation. The existence and proliferation of such content creates a hostile digital environment where victims and at-risk individuals feel unsafe engaging in online spaces. Victims and vulnerable populations who face sexual exploitation, non-consensual image sharing, or revictimization through digital platforms are effectively silenced and deterred from engaging in online spaces.¹⁹² As such, the government argues that restricting these categories of content is necessary to protect the most vulnerable groups in society.¹⁹³ Notably, the content encompassed by this duty is criminalized under the Code already.¹⁹⁴ The harms caused by these types of content have been recognized to cause individuals damage to their privacy, mental health, and dignity.¹⁹⁵ Proponents contend that safeguarding victims of harmful content outweighs any infringement per section 2(b).¹⁹⁶ Further, the Bill's opportunity for appeal involves considering the intent and context more deeply.¹⁹⁷ Thus, mechanisms such as limited proactive searches and independent judicial oversight help prevent the arbitrary removal of content.

Importantly, critics of the mandatory content moderation and removal provisions raise concerns of stifling free speech. As mentioned, the duty to act responsibly compels operators to mitigate harmful content exposure to users.¹⁹⁸ While Bill C-63 does not explicitly mandate specific mitigation measures,¹⁹⁹ the requirement to act against harmful content could reasonably suppress legitimate expression. The Canadian Minister of Justice has recognized that fulfilling this duty impacts whether content may appear online, and thus, inherently engages the freedom

192. *Id.*

193. *See id.* (Expression that degrades or dehumanizes is harmful in of itself. The restriction imposed is meant to prohibit these identified harms while leaving out low risk content).

194. *See* DEPT. OF JUST. (Can.), *supra* note 4 (asserting that the content Bill C-63 purports to restrict is currently prohibited under the Criminal Code).

195. *Id.*

196. DEPT. OF JUST. (Can.), *supra* note 4.

197. *Id.* (discussing elements of procedural fairness within the Act).

198. *See* Bill C-63, *supra* note 64, § 54.

199. *See* DEPT. OF JUST. (Can.), *supra* note 4 (discussing that Bill C-63 would allow operators of regulated services to have flexibility in determining how to mitigate risks).

of expression.²⁰⁰ Through this, the removal of user-generated content that is merely offensive, but not illegal, may be at risk for removal (i.e., the threshold for *inducing* hatred or violence can be overbroad).²⁰¹ For example, content that criticizes government policies or engagement in political debate has been subject to over-censoring and denial of social media services, with *Facebook* requiring individuals to clearly “indicate their intent” to avoid removal.²⁰² Further, this duty provides the DSCC discretion to determine what constitutes harmful content.²⁰³ Moreover, the DSCC has the authority to order an operator to cease providing social media services wholly in certain cases.²⁰⁴ These powers can be considered quite extensive.

Likewise, the duty to make certain content inaccessible raises concerns about potential overreach and the stifling of the freedom of expression. Similarly, harmful content that would still gain Charter protection may fall within the Bill’s scope not by purpose, but by effect. While there is an opportunity to appeal the final decision prior to the content being permanently inaccessible, this occurs after the initial takedown, and the length of the appeal process is unknown.²⁰⁵ Wide-reaching discretionary powers by the DSCC, coupled with vague terminology²⁰⁶, could incentivize the silencing of valuable speech and lead to unjust censorship. In

200. *Id.* (discussing how Bill C-63’s duty for operators to act responsibly may impact whether specified content will appear on a service, which may infringe the freedom of expression of its users).

201. *See* DEPT. OF JUST. (CAN.), *supra* note 4 (discussing the labelling of bot-disseminated harmful content generally).

202. *See* ANGEL DIAZ & LAURA HECHT-FELELLA, *DOUBLE STANDARDS IN SOCIAL MEDIA CONTENT MODERATION*, 5 (Brennan Center for Justice, 2021) (discussing that over reliance on government prohibitions has caused a wide range of political speech and human rights documentation will be caught in a “removal dragnet.”).

