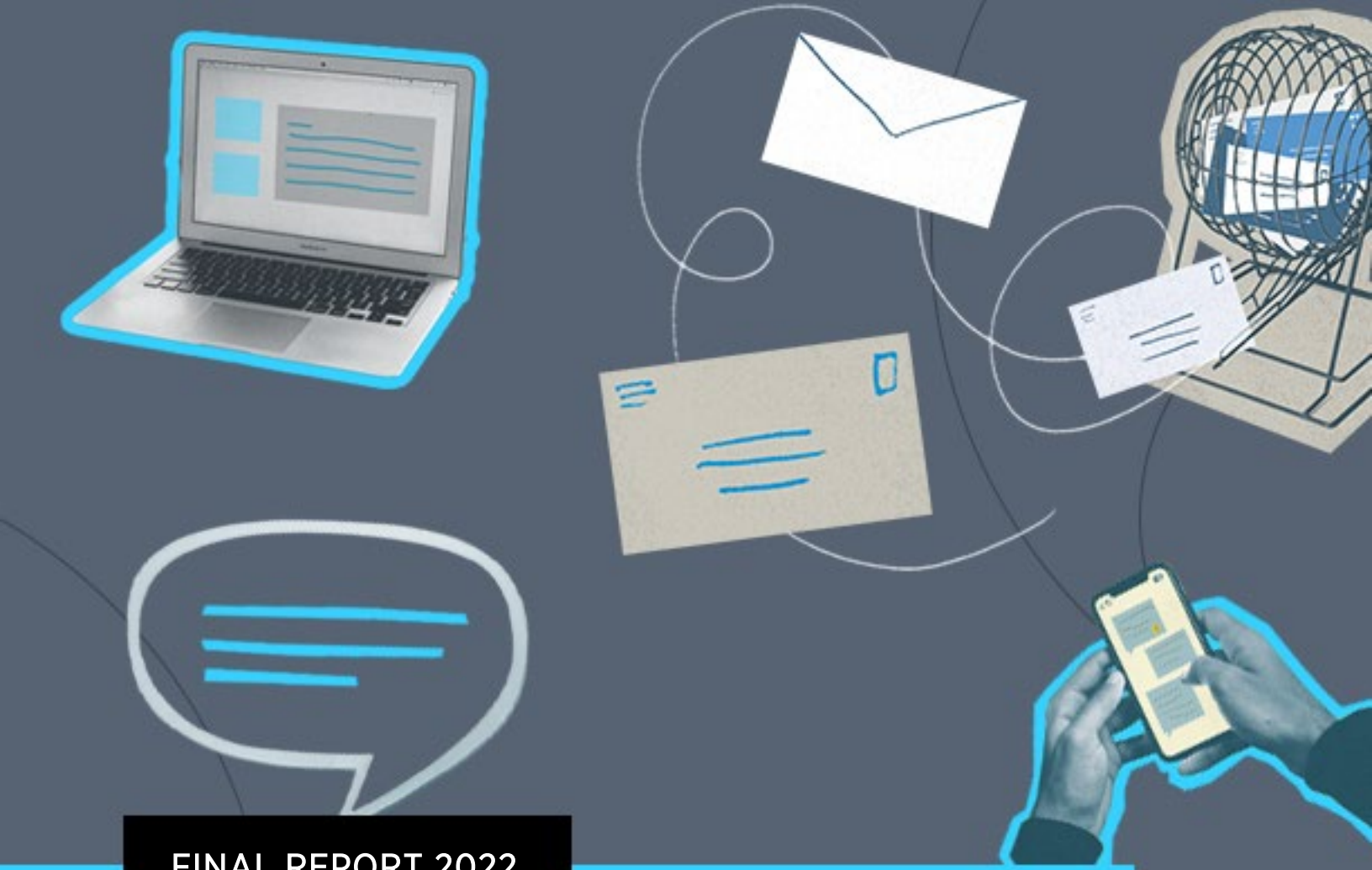


FINAL REPORT 2022

Canadian Commission on Democratic Expression

How to Make Online Platforms More Transparent and Accountable to Canadian Users

MAY 2022

DEMOCRATIC EXPRESSION
DÉMOCRATIQUE

PUBLIC POLICY FORUM
FORUM DES POLITIQUES PUBLIQUES
PPF

The Public Policy Forum works with all levels of government and the public service, the private sector, labour, post-secondary institutions, NGOs and Indigenous groups to improve policy outcomes for Canadians. As a non-partisan, member-based organization, we work from "inclusion to conclusion," by convening discussions on fundamental policy issues and by identifying new options and paths forward. For more than 30 years, the PPF has broken down barriers among sectors, contributing to meaningful change that builds a better Canada.

1400 - 130 rue Albert

Ottawa, ON, Canada, K1P 5G4

Tel: 613.238.7858

www.ppforum.ca

ISBN: 978-1-77452-114-4

TABLE OF CONTENTS

ABOUT THE INITIATIVE

DEMOCRATIC EXPRESSION
DÉMOCRATIQUE

The Canadian Commission on Democratic Expression is a three-year initiative, led by the Public Policy Forum that aims to bring a concerted and disciplined review of the state of Canadian democracy and how it can be strengthened. The centerpiece is a small, deliberative Commission which will draw on available and original research, the insights of experts and the deliberations of a representative citizen's assembly to assess what to do about online harms and how to buttress the public good. The Commission is designed to offer insights and policy options on an annual basis that support the cause of Canada's democracy and social cohesion. The Commission is supported by national citizen assemblies as well as by an independent research program.

This initiative grew out of earlier insights about the relationship of digital technologies to Canada's democracy covered by the Public Policy Forum's ground-breaking report, *The Shattered Mirror* and its subsequent interdisciplinary research outlined in the *Democracy Divided* report (with UBC) and through the Digital Democracy Project partnership with McGill university.

The initiative is delivered in partnership with MASS LBP and the Centre for Media, Technology and Democracy at McGill University's Max Bell School of Public Policy, who are executing the national citizen assemblies and research program, respectively.

To learn more about the initiative and how you can become involved, please visit www.ppforum.ca/project/demx. The initiative will run from April 2020 to March 2023.

This project has been made possible in part by the Government of Canada.
PPF would also like to thank the McConnell Foundation for their support.

Canada

La Fondation McConnell Foundation

EDWARD GREENSPON

FOREWORD

In the months leading to this second report of the Canadian Commission on Democratic Expression, we all bore witness to an internet-fuelled protest movement at home, and the invasion of Ukraine by the world's leading online and offline intruder into the democratic affairs of other nations. In both cases, the web and its most influential offshoot, social media, did their job of sharing, selecting, ranking and amplifying information, misinformation and disinformation.

That the internet boasts an unprecedented capacity to distribute the genuine social goods of news and views merits daily celebration. As with the printing press before it, this communications technology is endowed with powerful democratizing forces. It gives voice to the marginalized, enlarges access to knowledge and enables human connection to transcend physical limitations.

But its dark side must not be ignored. The intentional distortions of truth and the targeting of groups and individuals with constant barrages of hate, harassment and humiliation eats away at the cohesion and commonweal that undergird well-functioning societies. The very cause of democratic expression is being turned on its head as those victimized by online abuse are driven from the public square. The printing press has seen its excesses, too, but nothing matches the scale of the internet.

The Public Policy Forum has been working away at informational system policy questions for more than five years in reports such as *The Shattered Mirror, Democracy Divided, What the Saskatchewan Roughriders Can Teach Canadian Journalism, Harms Reduction* and *The Shattered Mirror Five Years On*. Each delves into the fallout of new information and communications technologies on the health of our democracy, and what can be done about the weakened state of journalism on the polluted state of online discourse without endangering free expression.

