March 6, 2024

The Online Harms Bill – Part 1 – Why We Need Legislation

By: Emily Laidlaw

Matter Commented On: 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*, 1st Sess, 44th Parl, 2024 (Online Harms Bill)

This is the first in a series of posts that will unpack the Online Harms Bill C-63. In this first post, I will explain how and why we got here, as there is a significant amount of misunderstanding about what this Bill is about and why we might need it before the merits of *this* Bill are examined. It also important to contextualize the Bill within the law of intermediary liability, the law that applies to technology companies that facilitate transactions between third parties. Unlike many other jurisdictions, Canada operates in a relative legal vacuum in this space.

Very Brief Overview of Bill C-63

Broadly, Bill C-63 proposes to regulate seven types of harmful content: content that amounts to child sexual victimization, content that is used to bully a child, content that induces a child to harm themselves, content that incites violent extremism or terrorism, content that incites violence, content that foments hatred, and intimate image content (including deepfakes).

The Bill would impose a duty to act responsibly on social media companies, and proposes creation of a body, comprised of a Commission, Ombudsperson and Safety Office, to oversee compliance with this duty.

The duties on social media would be graduated depending on the types of harms (and to ensure balancing fundamental rights). Designated social media (mostly large ones, e.g. Instagram, TikTok and LinkedIn (list of very large online platforms under EU law here)) would have a duty to act responsibly, which would entail identifying and mitigating the risks of harm of their services. They would be required to file a digital safety plan with the Commission and could be investigated for compliance with their risk management obligations. For content involving children, the duty would be stronger, to protect them from harm, and would specifically require that social media implement safety by design measures. For two types of content – child exploitation and intimate image abuse – social media would be obligated to remove such content and the Commission would have the power to order them to do so.

In addition, the Bill proposes several amendments to the *Criminal Code*, RSC 1985, c C-46, and the *Canadian Human Rights Act*, RSC 1985, c H-6. From a victim-centred perspective, including these amendments makes sense. The Bill provides various avenues for victims to address the harm

they have suffered, whether to a Safety Commission for social media, law enforcement, or the Human Rights Commission. However, these provisions are highly controversial, and many of the criticisms have merit, in particular a new hate crime offence for which one can be imprisoned for life, the availability of a recognizance order to keep the peace before a hate crime has occurred, and re-introduction of online hate speech complaints to the Canadian Human Rights Commission. The Bill also proposes uncontroversial amendments to what is colloquially called the 'Mandatory Reporting Act' (*An Act respecting the mandatory reporting of Internet child pornography by persons who provide an Internet service*, SC 2011, c 4).

I will examine the *Criminal Code* and *Canadian Human Rights Act* elements of the Bill later. For now, I will discuss the 'why' behind the core of the Online Harms Bill, namely the proposal to regulate social media. The devil is in the details with this Bill, and key points of controversy are evident in seemingly mundane provisions. In short, while amendments should be made to the Bill, the Bill gets the big things right. Some decisions – namely what to scope in and out of such a Bill – are difficult choices, and I will endeavour to highlight for readers those key decision points as they will likely be topics of debate as this Bill makes its way through Parliament.

Why Regulate Social Media for Online Harms

Companies have always been at the centre of the internet's architecture. In order to connect to the internet, search for information, consume and share ideas, bank, shop, store and secure our data and systems etc., a company intermediates the transaction. As a result, these companies gatekeep the internet and have tremendous power. Despite this, they have generally been treated in law as secondary actors, and often neutral ones, that simply connect between points A and B. I defame my neighbour on X (Twitter). I am the primary bad actor and X is secondary, acting as host of the content. Intermediary liability asks what legal duties should be imposed on these secondary actors.

The law and policy issues that have emerged are complicated. First, as the internet matured, many of these companies became big, and no longer simply hosted a platform for users to interact, but also offered a whole bunch of other products and services in multiple markets. Think of Amazon's varied businesses, whether connecting buyers with sellers, selling its own goods, its streaming service, cloud storage, among other things. These are the quintessential "big tech" of Amazon, Microsoft, Alphabet (Google), Meta, and Apple. They could no longer be called just intermediaries, as sometimes they were transacting directly with a customer, so how to view these companies became blurry. What was clear was that they were powerful (thus often called platforms although that still has an air of neutrality to it) and it was hard to slot them into existing legal frameworks.