203. *See* DEPT. OF JUST. (CAN.), *supra* note 4.

204. *See* David Fraser, *What’s the Harm? 10 key facts about Canada’s proposed Online Harms Act*, MCINNES COOPER (Apr. 30, 2024), <https://www.mcinnescoper.com/publications/whats-the-harm-10-key-facts-about-canadas-proposed-online-harms-act/> (asserting broadly that the Commission has the authority to require an operator to in any measures that will ensure compliance with the Act).

205. *See id.* (explaining that content deemed to be non-compliant with the terms of the Act must be removed immediately by the operator. The operator has an opportunity make representations to the Commission regarding the content at issue only after the content has been removed).

206. *Id.* (discussing the discretionary powers held by the Commission in determining when content removal is appropriate).

particular, social networking platforms may fear penalties and be overly restrictive to avoid violating the Bill. Moreover, past content moderation policies have been shown to over-enforce the viewpoints of marginalized groups such as communities of color, religious minorities, and LGBTQ communities.²⁰⁷ For example, Facebook's moderation of "dangerous individuals and organizations"²⁰⁸ to target terrorist groups has led to Muslim content being disproportionately misinterpreted as promoting and inciting violence, extremism, and terroristic views.²⁰⁹ In such cases, content produced by marginalized groups is traded off for rapid and wide-sweeping enforcement mechanisms that may be unable to account for context and intent.

B. Compelled Speech

Under Bill C-63, the DSCC can order an operator to publish a notice detailing a violation of the Bill or an undertaking entered with the DSCC.²¹⁰ Operators are also required to label content they reasonably believe is disseminated by bots and in which the content garnered prominence through such automated amplification.²¹¹ Proponents of these provisions underscore their consistency with the Charter to uphold principles of transparency and accountability.²¹² Largely, the publication of a notice or an undertaking remains discretionary; this matter is not left to the authority of the DSCC alone but rather under the guidance of the legislation to weigh the platform's freedom of expression against the objectives of the Bill.²¹³ Likewise, the government's objective

207. See DIAZ & HECHT-FELELLA, *supra* note 202, at 10—11 (discussing the minorities groups that have been disparately impacted and censored by content moderation policies for sharing politically charged media).

208. See *id.* at 5.

209. See *id.* (asserting that Canada's reliance on content sanction lists has led to an increased removal of political content posted by Muslim communities).

210. See Bill C-63, *supra* note 64, § 59(2).

211. See *id.* § 60.

212. See DEP'T. OF JUST. (CAN.), *supra* note 4 (explaining that operators are required to remain transparent with users regarding their duties under the Act to facilitate operator accountability).

213. See *id.* ("In exercising its discretion, the Commission would have to take into account any freedom of expression of an operator that may be affected and weigh it against

to label harmful content disseminated by bots is to occur only if there are reasonable grounds to believe it is prominent due to the automated amplification by bots.²¹⁴ In both cases, the labelling requirement is limited to details factual in nature, strictly pertaining to either the statutory compliance failure or informing the user of bot activity.²¹⁵ As such, the government contends that these labelling requirements constitute a minimal intrusion on the freedom of expression and are, therefore, justifiable.²¹⁶ Undeniably, these methods inform users and place them in a position to make informed decisions on validity with more background information. Simultaneously, social media platforms may be held accountable for their duties to users.

Ordering an operator to publish a notice with information that communicates a particular message can raise constitutional concerns. Potentially constituting compelled speech, section 2(b) of the Charter may be engaged as the operator is forced to communicate information they may not wish to express, or otherwise choose to publish. While the publishing requirement is not automatic when content is flagged, the DSCC has the discretion to determine when a publication order is made.²¹⁷ As such, if the order is made, the operator remains compelled to publicly and actively promote the DSCC's view; thus, this duty can engage compelled speech. Moreover, this is particularly relevant if the operator believes the DSCC's decision is flawed or misrepresents their moderation efforts. Likewise, despite intending to inform users and mitigate the manipulative effects of bots, requiring content labels can be similarly construed as compelled speech. Specifically, social media platforms may be forced to endorse a particular interpretation of content, violating

the purposes of the *Online Harms Act* in order to make a reasonable and proportionate decision.”).