If the medium is the message, as Marshall McLuhan asserted, what is the message built into our contemporary concourses of communications? What is the collective effect of their speed, volume, reach and ceaselessness? Do these qualities lead to snap judgement over deliberation, condemnation over empathy, virtual community over physical community, experience over expertise, emotion over reason?

Critically, how can we ensure that those programming the lightning decisions of the internet via algorithms bear responsibility for what they wreak and choose to weed out the bad effects? While it can look like open

space, the internet has long been colonized by giant corporate interests that exercise dominion over what gets circulated and emphasized on vast swathes of informational terrain. They make choices that ripple through societies with serious consequences.

The first report of the Commission on Democratic Expression, *Harms Reduction,* published in 2021, focused on the content of the internet, most particularly the challenges of online hate and related harms. It offered a six-step program starting with platform companies carrying a duty of responsibility to keep the internet safe for users, just as a landlord would for the tenants of a highrise complex. But it was just a start – the Commissioners felt the next round would need to go further in countering the structural biases that favour certain types of information regardless of truth or utility. If these abuses are more than random acts of the ill-intentioned but are somehow the product of systemic choices, what could and should be done about that?

That is the starting point for this second report by the Commission. It goes beyond the content we can all see to inquire into how it came to gain prominence, and how one person's hate or disinformation translates into the rallying cry of a mob or the discombobulation of social order. As readers of this volume will see, these outputs are the product of two related factors: the inputs of billions of pieces of data and the algorithmic closed box systems that manage, aggregate and distribute that data to maximize audience engagement and advertising revenue. This report examines what to do to ensure a better balance of power over these control systems.

I want to thank the nine Commissioners who have taken on this daunting task of discovery and deliberation with a civility that should serve as a model for the digital age. I also want to thank lead writer Chris Waddell, our research partners at McGill University's Centre for Media, Technology and Democracy and the members of the parallel Citizen's Assembly on Democratic Expression and its architects at MASS LBP. And last but always most, my colleagues at the Public Policy Forum.

Edward Greenspon
President & CEO
Public Policy Forum

EXECUTIVE SUMMARY

Following six months of study and deliberations, the Canadian Commission on Democratic Expression has settled on a series of principles and recommendations that underlie a practical course of action for the public, governments and social media platforms to protect democratic expression in Canada and to counter the harms created by content posted, shared and amplified through social media platforms. We recognize the complexity of the issues at play in a free and democratic, rights-based society, and offer these recommendations as a path forward. We encourage debate that may further refine them and lead to their implementation in the coming months by governments and regulators.

Principles

1. Free speech is fundamental to a democratic society. The internet enables more people to participate in public discussions and debates.
2. The rise of hatred, disinformation, politically polarizing content, conspiracies, bullying and other harmful communications online is undermining these gains and having a corrosive impact on democratic expression in Canada, both online and offline.
3. The status quo of leaving content moderation to the sole discretion of platforms has failed to stem the spread of these harms. Platform companies can find themselves in conflict between their private interests and the public good.
4. The notion that platforms are neutral disseminators of information is faulty. Platforms distribute content to serve their commercial interests and so must assume greater responsibility for the harms they amplify and spread while being mindful that free expression is the bulwark of a democratic society.
5. Public agencies must play a more active role in furthering the cause of democratic expression and protecting Canadians from online harms.

6. Any policy response must put individuals first, reduce online harms and guard against the possibility that will have chilling effects on participation particularly for Indigenous peoples and other equity seeking groups. This requires a balanced and multi-faceted approach.
7. Privacy protections are fundamental to human rights and democratic expression. Remedies to address online harms and protect democratic expression should centre on rights-based approaches, both in safeguarding citizens against possible privacy infringements and empowering individuals with greater control over their data and democratic expression.
8. Minors (under 18) are particularly vulnerable to online harms and so platforms which are accessed by minors have a special obligation to ensure that the appropriate safeguards are in place.
9. Many like-minded democracies are facing similar challenges to democratic expression in the 21st century and as such Canada should look to act multilaterally and in concert with other like-minded nations as much as it possibly can and where it makes the most sense.

These principles have led the Commission to an interrelated set of recommendations around three core themes of democratic expression: transparency, accountability and empowerment.

Recommendations

THEME ONE: TRANSPARENCY

1.1. Mandate and enforcement: Create and empower a regulatory gatekeeper to mandate and enforce access to data contained within social media platforms for research and oversight purposes.

The regulatory gatekeeper would enforce mandatory platform data access and sharing requirements outlined in legislation for each of three tiers: the general public; accredited researchers and journalists; and more specialized and detailed research in the public interest that without significant additional safeguards could have privacy implications.

1.2. Implement tiered data access: Mandate separate tiers of data access with safeguards for the public, for researchers, journalists and civil society groups.[1]

Introduce graduated tiers of data access obligations for platforms to give the public lower-level access (tier 1); to give researchers, civil society actors and journalists mid-level access (tier 2); and to give a narrow range of applicants conducting specialized public interest research access that requires more significant detailed safeguards to balance privacy protection with the need for greater platform transparency and accountability (tier 3).

1.3. Mandate universal digital ad transparency.

Introduce a mandatory duty for platforms to disclose regularly and archive, in a standardized format, specific information about every digital advertisement and paid content post on their platforms. Disclosure and archiving shall be universal, meaning that platforms shall present data using a standardized machine-readable format with common minimum disclosure standards. All this data will be preserved in a single, central archive with its content available to the three tiers of users detailed in earlier recommendations.

1.4. Introduce stronger protections for whistleblowers.

Given the economic, professional or personal risks whistleblowers may face, any recommendations to protect individuals who report and expose internal corporate malpractice should guarantee against legal, economic and reputational retaliation by their employers. The federal government could draw from other sectors in advancing private whistleblower protection.

THEME TWO: ACCOUNTABILITY

2.1. Increase the capacities of public agencies: Ensure existing regulators are properly empowered and equipped to operate in the 21st century digital world efficiently and effectively. In addition, create a new independent federal regulator (as noted above and proposed in the Commission's first report) vested with investigative, auditing and enforcement powers and responsibilities to ensure a new legislated duty to act responsibly is applied to platforms. Over time, the new regulator would also be entrusted with conducting systematic review of policies and regulations and suggesting proposals for reforms when needed.

With a primary mandate to oversee and enforce the duty to act responsibly, this new regulator would focus on platform systems and operations inside the 'closed box' to promote and ensure transparency and accountability, investigate perceived harms, assess platform liability, and determine and enforce remedies when platform liability is established. The regulator should be formally independent from the government, mainstream media and the platform. It is crucial for the regulator to be properly resourced and carefully designed, based on public consultations with a focus on fleshing out its scope and remit. The regulator should operate in a transparent and accountable fashion with legislated reporting to Parliament at regular intervals.

2.2. Mandate tiered obligations for different types of platforms and/or for services likely to be accessed by minors and adults.

All platforms regardless of size have a duty to act responsibly. However, obligations on individual platforms may differ based on the type and size of platforms according to their capability to comply with legislated

requirements. There would also be differentiation between the obligations placed upon platforms based upon use by minors (under 18), and adults.