Second, with the rise of social media and the many benefits it delivers (and yes there are many), a wave of toxicity and harm was unleashed that has profoundly rocked our society. Social media sites are perfect vectors for online abuse: cyberbullying, radicalisation, child sexual exploitation, fomenting hate, mob attacks often against our most vulnerable and marginalized, targeting journalists and elected officials, the list is endless. And at the same time social media evolved, first with photos, then videos and livestreaming. And then all of the internet seemed to become a space for discourse and harm: private messaging, gaming, virtual reality, and so on. The problems have become further complicated with the rise of generative AI and the spread of deepfakes,

disinformation, and harassment through synthetic means. As an attack vector, the challenges range from children bullying other children to state-backed information warfare that might upend elections and undermine democracy. That is a lot of harms, with different dynamics, to potentially take on.

For the most part social media have self-regulated all of the above through their terms and conditions of use, and content moderation policies. They built trust and safety teams, and even oversight boards. Multistakeholder bodies emerged such as the Global Internet Forum to Counter Terrorism (GIFCT) and the Global Network Initiative (GNI) that brought together government, civil society, academics, and industry to devise voluntary solutions. The problem remained that not all social media cared about self-regulation, and even those that did were not always above board in their practices. A change in ownership, as we saw with Twitter, could lead to starkly different approaches to content moderation. This has always been a problem with corporate social responsibility, that it is inherently unstable and dependent on corporate commitment (and whims), and it is difficult to make an industry practice.

The disparate approaches to "harms" create an unmanageable ecosystem from a governance perspective. A mainstream platform might closely moderate extremist videos, but it is more difficult to detect when it is posted on an alternative platform, with few rules, and then shared on that mainstream site. Similarly, users kicked off YouTube might jump to alternative platforms and continue their behaviour unabated. Trust and safety on social media are at a crucial inflection point. Many social media platforms that tried to moderate are suffering from fatigue and throwing up their hands. Twitter, Meta, Amazon, and Alphabet have all scaled back their trust and safety teams.

Third - and this is critical to understand the Online Harms Bill - the issue is no longer just about social media as intermediaries hosting third party content, but about social media as primary actors in amplifying or causing online harm through their *design*. Meta's decision to rate highly emotive reactions (anger, love) as five times more valuable than more neutral reactions (like) propelled divisive content to the top of our feeds (see Facebook Files). Pinterest and Instagram's recommender system pushed self-harm and suicide content to Molly Russell before her suicide, leading the coroner to conclude in their inquiry that it was likely a contributing factor in her death. These algorithms are a design feature for which social media are primarily responsible.

It is not just the recommender systems. For example, Discord has been used to livestream child sexual exploitation. Its content moderation system does not enable in-service reporting of livestreamed abuse. Yet the risks of livestreamed content are well-known, such as the livestreaming of the terror attack in New Zealand in 2019. It is arguable that Discord should design its content moderation system in a different way. 'Safety by design', as it is called, is an approach to social media regulation that moves beyond individual content decisions (quintessential intermediary liability) to how social media is designed and built, much like a car. It is more akin to consumer protection law and requires that the system design complies with certain safety standards. Of course, social media are not the same as cars, because of freedom of expression, and therefore the analogy only takes us so far.

Current (Lack of) Canadian Law

Canada does not have a comprehensive federal intermediary liability law. The law has developed primarily in two areas: defamation and copyright law. In defamation law, intermediaries are provided a conditional safe harbour from liability provided that they remove defamatory content once notified (see here for more). Defamatory content is notoriously difficult for intermediaries to assess, and the incentive with this type of notice and takedown model is to remove content even if its lawfulness is unclear. With the adoption of the Canada-United States-Mexico Agreement (CUSMA), the law is changing in Canada. Article 19.17 of CUSMA essentially imports US intermediary liability law, with the result that Canadian courts must not treat intermediaries as "publishers" and therefore must not hold them liable for defamatory content posted by third parties. In short, the conditional safe harbour has become a broad safe harbour.