214. *Id.* (explaining that there must be reasonable grounds to believe bot activity both occurred and enhanced the content's prominence).

215. *Id.* (“The only information that could be required to be published is factual in nature and relates to an operator's non-compliance with statutory requirements.”).

216. *Id.* (detailing that one of the Act's purposes is to ensure operator transparency and accountability).

217. Bill C-63, *supra* note 64, § 119.

their right to express themselves freely.²¹⁸ For instance, consider a bot network amplifying contentious political discussions. The platform may be obligated to label the content as bot-disseminated harmful content, and as a result, the operator applies the label for users to see. The public could construe the platform as taking a stance in the political discussion based on the content label. More concerning, legitimate online activism may be stifled by those who implement bots for amplification (e.g., environmental organizations employing bots to raise awareness about climate change), devaluing the messaging for users.²¹⁹ Alternatively, this could be implemented as a tactic for targeting the removal of content that a party may want to silence. As such, the subjective and vague interpretation of applying notice orders and content labels could lead to arbitrary decisions that unjustifiably engage section 2(b) of the Charter.

C. Hate Speech

Bill C-63 defines hatred as “the emotion that involves detestation or vilification and that is stronger than disdain or dislike”.²²⁰ As such, the government contends this is only to target the most extreme forms of expression.²²¹ Following this logic, most political speech and social discourse should be “untouched”.²²² Further, the definition in the Bill is consistent with the SCC’s jurisprudence on hate speech generally. In both the cases of *R. v. Keegstra*²²³ and *Canada (Human Rights Commission) v. Taylor*,²²⁴ hatred was determined to denote extreme manifestations of detestation or vilification that go

218. See *Slaight Commc’ns. Inc. v. Davidson* [1989] 1 S.C.R. 1038 (Can.) (explaining that compelled speech is a violation of freedom of expression).

219. Bill C-63, *supra* note 64, § 60.

220. *Id.* § 319(7).

221. See Clayton DeMain, *Feds Claim Online Harms Act Won’t Infringe on Charter Rights*, TRUE NORTH (June 7, 2024), <https://tnc.news/2024/06/07/feds-online-harms-wont-infringe-Charter1/> (describing what the federal government contends is the target of Bill C-63).

222. DEPT. OF JUST. (Can.), *supra* note 4 (explaining that most speech will be unaffected).

223. [1990] 3 S.C.R. 697 (Can.).

224. [1990] 3 S.C.R. 892 (Can.).

beyond mere dislike or offence. By purportedly using this definition as a guiding standard in the Bill, the government appears to provide support towards a more identifiable legal threshold for distinguishing hate speech by mimicking already precedent-setting case law. Likewise, the reintroduction of a hate speech provision in the CHRA can help provide individuals with recourse against those who post hate speech. Given this, the federal government claims that to ensure equal opportunity without being hindered by discrimination, these provisions remain necessary.²²⁵

Importantly, the definition of hatred within Bill C-63 is still hotly contested concerning its scope and application despite the tie to case law. To constitute “hatred”²²⁶ for hate propaganda under Bill C-63, critics argue that despite the government’s claims of clarity, the definition and its application remain largely left to interpretation in practice; in particular, that while the definition may tie back to existing jurisprudence, opponents contend that applying case law principles to a statutory definition introduces ambiguity and risks overly broad enforcement.²²⁷ The risk noted here is encompassing expression that may fall short of what constitutes hate propaganda. Critics posit that legitimate expression of criticism or dissent will be silenced, particularly from marginalized groups that tend to use strong language when expressing lived experiences and challenging current power dynamics.²²⁸ Historically, anti-hate laws have been disproportionately used against members of marginalized communities, even when their speech does not meet the legal threshold for hate speech.²²⁹ Thus, despite the government underscoring this provision’s support from case law, it still may remain too vague for broad application.