2.3. Legislate intermediary liability protections and exceptions for platform liability.

Clarifying when platforms can be held legally liable for harmful consequences generated by platforms' algorithmic recommender systems and amplification tools would encourage platforms to better moderate user-generated content. Users should be allowed with care to express freely their views online (within the limits of what is permitted under Canadian law). At the same time, making platforms strictly liable for harms arising from user expression can result in over-removal of material and censorship. Nevertheless, the degree to which problematic content is amplified and the impact that recommender systems have on public opinion should be taken into consideration when addressing platforms' role in democratic societies.

2.4. Empower regulatory entities to develop and implement a rights-based Algorithmic Accountability Framework which includes Algorithmic Impact Assessments (AIAs), Human Rights Impact Assessments (HRIAs) and algorithmic audits.

Relevant regulatory entities should be empowered to develop and implement a robust algorithmic accountability framework which is centred on rights-based approaches to algorithmic governance. Centring rights in accountability frameworks for automated decision-making systems follows international standards for addressing the heightened risks to safety and fundamental freedoms, such as the right to freedom from discrimination.[2]

2.5. Develop a Code of Practice on Disinformation.

Canada should develop a Code of Practice on Disinformation establishing commitments and requirements in collaboration with major online platforms. The overall aim of the Code is to promote the creation of platform-based policies and procedures to address disinformation, including demonetization of problematic content, increased transparency of political and issue-based advertising, empowering users to better control their online activities and enabling privacy-compliant access to data for fact-checking and research activities.

THEME THREE: EMPOWERMENT

3.1. Support Indigenous knowledge, relationships and protocol development and Indigenous data governance for Indigenous communities.

Support meaningful participation of Indigenous peoples, and ensure Indigenous relationships and protocols are built into the development of technological and social policies, tools and mechanisms. The federal government should collaborate with Indigenous peoples, communities and organizations to ensure that Indigenous data governance rights are respected and that Indigenous peoples have the means to pursue self-defined agendas. Additional support should include funding, the creation of new legislative proposals, literacy programs around data ownership and self-determination, and other needs identified in partnership with Indigenous peoples and communities.

3.2. Substantially strengthen civic education respecting rights, digital literacy and access to quality information to support equity-seeking groups and community-led programs.

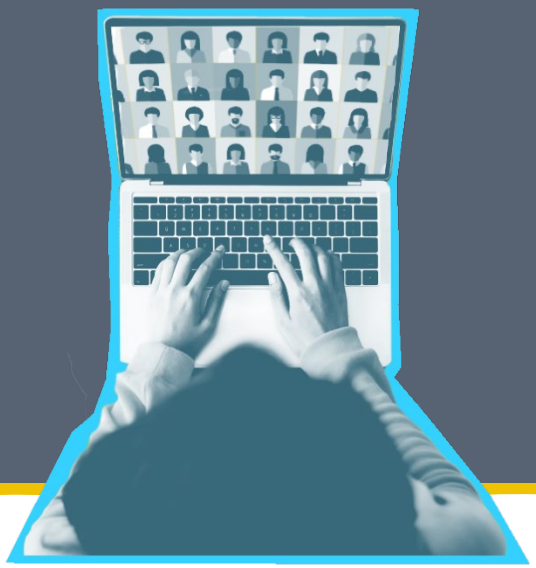
Public education and digital literacy initiatives should provide the public with an understanding of their rights and freedoms, how digital media works, how it can impact public opinion, and how structural biases operate within it and reinforce inequities in real life. This includes equipping citizens with skills to identify biases and assess the reliability of information, how to search, navigate, synthesize and evaluate content online, and how to meaningfully participate in communities online. Underrepresented groups should be supported through targeted policies and programs that strengthen equity, including funding for the digital production of Indigenous cultures and knowledge. Programs should be offered in multiple languages including Indigenous languages.

3.3. Mandate Interoperability and Data Mobility.

Information systems should be able to regularly interact and exchange information with one another which would allow alternatives such as start-ups and platform co-ops to connect with existing services. Canada should ensure the interoperability of digital services to empower individuals with greater choice and control over their interactions online. Additionally, Canada should introduce the right to data portability – giving individuals the right to have their personal data transmitted directly from one platform to another, without hindrance.

3.4. Modernize Canada's Privacy Legislation.

Canada should update its privacy legislation to a rights-based framework for current and future technological developments. The Privacy Commissioner of Canada should be given greater authority to modernize Canada's current privacy legislative framework and decide how private platform companies can collect, process and target individuals' data.



COMMISSION'S PREAMBLE

A year ago, the Commission's first report began by noting: "We embrace the astounding ways in which the internet, and social media in particular, have lowered barriers to participation in the public realm. They have strengthened our democracy by giving individuals new ways of making themselves heard, new ways of organizing politically, new ways of engaging with elected representatives and new ways of holding power to account.

Today, anyone can own a corner of the internet. More people than ever can benefit from enhanced access to knowledge, community and collective action. But there is also the darker side. Along with a more open and accessible public square has come a less trustworthy and safe one. This represents one of the central paradoxes and challenges of our times – and of this paper."[3]

A year later, all that still remains true. Yet it is also true that the rising tides of hatred, disinformation, conspiracies, misogyny, bullying and other harmful communications online have become a flood. This tsunami of social and democratic harms is driving women, minorities, Indigenous peoples and others from the digital public sphere.

Online harassment affects and shapes our offline experiences with damaging results for democratic expression. It distorts politics on a daily basis and can damage relationships in families, in workplaces, in schools and in communities.

Disinformation undermines public debate by depriving us of a shared understanding of the facts. Separating citizens into information silos compromises our ability to make collective decisions. This is particularly serious in a diverse nation like Canada in which accommodation and social cohesion are necessary values.

We cannot allow those intent on frustrating democratic expression in Canada to hijack this great opportunity for democratizing enrichment. The laundry list of harms is too long to ignore misogyny, racism, anti-Semitism, Islamophobia, white supremacy, homophobia, disinformation, alternate facts, bullying, phony consumer reviews, seniors' fraud, counselling of suicide, conspiracy theories, attacks on electoral integrity, incitement to violence – on it goes. It has reached a point where the targets of harassment sometimes feel their health and safety requires them to withdraw from the digital square – the very opposite of democratic expression.

Internet platforms have been insufficiently diligent in turning back these harms; indeed, their systems are in many respects complicit. Harmful and hateful speech are not so much anomalies as the logical products of social media's structures, design, policies and practices.

As occurred last year, the Commission worked in parallel with the second PPF-organized Citizens' Assembly on Democratic Expression, comprised of 42 Canadians from all 10 provinces and three territories who developed their recommendations through a series of online and face-to-face meetings in the fall of 2021. Their report in January 2022 should be read in conjunction with this report.[4] The process of the Citizens' Assembly was remarkably reassuring to observe. While the online debate and discussion of these issue is predictably divided and toxic, when you get 42 Canadians from all walks of life and ideological perspectives together it is remarkable how civil the discussion becomes. While all entered the process with their own experiences, subjectivities and perspectives of the problem and role of government in the solution, they left astonishingly unified. This reality was front of mind for the commissioners as we deliberated this year. Citizens want action and are far more unified than the online and media discourse might make us believe. We drafted our recommendations in this report with their latest recommendations from the Citizens' Assembly in mind.

As the Commission noted in last year's report, ultimately it is the role of governments to protect its citizens against social harms, stand up for the targeted and assert the greater public interest through the appropriate governance of platforms, search engines and other intentional or incidental purveyors of this material. The recommendations in this report suggest how governments can strike a balance in doing that.

We know that there is much unlawful content online: hate propaganda, harassment, threats, that ought to be challenged. Enforcement of existing laws is a responsibility of governments. In the long term, failure to enforce current regulations and laws threatens the rule of law. This report recommends new and additional measures to curtail online harms, it should not be read as dismissing efforts to enforce our existing legal regime.