Canadian copyright law creates a notice-and-notice regime for intermediaries, requiring that internet service providers (ISPs) pass on notices of infringement to their customers. Failure to do so risks imposition of an administrative monetary penalty, not liability. Provincially, Quebec is the only province with a comprehensive law that applies to intermediaries for illicit activities, operating like a conditional safe harbour (see here and here for more). Otherwise, a court order to remove content can be obtained, but that is different than liability as explored here (e.g. see *Criminal Code*, ss 83.225(5) [terrorist propaganda], 320.1 [hate propaganda], 164.1(5) [intimate images, voyeurism and child pornography]; and various civil tools).

There are many laws that apply to an aspect of what intermediaries do, in particular, privacy, competition and anti-SPAM law (CASL), and the controversial recent passage of Bills C-11 and C-18. However, for online harms proper, the above are all we currently have. The result is that, in Canada, social media have shockingly few legal obligations concerning how they manage the risks of harm of their services given their impact on society. For individuals and groups, they are at the mercy of what social media chooses to do or not do. In terms of safety by design, unless the focus is on the collection, use and disclosure of data (see eg *Personal Information Protection and Electronic Documents Act*, SC 2000, c 5 (PIPEDA); and similar) or misleading advertising (see *Competition Act*, RSC 1985, c C-34), we do not have a law in Canada that mandates safety standards for social media.

Other Jurisdictions

With Pierre Trudel I co-chaired the expert panel that advised the Federal Government on this law, and our group often discussed the late mover advantage of Canada, and the risk that the advantage might be lost if the government did not act soon. Contrast our lack of law with the European Union (EU). Since 2000, intermediaries as hosts have had a conditional safe harbour for illegal content posted by third parties provided that they remove this content expeditiously upon actual notice (see E-Commerce Directive). This operates as a notice and takedown model applied to all illegal content, whether privacy, copyright, hate speech, terrorist propaganda, and so on. The regime was studied for many years, and had its flaws, but the EU has an almost 25-year head start on Canada in this space.

Notably, the EU and United Kingdom (UK) have also moved onto a second iteration of these laws, with the *Digital Services Act* (DSA) adopted by the EU Parliament in 2022 and the UK's *Online Safety Act* (OSA) passed in 2023. Crucially, these laws tackle the systemic design problems with platforms. The laws are different from each other, but at their core they impose risk management obligations on platforms. By this I mean that platforms must identify risks of harm for certain categories of content and take steps to mitigate those risks. They must be transparent about their practices, and a regulator is tasked with oversight. The OSA is more focused on children, while the DSA is broader (and includes disinformation).

Australia has had an eSafety Commissioner since 2015. Its remit has grown over time from protecting children from cyberbullying to include image-based abuse, adult bullying, and abhorrent violent material. The Commissioner has the power to order content removal in narrow circumstances, develop codes of practice with industry, and public education. Recently their oversight expanded to include a systemic focus, with oversight of Basic Online Safety Expectations for technology companies.

Even the US is taking on platforms. In 1996 it passed the earliest intermediary liability law in the world. Section 230 of the *Communications Decency Act* creates a broad immunity from liability for intermediaries, except for federal criminal, communications privacy, and intellectual privacy matters. The notable exception is copyright, which requires swift removal of copyright infringing content upon notice (see *Digital Millennium Copyright Act*). This broad immunity is the blueprint for article 19.17 of CUSMA. However, even the US is poised to adjust its blanket immunity approach that has shaped so much of our online worlds. There are currently hundreds of federal and state laws (proposed and passed) that impose obligations on social media. The legality of the Texas and Federal laws are currently being argued before the US Supreme Court. Most consequential is the Federal *Kids Online Safety Act*, which proposes a duty of care to prevent and mitigate harms to minors with oversight by the Federal Trade Commission.