Bill C-63 additionally outlines amendments to the CHRA to address online hate speech by making it a discriminatory practice to communicate hate speech in a way *likely* to encourage hatred or

225. DEPT. OF JUST. (Can.), *supra* note 4.

226. See Bill C-63, *supra* note 64, § 319(7) (defining hatred).

227. See THE DEMOCRACY FUND, *Bill C-63: What You Need To Know* (Oct. 10, 2024), https://www.thedemocracyfund.ca/online_harms_bill_brief.

228. See Ahmad, *What’s in Bill C-63*, *supra* note 5.

229. DIAZ & HECHT-FELLELLA, *supra* note 202, at 3, 7.

severe dislike of an individual or a group.²³⁰ The Commission is empowered to receive and adjudicate complaints related to online hate speech.²³¹ Specifically, the Commission would be granted the authority to prosecute and penalize Canadians for expressing offensive opinions that are not criminal but deemed hateful by unelected officials under the implementation of the Bill.²³² Here, a citizen may file an anonymous complaint against another for technically non-criminal speech.²³³ These amendments could be subject to weaponization, silencing critics of the government or dominant social groups. For example, consider if a member of a racialized group posts a critique of immigration policies to express frustration with systemic discrimination, using strong language. Wholly intending to raise awareness for injustice and advocate for change, a disagreeing individual might file a complaint alleging hate speech. Even if the Commission decides to dismiss the complaint, the process may be time-consuming and burdensome.²³⁴ While aiming to protect certain groups, this provision has the risk of being self-defeating; overall, it could exacerbate self-censorship and prevent engagement in robust discourse.

D. The Open Court Principle and Issues of Confidentiality

Bill C-63 grants powers to the CHRC with several provisions aiming to protect confidential information (e.g., trade secrets, financial information, commercial data, the identities of employees who submit to the Commission, and principles of confidentiality generally).²³⁵ This includes amendments that empower the Commission and the CHRT to issue confidentiality orders on complainants, witnesses, and alleged victims.²³⁶ The Bill also permits the Commission to hold private hearings, even

230. Bill C-63, *supra* note 64, § 319(7).

231. Bill C-63, *supra* note 64, § 88.

232. *Id.* § 119(1).

233. See JUST. CTR. FOR CONST. FREEDOMS, *Online Harms Act Threatens Free Expression in Canada*, (Feb. 29, 2024), <https://www.jccf.ca/online-harms-act-threatens-free-expression-in-canada> (explaining that citizens may report each other based on fear).

234. Ahmad, *What's in Bill C-63*, *supra* note 5.

235. See Bill C-63, *supra* note 64, § 88(2).

236. *Id.* § 127.

where the default is public.²³⁷ The confidentiality provisions are drafted with the intent to protect witnesses, alleged victims, and hate speech complainants from intimidation, threats, and discrimination.²³⁸ The federal government posits these measures are necessary to ensure members involved in such matters feel safe and secure to come forward with complaints, preventing further victimization.²³⁹ Moreover, the decisions on confidentiality are made on a case-by-case basis by the Commission and the CHRT, rather than on an automatic default basis.²⁴⁰ As such, both these administrative bodies are subject to the provisions of the Charter and must exercise their discretion in a manner that complies with it.²⁴¹ Further, any decisions may be subject to judicial review.²⁴²

The risk of Charter violations under these amendments may be pressing. While the intended purposes of these provisions are to protect individuals from retaliation, protect privacy, and encourage whistle-blowing, some opponents believe they *prima facie* violate the open court principle;²⁴³ this principle guarantees public access to the court and adjudicative proceedings.²⁴⁴ Despite restricting access to safeguard confidential business information to protect commercial interests and sensitive information overall, these components are protected by section 2(b) of the Charter on face value.²⁴⁵ Given the provisions allowing for confidentiality orders and private hearings can cause concern for a section 2(b) Charter violation via the open court principle, the challenge lies in ensuring judicial application of the confidentiality provisions;

237. *Id.* § 88(2).