However, there is much online that is 'lawful but awful' that undermines and harms democratic expression in ways we have highlighted. But it is important to understand that the harms to democratic expression online have direct effect on the "offline" world as well, including how politics play out in the voting booth, and how

people interact in the workplace, in schools and in public. In particular global contexts, online incitements to violence and forms of harassment are directly implicated in offline events that have enormous implications for democracy, including social media's role in the February 2022 Canada-U.S. border blockades and occupation of downtown Ottawa by protesters, as well as the January 6, 2021, attack on the U.S. Capitol and the genocide in Myanmar.

In this world of 'lawful but awful' expression, users are both the victims and also the perpetrators. Individuals and organizations post misogynistic and bullying content as only two examples of the many harms noted above. Other users share those harmful posts, amplifying the harms created and spreading them to a wider audience. Users must take greater care in ensuring they do not post material that is harmful, while platforms must address how this material gets amplified and shared, as well as the removal of such material when it is the subject of a complaint by a user or by those hurt by the material, or it presents an imminent threat to a person.

More generally, those duties are grounded in who we are as a society. With our Charter of Rights and Freedoms, Canadians pride ourselves on living in a rights-based society under the rule of law grounded in respect for and protection of human rights. This includes individual rights and also rights of marginalized groups and communities within a multicultural nation. The rise of online disinformation and misinformation pose new challenges to our society requiring a multi-faceted response by our governments and institutions as proposed in the recent second Citizens' Assembly report.

What will democratic expression mean and look like in the future? What was once spoken has become multimedia and is given instant global reach and impact through social media platforms. This includes the posts of anonymous actors and automated bots amplifying messages, designed to create the illusion of consensus around lies and conspiracy theories. Threats to democratic expression can emerge even from the content of video games and new technologies such as augmented reality and the place the metaverse may play in all our futures.

Platform companies must commit to the protection, promotion and enhancement of human rights, equality and the responsibilities that accompany that. This is especially so for children and their rights. Children, adolescents and teens live on social media platforms as much or more than adults. The impact of what they post, see, hear, read and share shapes the development of their brains and identities in ways that don't apply to adults. The duty of social media platforms to act responsibly must specifically ensure in spirit and action that children's activities on the platforms receive dedicated attention and protection. Regulators have a duty here to craft and enforce best practices regarding children and social media that respects the distinct challenges that emerge from different content and interactions. The principle of data minimization should be strictly applied when it comes to children.

Finally, the right to privacy must not be compromised. By their nature, threats to and limits on privacy directly constrain and risk undermining the equality of individuals and groups to exercise their rights to democratic participation and expression. That cripple's democracy.

Citizens cannot participate in a democracy without a right to privacy and the ability to exercise some control over what others know about them.

As it pertains to social media platforms, their systematic data capture infringes on the right to privacy. There is emerging consensus around the world that systematic data collection may additionally interfere with other human rights as well, given that the right to freedom of opinion is potentially impaired by gratuitous data collection and profiling.[5] Public awareness of online surveillance can also lead to self-censorship, presenting a "chilling effect" of data capture on public participation and engagement.[6] In its first-ever report on disinformation, the United Nations Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression finds freedom of expression safeguards and privacy safeguards to be complementary. Canada's efforts to protect and promote democratic expression should be seen as concurrent to its modernization of data privacy legislation. Without a joint and coordinated effort, governments and individuals will be forever playing catch up – applying small band-aids rather than addressing the cause of wounds.

We believe these threats to democratic expression can be countered by introducing and implementing on social media platforms an enforced duty to act responsibly under three interconnected themes in defining the relationship between social media users and the platforms: transparency, accountability and empowerment.

The objective of our recommendations is to achieve a more equitable balance of power between social media users and the platforms they use that will benefit both.

Rick Anderson

Wendy Chun

Nathalie Des Rosiers

Amira Elghawaby

Merelda Fiddler-Potter

Philip N. Howard

Vivek Krishnamurthy

Beverley McLachlin

Taylor Owen

CHAPTER ONE: DEFINING THE PROBLEM

Social media has changed so much of our lives in the past two decades.

Canadians routinely use social media to keep in touch with family and old friends both near and far, sharing the joys and sorrows of their lives, their photos and videos, their hopes and dreams. It has brought people together in groups building new friendships around their shared interests. It has become the way many get their news and information and then share it with their friends and acquaintances. Social media has become a part of daily life that many look forward to checking in with regularly, whether it is through posting material to inform and impress their friends or just killing time scrolling through the posts of others looking for things that are different, unexpected or funny that catch their eye.

As the number of users grows on social media platforms, so does their collective impact. Platform companies have a major influence on shaping and distributing news and information with a global reach and audience, delivering a message from anyone at any time to everyone or to a deliberately selected micro-audience. The ability for users to remain constantly engaged can generate massive advertising revenue and profits for the platforms.

As noted in the Commission's first report, published in January 2021, the platforms are not neutral disseminators of information. They curate content to serve their commercial interests and they have a responsibility for the harms that content can amplify and spread.

Specifically, we are concerned about harms to democratic expression.

They include misinformation and disinformation, lies, threats, slander, intimidation, bullying and humiliation, directed at politicians, public officials, organizations and interest groups, journalists, ethnic communities and members of the public.

Many of these harms are not new, but the scale of social media platforms and their ability to amplify messages increases the scale of the harms and the risks they pose to democratic expression.

There are legal remedies for some of these harms that involve violations of privacy and the potential for claims of civil liability for defamation. Other harms violate the Criminal Code or human rights legislation, everything from incitement to violence, death threats and to racist assaults aimed at minority communities. All these have become increasingly common and frequent in recent years.

Legal action can and should be taken against those who break the law. Much can and should be done with the current legal tools. However, many of the harms are not illegal, making a legal response difficult or impossible. For example, radicalization is a consequence of the harms to democratic expression produced when the sharing of conspiracy theories on social media amplifies the message and thus can create a false sense of consensus that attracts a community of believers and supporters. Such activity can be 'lawful but awful.' It undermines democratic expression and threatens democracy in the following ways:

The Harms	Examples
Greater toxicity in politics and difficulty in persuading people to seek public office	Misogynistic, slanderous and racist posts that threaten politicians, journalists, public officials
A rise in hate speech, threats to social stability and attacks aimed at minority communities	Posts that blame equity-seeking communities for public health crises
Stresses on our democratic system and attempts to undermine the legitimacy of institutions	Questioning election results with no evidence
Misleading and false advertising	Lies and distortions about government and political party policies and statements by politicians
Election misinformation and disinformation	Attempts to suppress voting with disinformation about polling locations, hours, etc.
Health disinformation and misinformation	Circulating lies about vaccines, vaccinations and other health measures while promoting "remedies" for financial gain
Attempts to block access to public events and institutions	Encouraging people to blockade health care facilities or prevent public political events
Conspiracy theories	Claims about sources of disease, the motives behind specific government policies
Growing coarseness and lack of civility in public dialogue	Extreme partisanship and attributing motives to opponents that inhibits policy discussion and consensus building

Regulating the role of social media platforms is undoubtedly complex given the scale and nature of the different harms they pose to democratic expression, human rights and the public good. Much of the global debate about the activities of social media platforms, harms they may cause and calls for regulation and oversight has focused on the outputs – what users see.

The Commission strongly believes the time has come to focus on the algorithmic systems that amplify content in addition to addressing outputs.