As is evident, online safety regulators are being created in several jurisdictions, such that a Global Online Safety Regulators Network has been created for collaboration, with members from Australia, France, Ireland, South Africa, Republic of Korea, UK, and Fiji. Canada's Department of Heritage currently has observer status.

The Journey to Bill C-63

In 2021, the Federal Government published a discussion guide and technical paper outlining their proposal for online harms legislation. It was widely criticized, including by me (with Darryl Carmichael here). It focused on an 'old school' style of online harms law, requiring swift content removal within 24 hours modelled on Germany's *Network Enforcement Act*. It did not focus on risk management obligations or safety by design. The federal election was called soon after. After the election in September 2021, the Government regrouped. The Department of Canadian Heritage led the online harms file and proceeded to consult widely. They put together an expert advisory group, mentioned above, held a Citizens' Assembly on Democratic Expression, which published a report of recommendations, and held roundtables across Canada and with various stakeholders through 2022 (see here). In early 2023, the Government published a 'What We Heard' summarizing various consultations. One could glean from it the approach the Government might

take, but then everything grew quiet for a while. Sometime in the summer of 2023 the file moved from Heritage to the Department of Justice. The Bill was tabled on February 26, 2024.

Next Steps

There is much to debate about the Bill, and I will examine each part in turn in the coming weeks. The issues include, among others:

- ***Scope of the Bill – the regulated***: The Bill applies to social media services. It excludes private messaging, gaming, and search engines. It also only applies to large social media services unless designated, thus smaller platforms that are hotbeds for illegal behaviour would be excluded.
- ***Scope of the Bill – the harms***: The question here is whether the harms selected to be addressed by this legislation are appropriately scoped to ensure protection of freedom of expression and harm reduction. For example, I had initial concerns seeing cyberbullying on the list, but the definition is narrow (focused on risks of serious harm that was communicated intentionally to threaten, intimidate, or humiliate) and it seems to be modelled on the approach in Australia (see s. 6 of Australia's *Online Safety Act*).
- ***Structure of Commission/Ombudsperson/Office***: What powers should a body have, and how is this power constrained; will it be effective to redress the wrong and what is missing. For example, is flexibility to deal with evolving technology and threats balanced against the need for legislative clarity? The broad powers of inspectors, and regulatory powers of the Commission, have already been highlighted as issues of concern.
- ***Fundamental rights***: How are fundamental rights protected or neglected in this Bill? It is clear that the Commission is obligated to account for fundamental rights, but no similar obligations are imposed on social media companies. In my view, safety includes fundamental rights, and therefore companies should be obligated to protect rights as part of safety by design.
- ***AI Laws****:* Social media are AI ecosystems. AI is embedded in both the threats and solutions. Any Digital Safety Commission will be an AI regulator, and it is unclear how this body would and should interact with the proposed new AI and Data Commissioner in Bill C-27.
- ***Criminal Code:*** The Bill proposes to make significant amendments to the *Criminal Code* that include creation of a new hate crime offence, which risks imprisonment for life. This offence makes available a recognizance order to keep the peace where there are reasonable grounds to fear a hate crime will happen in the future. These provisions should likely be removed from the legislation but will be examined in more detail.
- ***Canadian Human Rights Act:*** The Bill proposes to re-introduce s 13 to the *Canadian Human Rights Act*, which would enable complaints of online hate speech to be made to the Commission. Section 13 was repealed in 2014. The proposed s 13 is narrower than the old version and implements the standard set down in *R v* Whatcott, 2013 SCC 11 (CanLII). Professor Jennifer Koshan and I, with several students, advocated for the re-introduction of s 13 in 2021. However, the concerns that the Commission will be overwhelmed with complaints, and of weaponization of the process, are valid and will be examined in due course.

This post may be cited as: Emily Laidlaw, “The Online Harms Bill – Part 1 – Why We Need Legislation” (6 March 2024), online: ABlawg, http://ablawg.ca/wp-content/uploads/2024/03/Blog_EL_Online_Harms_Bill_Context.pdf

To subscribe to ABlawg by email or RSS feed, please go to http://ablawg.ca

Follow us on Twitter @ABlawg