238. See DEP'T. OF JUST. (Can.), *supra* note 4 (explaining that confidentiality orders are available if necessary).

239. *Id.* (explaining that confidentiality may be necessary to protect victims).

240. Bill C-63, *supra* note 64, § 88(2).

241. *Doré v. Barreau du Québec*, [2012] 1 SCR 395 (Can.).

242. *Id.*

243. See DEP'T. OF JUST. (Can.), *supra* note 4; see also CANADIAN CIV. LIBERTY ASSOC., *supra* note 178.

244. See SUPREME COURT OF CANADA, *The Open Court Principle* (Sep. 7, 2025), <https://www.scc-csc.ca/about-apropos/work-travail/transparent/> (explaining Canada's open court principle).

245. See *Section 2(b) Freedom Of Expression*, GOV'T OF CANADA, (Dec. 1, 2024), <https://www.justice.gc.ca/eng/cs/sj/sjc/rfc-dlc/ccrf-cedl/check/art2b.html>.

this requires balancing transparency and accountability with individual rights protection and legitimate confidentiality interests in mind.

VI. THE RIGHT PATH FORWARD? APPLAUDING THE RIGHT EFFORTS IN THE WRONG DIRECTION

In light of the potential violations of section 2(b), the question remains whether the possibility of these infringements may be justifiable. Overall, it is likely that Bill C-63 will need to undergo further revision and consultation. This is largely based on case precedent, Charter rights, and the extent to which limitations on section 2(b) can be reasonably justified.²⁴⁶ That said, not all provisions within the Bill must be further rectified. Nonetheless, careful consideration must be taken when making such amendments. Moreover, further questions need to be asked concerning the sufficiency of the regulations to address the overarching problem of online harms amplified through social media use.

Aligning with the objectives of Bill C-63, SCC case precedents have consistently upheld laws restricting harmful content to protect vulnerable individuals and public safety. In *R. v. Keegstra*, the Court upheld section 319(2) of the Code, prohibiting hate speech that incites hatred against identifiable groups,²⁴⁷ which supports Bill C-63's regulation of hate-fomenting content. Further, the court in *Saskatchewan (Human Rights Commission) v. Whatcott* upheld the Saskatchewan Human Rights Code, restricting hate speech in publications.²⁴⁸ In *R. v. Zundel*, the SCC considered whether a law prohibiting the publication of false

246. See CANADIAN CIV. LIBERTY ASSOC., *supra* note 178 (expressing concern that Bill C-63 contains overbroad restrictions on expressive freedom, privacy, and protest rights, and urging substantial amendments to address the Bill's sweeping powers before it is passed).

247. See *Keegstra*, [1990] 3 S.C.R. at 780 (Can.) (holding that the prosecution of a high school teacher in Alberta for anti-Semitic statements made during his class was a reasonable and justifiable limitation on freedom of expression).

248. See *Whatcott*, [2013] 1 S.C.R. at 542 (Can.) (upholding § 14 of the Saskatchewan Human Rights Code, and explaining that publications likely to expose vulnerable groups to hatred may be restricted, while still protecting expression that contributes to public debate).

news violated the freedom of expression under section 2(b) of the Charter.²⁴⁹ Striking down the law as overly broad, the court acknowledged that hate speech and material that could harm public order could be limited.²⁵⁰ Thus, these cases demonstrate that while freedom of expression is fundamental, online hate speech or material that misrepresents facts or harms individuals may justify restrictions. Likewise, other criminal offences under the Code have been challenged constitutionally, where the protection of children has been recognized as a compelling reason to limit otherwise broadly prohibited conduct.²⁵¹ These core principles remain reflected in Bill C-63.²⁵²