Our attention is focused on the systems developed and deployed by social media platforms and the incentives within those systems that can lead to harms and their amplification. These 'closed boxed systems' comprise at best a very opaque world for governments, regulators, social media users and the public.

One way of conceptualizing this system is as inputs, the closed box and the outputs.

- **The Inputs** – the material and personal data that users and advertisers post and provide to social media platforms as well an expansive range of online and offline data, as well as inferred variables about our lives, beliefs and preferences.
- **The 'Closed Box Systems'** – the automated processes designed and deployed by platform companies within which they manage, aggregate and distribute that material and data collected; and
- **The Outputs** – the content that users see – or don't see – posted on their pages and sites.

In Canada, the federal government has introduced several pieces of legislation, none of which became law before parliament was dissolved in August 2021 for a September election. Those legislative proposals largely concentrated on trying to minimize the harms that flow from outputs.

Similarly, the Commission's first report published in 2021 focused largely on the outputs. Specifically, it outlined an integrated program of six practical steps that rejected a policy of aggressive takedown of harmful content in favour of a citizen-centric approach that places responsibility for hateful and harmful content firmly on the users who post such material and the platforms that amplify their messages.

Inputs can be constrained by Criminal Code provisions on hate speech. (It is worth noting that the policies of mainstream platforms constrain what people can post on the platforms to a far greater extent than permissible under law, due to free expression rights.) False and misleading advertising can be addressed by

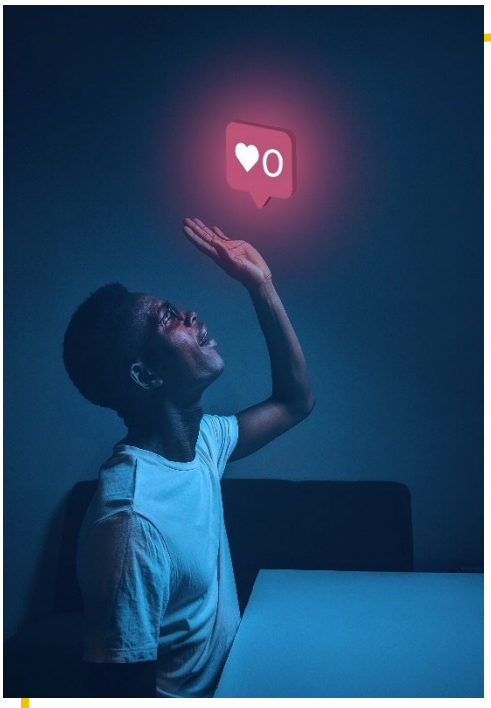
the Competition Bureau. Federal and provincial legislation, if it is enforced, may protect privacy rights of citizens, and constrain what social media users can do and post. To varying degrees these constraints also exist on outputs.

Opaque algorithmic systems have the least external oversight and constraints. That is why we believe governments would be wise to now focus their attention on understanding how social media platforms manage those systems – and enlist the assistance of regulators, researchers and users, and those affected by the content the platforms distribute. This represents an essential precondition for governments and regulators to make platforms accountable for eliminating harms to democratic expression and introducing measures to prevent the harms from happening.

Requiring transparency around how platforms program and manage the workings of their algorithmic systems is the first of three principles we believe are paramount in protecting against harms to democratic expression.

But transparency is a means to an end, not the end itself. Through transparency, governments, regulators, researchers, social media users and those affected by the content on social media platforms can assess the accountability of platforms for the spread of harmful material that threatens democratic expression.

To reduce the harms and to allow democratic expression to thrive, we believe transparency and accountability must be accompanied by a third principle – empowerment. Knowing how users are targeted, how their expression is made visible, amplified or silenced, and how their personal information is used without their informed consent or knowledge runs counter to both their basic rights and fundamental principles of democratic expression. The result is an imbalance of power in which users and those affected by what appears on platforms have little or no recourse or mechanisms to hold companies to account for potential harms. At worst, users faced with this overwhelming spectre of capture will simply not exercise their right to democratic expression. Empowerment helps shift greater power and control over their online and offline interactions into the hands of users and the public.



This report and our recommendations focus on how government and regulatory action on those three themes – transparency, accountability and empowerment – hold the key to preventing the harms to democratic expression we see flowing from the current unrestrained activities of social media platforms.

VALUES AND PRINCIPLES

The Commission's first report in 2021, *Harms Reduction*, identified a set of principles that guided its deliberations and recommendations, and these are noted in the executive summary of this report.

We remain committed to those principles and have added to them. At the core of all responses to address harms to democratic expression are our society's basic values. The internet now shapes almost everything in the lives of Canadians – how we work, how we spend and save, how we meet people, how we are entertained, how we use our spare time, how we relate to governments and other institutions and the ways we communicate.

But we believe that has not and should not change our society's underlying principles and values – the importance of privacy; respect for others and their points of view; protecting the vulnerable and young people; a recognition that we are a diverse, multicultural and multiracial society in which all have equal rights that must be protected; ensuring freedom of expression within the bounds placed upon it by Canada's laws and court decisions; and promoting an open and competitive economy that benefits all. These values and principles form the basis of Canadian society and remain important cornerstones of our democracy. They are integral to our deliberations and our recommendations.

There are steps that can be implemented in Canada to protect privacy rights and freedoms while countering harms and threats to democratic expression, keeping true to the principles the Commission outlined a year ago and consistent with the recommendations in our original report. But what we now propose will be more effective if Canada's actions are coordinated with the growing international consensus that has formed on steps rights-respecting democracies can jointly take in the interests of democratic expression.

All our nations face similar threats in the 21st century and Canada should adopt common responses in concert with other democracies, working together through multilateral and multistakeholder initiatives such as the Freedom Online Coalition, unless there is a clear reason and rationale for taking an alternate approach.

Notes on Freedom of speech in the Canadian context as it applies to regulation of social media platforms and the internet

The Right Honourable Beverley McLachlin, PC, CC

Freedom of expression enjoys constitutional protection in Canada. Section 2(b) of the Charter of Rights and Freedoms provides:

> 2. Everyone has the following fundamental freedoms:
>
> b. freedom of thought, belief, opinion and expression, including freedom of the press and other media of communication.

Purpose

The protection of freedom of expression is premised upon fundamental principles and values – the value of the search for and attainment of truth, participation in social and political decision-making and the opportunity for individual self-fulfillment through expression: *Irwin Toy Ltd. v. Quebec (Attorney-General)*, [1989] 1 S.C.R. 927 and 976.

Interpretation

The Courts have interpreted Section 2(b) broadly to apply to anything that has expressive content not removed by the method or location of the expression – i.e., expression that takes the form of violence or threats of violence: *Canadian Broadcasting Corp. v. Canada (Attorney-General),* 2011 SCC 2.

Physical violence is not protected, nor are threats of violence: *Irwin Toy, supra; Suresh v. Canada (Minister of Citizenship and Immigration),* [2002] 1 S.C.R. 3 at paragraphs 107-108. In other respects, the form or medium used to convey a message is considered part and parcel of the message and protected by Section 2(b): *Weisfeld* (F.C.A.) Otherwise harmful speech is protected – hate speech, child pornography and misinformation enjoy Section 2(b) protection.

The reference in the guarantee to "other media of expression" makes it clear that it applies to the internet. Online messages of all types (possibly with the exception of threats of violence) are presumptively protected by the constitutional guarantee of freedom of expression. The content of the expression does not remove the Section 2(b) protection; it covers even odious and hateful expression.

Limits under Section 1 of the Charter

The Charter guarantee of free expression is not absolute. The state may impose limits on free expression under Section 1 of the Charter, which provides that the rights guaranteed are "subject only to such reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society." In a democratic society, rights – including freedom of expression – must sometimes be limited to prevent harms to others. Most modern guarantees of rights follow this model and expressly recognize that freedom of speech can be limited in light of conflicting values and concerns.

Limits on expression may be imposed by statutes or arise from inherited common law wrongs, like defamation. The limits on speech imposed by governments are varied. Parliament has passed laws making certain kinds of speech crimes – for example, crimes against hate speech, pornography and sedition. Parliament and the legislatures may authorize agencies created by statute to take action against harmful speech; examples include federal and provincial human rights legislation and media constraints under the CRTC. Municipal noise and protest by-laws are yet other examples of legal limits on expression. Finally, individuals may bring civil suits to limit or provide reparation for harms recognized by law, like defamation or breach of local noise by-laws. Individuals may obtain court injunctions to prohibit unlawful speech.

Ultimately, independent courts – not the government – decide the limits of free expression. If challenged, laws restraining expression passed by Parliament, the legislatures and municipalities must be proved in court to be reasonable and justified in a free and democratic society. Actions taken and regulations made by governments and government agencies are subject to the same judicial scrutiny. Shutting down speech in Canada has generally required a court order or injunction. These protections ensure that limits on free expression will not exceed what is reasonable and justified in a free and democratic society.

We also believe we can learn much from Canada's Indigenous peoples' culture and traditions. Willie Ermine, Assistant Professor with the First Nations University of Canada in Regina and from Sturgeon Lake First Nation located in north central Saskatchewan, writes about the need to create "ethical spaces," which he describes as a neutral ground – where you can get out of your own rules and spaces – released from our allegiances and mental cages, it's an ethical space where human to human contact can occur.[7]

This involves infusing our communications with human values, issues about heart and soul, and democratic and higher values than the extreme comments that are too often spilled out on social media. It requires the engagement of everyone in that communication process – whether it is users or software developers and programmers who design algorithms to generate specific responses and the platforms that use them – acting in a common spirit, instilling ethical spaces and humanity into technology.

Members of the Citizens Assembly on Democratic Expression came to a similar conclusion, noting "ethical practice strives to cause no harm to individuals or the public. Harms can be physical, emotional, mental, political or financial. Ethical practice is a foundation for values such as accountability and transparency and includes the duty to act responsibly in every online environment."[8]

That starts, as the Commission noted in its first report, with a platform code of conduct that treats all users equally in their interactions with the platforms. It includes the impact on users of what they see and do and how the platforms shape and affect the relationships that users have with each other, with other communities and with Canadian society as a whole.

THE POWER IMBALANCE

How can individuals, communities and governments delineate and enforce that responsibility in the face of such an imbalance of power?

To this point, the Citizens Assembly's report stressed: "Social media platforms have gone unchecked. Profitability has prevailed over privacy and data ownership. Inadequate data protection policies and the lack of recognition of our data privacy rights have disempowered us. It is overdue that we take a proactive approach for the public good." [9]

That brings us to the core of our mandate – the opaque and increasingly automated systems through which platform providers shape public discourse without recourse for potential harms or adequate protections to mitigate potential risks.

As experts told us, this echoes the 'incredible' power imbalances between digital platforms, on the one hand, and users and independent researchers and journalists with a public interest orientation, on the other.

Platforms are responsible for establishing policies about what to permit on their sites. Then those standards are applied through programming to identify hatred, disinformation, conspiracies and other harmful written and video communications and images that may be subject to removal and sanctions applied against the posters.

Many of the controls exercised by the platforms on users, content and advertisers lie within the currently opaque systems, including the constantly adjusted algorithms used to categorize users and to match content and advertising to perceived user preferences. As experts repeatedly told the Commission, users have no idea what the platforms know about them and how platforms determine those preferences and have little ability to shape or alter that information and limited recourse to have hateful or offensive content removed.

We also heard from experts that users and researchers don't know what societal or cultural biases are inherent in the algorithms, or how or if platforms are compensating for that. Nor do they even know whether the platforms are aware of their own algorithmic biases that may distort the outputs provided to specific groups in society, based on such demographic criteria as religion, income, racial or ethnic origin, gender, education or minority status within a majority setting. They have testified that the automated systems designed to protect marginalized groups fail to do so and further censor their attempts to address discrimination online.

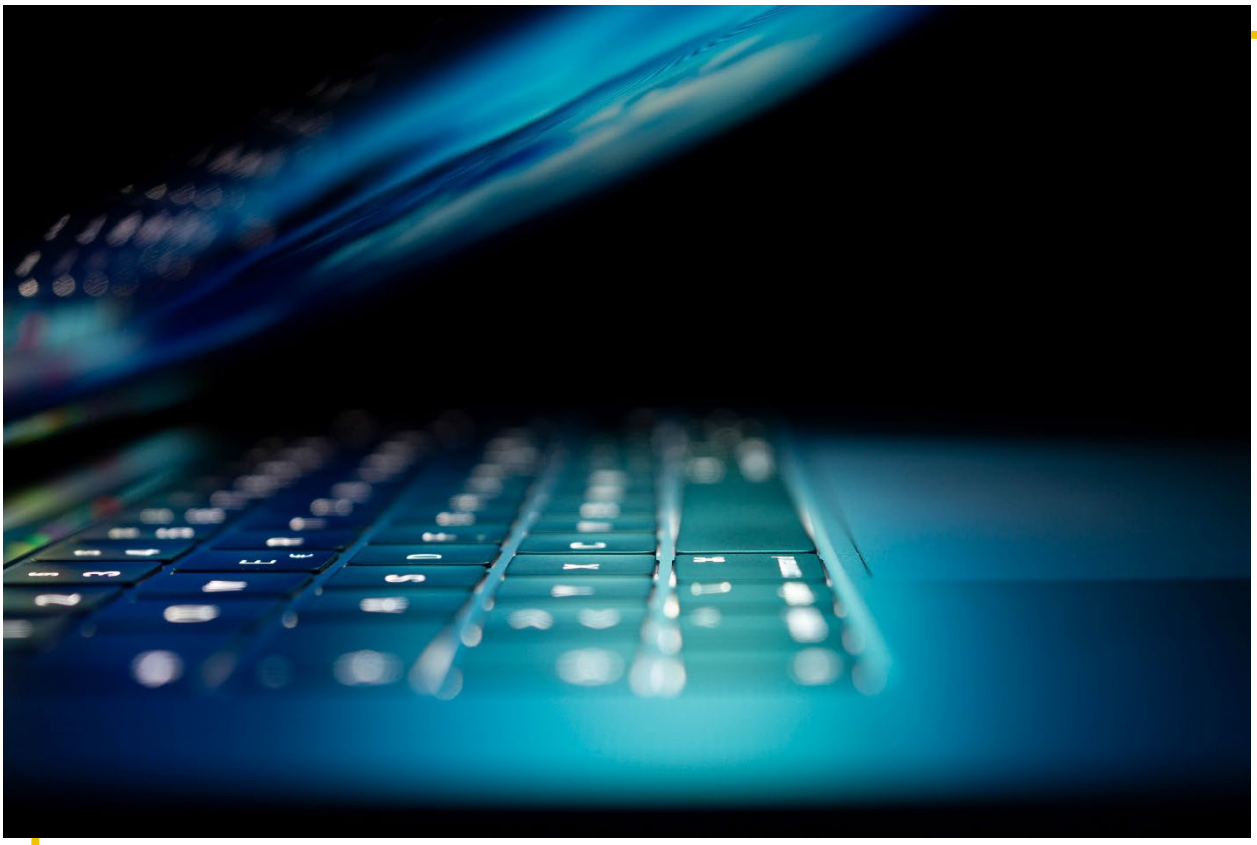
Based on these testimonies, we suggest three ways to establish a more balanced distribution of power between users, those affected by an algorithmic system and the platforms.