Following this, Bill C-63's duty to act responsibly and duty to make content inaccessible are likely justifiable under section 2(b) of the Charter. While the concerns regarding these provisions are valid, the drafters intended to strike a balance between upholding the freedom of expression and combating proven harmful content online.²⁵³ Despite limiting specific types of expression, section 2(b) of the Charter is not hindered beyond a justifiable means. Firstly, the Bill centers on specifically known established categories not given protection under the Charter.²⁵⁴ Specifically, the seven content categories are: "intimate content communicated without consent; content that sexually victimizes a child or revictimizes a survivor; content that induces a child to harm themselves; content used to bully a child; content that fomented hatred; content that incites violence; and content that incites violent extremism or terrorism."²⁵⁵ Historically, these categories of content have been accepted as outside the scope of protected expression.²⁵⁶ As

249. See *R. v. Zundel*, [1992] 2 S.C.R. 731, 732 (Can.).

250. *Id.* at 733—34 (striking down the law as overly broad and noting that s. 181 supports criminalization of expression only when sanctions are narrowly confined to serious concerns).

251. See *R. v. Butler*, [1992] 1 S.C.R. 452, 454 (Can.); see also *Canadian Found. for Children, Youth and the Law v. Canada (Attorney General)*, [2004] 1 S.C.R. 76, 77—78 (Can.); see also *R. v. Sharpe*, [2001] S.C.R. 45, 52—53 (Can.).

252. Ahmad, *What's in Bill C-63*, *supra* note 5.

253. See DEP'T. OF JUST. (Can.), *supra* note 4 (explaining that searches or seizures under § 8 must reasonably balance privacy and state interests).

254. See *id.* (discussing categories of content the Bill targets).

255. *Id.*

256. *Id.*

such, there is a strong social interest in protecting children and victims of sexual exploitation, as well as preventing terrorism and acts of violence.²⁵⁷ Further, the freedom of expression is premised upon the “search for and attainment of truth, participation in social and political decision-making, and the opportunity for individual self-fulfillment through expression.”²⁵⁸ By limiting its scope to specific categories, the Bill aligns with previous case law, minimizes the effect on legitimate expression, and fulfills the Charter’s fundamental principles and values.

Similarly, the duties to publish notices or violations and to label bot-disseminated content under Bill C-63 are likely justified under the Charter. While there is no specific case precedent in Canada that directly mandates online companies to label content,²⁵⁹ several regulatory frameworks in similar industries suggest that the government can impose regulations requiring platforms to disclose or label content in certain contexts. For example, the Canadian Broadcasting Act provides the Canadian Radio-Television and Telecommunications Commission (“CRTC”) with the authority to regulate most types of media broadcasts, such as television and radio.²⁶⁰ Under their authority, regulatory labelling already occurs, such as content ratings for television programs (e.g., “C” for children, “G” for general audiences, and “18+” for mature audiences) and identification of paid-for advertisements.²⁶¹ Following this, compelling platforms to publish notices about violations and label bot-generated content

257. *Proposed Bill to Address Online Harms*, GOV'T OF CANADA (May 14, 2025), <https://www.canada.ca/en/canadian-heritage/services/online-harms.html>.

258. See Charter of Rights, *supra* note 5 § 2(b) (citing *Irwin Toy Ltd. v. Quebec (Attorney General)*, [1989] 1 S.C.R. 927, 976 (Can.); *Ford v. Quebec*, [1988] 2 S.C.R. 712, 765–66 (Can.)).

259. Sarah E. Whitmore, Lara Guest & Adrienne Oake, *Risks of Liability Emerging for Online Platforms in Canada*, TORYS LLP (Nov. 23, 2021), <https://www.torys.com/en/our-latest-thinking/publications/2021/11/risks-of-liability-emerging-for-online-platforms-in-canada>.