First, mandate greater transparency about otherwise closed automated systems so users, those affected by an algorithmic system, researchers, regulators and governments have a greater ability to assess and audit how algorithms are used. This could take a number of forms tried internationally, such as transparency registrars, audibility of platforms and algorithmic systems, and protected access to publics, civil society groups and independent parties evaluating these systems.

Second, use this increased transparency created by opening up the systems to determine and assess whether content promoted by platforms contributes to harms to democratic expression and whether and how platforms should be held accountable. Ask questions, pass judgment and enforce sanctions against platforms that host and amplify material that contributes to harms to democratic expression.

Third and most important, give users more power to determine, manage and control the information about them that is collected and used by platforms.

Our report's recommendations centre on how to accomplish these three objectives.



RECENT DEVELOPMENTS

Throughout the year since the Commission's first report, the extent of international scrutiny and proposed legislative remedies to address online harms and threats to democratic expression has grown considerably, with related extensive media coverage. This societal and public policy debate is evolving rapidly.

As the Aspen Institute in the United States highlighted in its November 2021 final report of its Commission on Information Disorder,[10] in describing the United States: "The past decade has been marked by a tremendous shift in the social, cultural and political fabric of American life. As we close in on the end of a second year of the COVID-19 pandemic, the seams are splitting and the threats to communities and to livelihoods have moved from internet chat rooms to the ICU. We see how our information ecosystem is failing the public, and how the absence or loss of trust in government entities, community institutions and journalism, combined with a growing number of bad actors and conflict entrepreneurs who exploit these weaknesses, have led to real harms, sometimes with fatal consequences. Public discourse is deeply polarized and acrimonious; we are distrustful of each other and of powerful institutions (sometimes for good reason). Many have become groundlessly skeptical towards scientific research and reject substantiated facts. Moreover, amongst the mandated lockdowns and abrupt shifts to online everything, this past year underscored how critical it is for us to connect with each other in true dialogue and meaningful discourse. Our growing incapacity to bridge these divides and make vital connections in our lives is having a corrosive effect."[11]

Canada is not the United States but there are similarities.

Our political culture and institutions, linguistic heritage, patterns of immigration, multicultural diversity, social and legal structure and media environments are different than the United States. In many ways they are substantially different.

We think it would be a mistake for Canada simply to adopt either the societal analysis or the remedies being proposed and debated in the United States.

Social media platforms were not just key for politicians trying to reach voters. Many Canadians used social media positively during the election to discuss and debate parties' policies and who they should support, to share videos and jokes and to provide running commentary during televised debates.

But for some, social media platforms have also become the conduit for spreading disinformation and conspiracy theories such as those alleging COVID-19 did not exist and lying about the risks and side effects of vaccines. Sharing that on social media platforms during the COVID-19 pandemic fuelled the formation of anti-vaccination groups, some of which harassed and threatened health care workers and tried to block public, employee and patient access to health care facilities and hospitals.

And as the past year in Canada has demonstrated, while they may not be as prevalent, many of the same underlying stresses on society and civilized discourse and disagreement identified by the Aspen Institute can be found here as well.

The responses are similar as well, such as the extremely well-organized and well-financed occupation of downtown Ottawa for more than three weeks in February 2022, by a combination of white supremacists, anti-vaccination conspiracy theorists, those opposing all pandemic measures and extremists promoting the overthrow of the federal government, using trucks as weapons. A similar example was the series of Canada-U.S. border truck blockades during the same month in Ontario, Manitoba, Alberta and British Columbia.

At all these, as in the 2021 federal election campaign, many prominent slogans, images and chants from demonstrators mimicked those of similar groups in the United States involved in the January 6, 2021, storming of the U.S. Capitol. A significant share of the funding for the protests came from the United States as well, but the protests were led by Canadians and backed by millions of dollars in Canadian funding, much of it in small amounts from individual donors.

Canadians must not minimize the importance of these, to this point, uncharacteristic confrontations. Our analysis and recommendations are rooted in our legal, cultural and social norms and history as well as the values and principles of our citizens.

On the public policy front, the past year has also seen a growing number of legislative and regulatory proposals that like this report, focus on transparency, accountability and user empowerment.

In the European Union, the Digital Services Act (DSA) is close to passage and the Digital Markets Act (DMA) may be passed in the spring of 2022. Together the two are designed "to create a safer digital space in which the fundamental rights of all users of digital services are protected; and to establish a level playing field to foster innovation, growth and competitiveness, both in the European Single Market and globally."[12]

For instance, the DSA includes a requirement for platforms to make their data accessible to researchers as well as disclosing all online advertising and targeting parameters.

Proposed legislation in the United States has a focus on algorithmic accountability including regulations that would require certain entities using personal information to conduct impact assessments and "reasonably address in a timely manner" any identified biases or security issues.

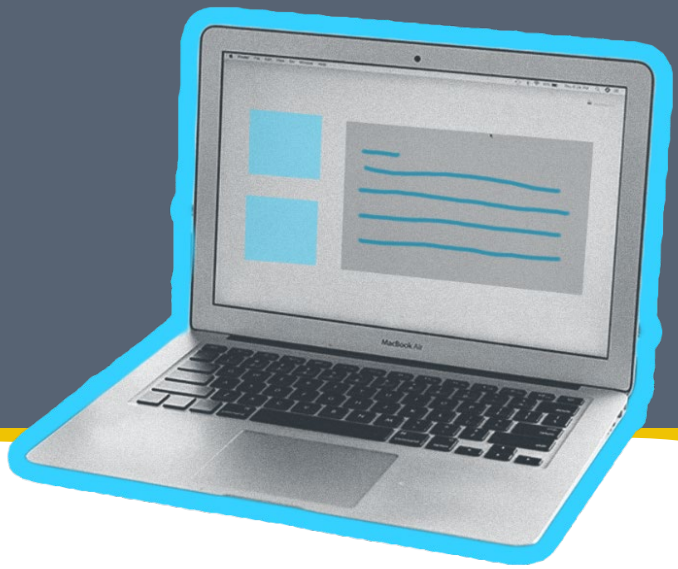
Promoting user empowerment, the EU's General Data Protection Regulation (GDPR) entitles Europeans to eight user rights: the right to information; the right of access; the right to rectification; the right to erasure; the right to restriction of processing; the right to data portability; the right to object; and the right to avoid automated decision making. National data protection authorities in the 27 EU Member States uphold these data protection rights to protect both individual interests and the public interest in ensuring that privacy laws are respected.

In Canada, the federal government has responded with several pieces of legislation, none of which became law before parliament was dissolved in August 2021 for a September election.