260. Broadcasting Act, S.C. 1991, c. 11, s. 5(1) (Can.) [hereinafter Broadcasting Act].

261. *Ratings Classification for Canadian English-Language and Third-Language Broadcasters*, CANADIAN BROAD. STANDARDS COUNCIL (2025); *Television Broadcasting Regulation*, 1987, SOR/87-49, s. 2; Broadcasting Act, *supra* note 260, at s. 10(1)(iv) (discussing CRTC-approved TV ratings categories including C, G, and 18+; requiring broadcasters to log advertising material and identify the advertiser for commercial messages).

could be argued as compelling speech that infringes on freedom of expression. However, similar to the CTRC's regulations on publicly labelling content, these provisions are designed to promote transparency, accountability, and user empowerment.²⁶² Further, publications are limited to factual information regarding non-compliance with the proposed Bill.²⁶³ The Commission exercises this discretion on a case-by-case basis and is mandated to consider Charter protections.²⁶⁴ Thus, the labelling requirement aims to provide users with information about the origin of the content, without forcing content removal.²⁶⁵ Providing users with this information promotes informed decision-making for publicly consumed content.²⁶⁶ As such, the provisions address potentially harmful online content when artificially amplified or disseminated through deceptive means.²⁶⁷ Focusing on transparency and user awareness overall, the Charter is likely engaged in a matter consistent with the principles of a free and democratic society.

In contrast, Bill C-63's provisions expanding the definition of hate speech and the confidentiality mechanisms may be found to constitute an unjustifiable infringement. Despite the government's attempt to align the definition of hatred with case law, concerns remain about its vagueness, overbreadth, and potential for subjective application. Previous case law supports the mechanism employed to reduce the occurrence of propagating hate and harmful content generally on online platforms, but does not go so far as to inflate the definition and to expand that application in a manner that could cause inconsistencies or

262. See DEP'T. OF JUST. (Can.), *supra* note 4 (explaining that the Bill requires the labelling of bot-disseminated content to inform users and mitigate manipulative effects and that the Digital Safety Commission can order operators to publish notices of violations to ensure transparency and accountability).

263. *Id.*

264. See Bill C-63, *supra* note 64, §§ 87, 88(2), 119 (explaining that the Commission's power to order the publication of a violation notice is at its discretion and, in exercising that discretion, the Commission must take into account the Charter-protected freedom of expression of the operator).

265. *Id.* § 60.

266. See DEP'T. OF JUST. (Can.), *supra* note 4.

267. *Id.*

overbreadness.²⁶⁸ As such, the suppression of legitimate dissent or opinions on sensitive topics, in particular by marginalized groups, remains quite pressing through such a reformed definition. Thus, there is a real risk that merely offensive or distasteful expressions will be captured. In addition to this, the confidentiality provisions might also undermine the transparency and accountability principles that the government proclaims to prioritize. Despite good intent, restricting access to information and limiting public scrutiny of proceedings only heightens concerns about the subjective nature and vagueness of the amendments.²⁶⁹ In considering the already existing backlogs at the Commission, and the application by unelected officials, the provisions may limit the freedom of expression through eroding fundamental legal cornerstones.²⁷⁰ Resultingly, these risks likely outweigh the government's stated objective of combating hate speech and will need to be reexamined.

With all this in mind, it is important to recognize that Bill C-63 contains necessary provisions to combat harmful content online; however, it is even more imperative to note that they are not sufficient to remedy the problem alone. Specifically, government action through this Bill targets the outcome, not the root cause of the problem.²⁷¹ As such, these steps remain insufficient if the underlying business model that propagates these harms continues to prosper. The design, algorithms, and infrastructure of social media platforms are tailored to incentivize engagement, often by spreading sensationalist and divisive content;²⁷² this incentivization strategy is consistently applied at the expense of accuracy and overall user well-being.²⁷³ Profit

268. See *Saskatchewan*, [2013] 1 S.C.R. at 501—02, 513 (Can.); see *Keegstra*, [1990] 3 S.C.R. 697 (Can.); see *Canada (Human Rights Commission) v. Taylor*, [1990] 3 S.C.R. 892 (Can.).

269. See Bill C-63, *supra* note 64, §§ 88(2), 126, 127(1), 127(3); *Canadian Found. for Children, Youth and the Law v. Canada (Att'y Gen.)*, [2004] 1 S.C.R. 76, 95, 166 (Can.).

270. CANADIAN CIV. LIBERTY ASSOC., *supra* note 178.

271. See Bill C-63, *supra* note 64, §§ 60, 87, 88(2), 119.

272. See Hannah Metzler & David Garcia, *Social Drivers and Algorithmic Mechanisms on Digital Media*, 19(5) PERSPECT. PSYCHOL. SCI. 735, 736—37 (2024) (explaining how algorithms designed for profit rather than societal flourishing carry risks).

273. *Id.* at 736—38.

maximization, driven largely by engagement metrics, consistently leads platforms to prioritize content that generates the highest engagement, irrespective of harm.²⁷⁴ Thus, a fundamental shift is required for how social media platforms operate.

Exploring additional alternatives is needed. One way forward may be the implementation of ethical design principles that optimize user well-being, rather than solely engagement-based advertising.²⁷⁵ While beyond the scope of this Article, it is also reasonable to question whether the current trajectory is reversible. Given that existing business models are structurally designed to incentivize harmful content for profit, the potential pathways for meaningful transformation may be foreclosed or significantly constrained by entrenched economic incentives.²⁷⁶ Nonetheless, it is crucial to consider whether the deeply entrenched profit-driven structures are amenable to genuine change. Further questions need to be asked. Are there economic incentives or regulatory mechanisms that could motivate a shift toward a more ethical business structure? Could a regulatory structure, combined with a collective user behavior shift, compel platforms to adopt more responsible business models? Is a responsible business model possible? Without addressing the economic incentives that drive harmful content, regulatory measures will only offer a partial solution. As such, meaningful and sustainable change requires more than regulatory measures to adjust the existing social media landscape.

VII. CONCLUSION

Social media platforms are intertwined in the daily lives of Canadians. This great influence demonstrates the potential for severe and negative consequences throughout society on a massive scale. As explained in the Department of Justice's Charter Statement for Bill C-63, this Bill aims to balance fundamental freedoms with safeguarding public safety.²⁷⁷ As such, while the provisions regarding the four duties imposed on

274. *Id.* at 736—37.

275. *Id.* at 737—38, 741—42.

276. *Id.* at 735, 743.

277. *See also* DEPT. OF JUST. (Can.), *supra* note 4.

operators will likely be upheld as justifiable limits on freedom of expression, the provisions regarding hate speech and confidentiality measures through statutory amendments could be likely found to violate the Charter beyond a defensible means. Ultimately, Bill C-63 has not been enacted, and the constitutionality of the Bill has not yet been evaluated by the courts;²⁷⁸ thus, it remains unknown how the courts will rule on the competing interests of protecting Canadians from online harms while safeguarding their rights.

Procedural safeguards and accountability measures are crucial to mitigate Charter violations and promote a safer online environment. Government legislation remains vital to set standards and deter online harmful behavior that can manifest into severe real-world consequences. However, these steps alone remain insufficient if the underlying business models of social media platforms remain unchanged. Solely focusing on government regulation ignores the fundamental issue itself, the design and algorithms that incentivize the spread of online harmful content to generate greater engagement and therefore, superior revenue opportunity. Thus, even with legislative regulation and content moderation, the underlying incentive structures will continue to encourage harmful behavior. As identified earlier, there remains the question of whether it is realistic to achieve a shift towards an economic business model that prioritizes user well-being. Irrespective of this notion, it remains clear that a legislative measure alone is not enough to combat online harms. This Article examined the constitutionality of the provisions within Bill C-63, while also underscoring that government regulation is a necessary but insufficient condition to address the causes of online harm. As such, to fully redress the online harms facilitated through social networking sites, the underlying architecture that bolsters these harms must be scrutinized and dismantled.

278. GOV'T OF CANADA, *supra* note 257; *see also* DEP'T. OF JUST. (Can.), *supra* note 4 (explaining that the Bill died on the Order Paper with the dissolution of Parliament and did not pass into law, having only reached its second reading in the House of Commons).