In November 2020, the federal government introduced two pieces of legislation addressing aspects of the activities of platforms and their impact on democratic expression: Bill C-10 – An Act to amend the Broadcasting Act and to make related and consequential amendments to other Acts, and Bill C-11 – An Act to enact the Consumer Privacy Protection Act and the Personal Information and Data Protection Tribunal Act and to make consequential and related amendments to other Acts. A third bill followed in June 2021: C-36 – An Act to amend the Criminal Code and the Canadian Human Rights Act and to make related amendments to another Act (hate propaganda, hate crimes and hate speech). In addition, the federal government released a technical paper in July 2021 outlining an approach to address harmful content online and sought public feedback on the proposals via a consultation process.[13]

Only Bill C-10 received extensive parliamentary scrutiny, but that didn't prevent a largely negative response to the government's package of proposals, even as the COVID-19 pandemic and a national mass vaccination campaign dominated the headlines in the late spring into the summer. Much of the criticism coalesced around perceived threats posed by the legislation to free speech and freedom of expression as well as individual privacy.

In February 2022, the federal government responded to its Online Harm Consultation by publishing a 'What We Heard' report, which outlines feedback from the consultation process and announced the creation of an Expert Panel to advise on how to adjust proposals to address concerns.



CHAPTER TWO: COUNTERING THE THREATS TO DEMOCRATIC EXPRESSION

RECOMMENDATIONS

Our recommendations are clustered under three interrelated themes – transparency, accountability and empowerment. The three work together to provide users, advertisers, governments, regulators and the public with the material to understand and counter how social media is used, and to help determine how platforms shape the contours of democratic expression that can end up resulting in online harms.

Each theme begins with a definition of the theme and some supplementary background information, followed by individual recommendations related to that theme. Each recommendation is accompanied by a brief explanation of its rationale and details of where similar recommendations are being proposed or implemented by governments and their regulators in other countries.

THEME ONE: TRANSPARENCY

Definition

Transparency refers to a variety of disclosure mechanisms through which platforms provide information about their operations, including automated decision-making systems. Usually this information is not public, which often makes it difficult for those outside the platforms to identify and mitigate discriminatory impacts. The aim of meaningful transparency – transparency that serves those most impacted by platform opacity – is to demonstrate compliance with legal and/or regulatory requirements and to bolster accountability mechanisms, including individuals' and regulators' ability to challenge hidden or automated decisions. In this

sense, meaningful transparency can partially serve efforts to identify unjust outcomes and harms, hold powerful actors publicly accountable and improve overall governance. Transparency can also provide users with the necessary tools to make informed decisions about their behaviour, both online and offline.

Context

Digital platforms use personal data and automated tools and algorithms that affect the public sphere. We heard testimony from experts that the personal data of users collected by platforms can be used to target other users and to enable and amplify divisive and harmful content, which causes real-world harms. Yet these platform processes remain hidden from the public and from independent actors working to mitigate discriminatory patterns. Opaque decision making about the public directly impedes democratic values and rights, including those of privacy, fairness and due process. While many platforms report on their activities with regards to specific content, this reporting is entirely self-regulated, often incomplete and cannot be independently verified. Platforms also largely resist releasing information on the type of data they collect or permitting independent experts to evaluate the performance of their algorithms. In part legitimately citing concerns around privacy, intellectual property, trade secrets and the potential for malicious use of whatever they make public. Researchers and journalists working in the public interest are repeatedly denied access to public content on platforms. The lack of transparency makes it nearly impossible for governments and regulators to verify if the companies are complying with the law.

Legislative proposals from Europe, the United States and elsewhere have attempted to strengthen or mandate transparency through regular reporting, user notifications, data access for public interest and public-facing audits of platform practices.[14] But transparency measures must also ensure individual privacy is retained. That includes the privacy of those whose data is collected and shared with third parties. That privacy is essential because such data must be sufficiently contextual and granular to identify broader trends and patterns of discrimination,[15] while protecting user privacy and preventing overreach by governments, corporations and researchers. Many experts and practitioners agree that greater transparency, in and of itself, is an insufficient accountability mechanism given that self-reported aggregate data rarely offer true insight into content practices. Just because data is made transparent does not mean that its collection cannot cause harm. We therefore also support the government's efforts to modernize our data privacy regime to ensure there are far greater protection of user data in Canada.

While transparency is only a partial solution, it remains a key mechanism for enhancing public understanding about the design, technical and financial processes that govern today's social media platforms.[16] Our recommendations in this area are designed to strengthen efforts to hold companies to account for societal harm and to equip the public, researchers, journalists and policymakers with tools to highlight and then address structural inequalities and their amplification online.

Recommendations

1.1. Mandate and enforcement: Create and empower a regulatory gatekeeper to mandate and enforce access to data contained within social media platforms for research and oversight purposes.

The regulatory gatekeeper would enforce mandatory platform data access and sharing requirements outlined in legislation for each of three tiers: the general public; accredited researchers and journalists; and more specialized and detailed research in the public interest that without significant additional safeguards could have privacy implications.

Entrusting a regulatory entity with the task of administering the release of platform data would ensure a privacy-protected, secure pathway to access data. The entity, perhaps operating under the auspices of the Tri-Council,[17] would guarantee that only qualified researchers and journalists with public interest have access to higher levels of data, thus safeguarding individuals' privacy and security. The entity would need to consider user privacy in all decisions to grant access and take appropriate measures to safeguard sensitive data from being disclosed. Its tasks would include administering data to researchers, approving requests and creating an advisory board (composed of industry and academic representatives). The entity would also have enforcement powers to ensure cooperation by both platforms and researchers. Only allowing vetted researchers, journalists and academics mid-to-high access helps safeguard data security and confidentiality.

Precedents

With the European Union's proposed Digital Services Act, national authorities ("Digital Services Coordinators") would be able to order platforms to provide data access to vetted researchers (Article 31). In the United States, the proposed Platform Transparency and Accountability Act,[18] suggests the creation of a "Platform Transparency and Accountability Division" within the Federal Trade Commission. The Division's main power would be to develop and establish recommended standards, criteria and approval process for researchers, research projects and platforms.

1.2. Implement tiered data access: Mandate separate tiers of data access with safeguards for the public, for researchers, journalists and civil society groups.[19]

Introduce graduated tiers of data access obligations for platforms to give the public lower-level access (tier 1); to give researchers, civil society actors and journalists mid-level access (tier 2); and to give a narrow range of specialized applicants conducting specialized public interest research access that requires more significant detailed safeguards to balance privacy protection with the need for greater platform transparency and accountability (tier 3).

Requiring digital platforms to better inform users about how they operate will allow individuals to make more informed choices about how they use social media.

Tier 1 ensures individual users from the general public can obtain the basic demographic and related data about themselves that the platforms use to determine what shared content and advertising is promoted by the platform to that user. Such data must be provided by the platform to a user based on an application by that user to a specific platform. In addition, tier 1 should provide public transparency into high-visibility and/or high-engagement public content – i.e., public posts by accounts with very high numbers of followers and public posts that receive high levels of views (impressions) and/or engagement.

Tier 2 mandates for data access should be tied to specific research/journalistic objectives (e.g., the performance of independent risk assessments, the evaluation of platforms' impacts in particular areas of policy focus, how online harms are propagated, etc.), but would allow the independent researchers and journalists to define what precise forms of data will be required to carry out the research in question. Emphasis should be placed on mandating data that informs and promotes long-term harm prevention and on providing data to enable the identification of broader structures and patterns of harm, especially discrimination and inequality. Tier 2 recipients will be required to comply with specific detailed ethical, security and privacy requirements. Tier 2 access would be controlled by a body, perhaps operating under the auspices of the Tri-Council.

Tier 3 applicants must meet more stringent requirements on data access and more detailed provisions for ethical use and privacy based on the nature of the data being used and the purpose for which it is obtained.

This recommendation allows researchers, journalists and civil society actors working in the public interest to identify potential harms before they occur and to hold platforms accountable for societal harms. Ensuring that independent researchers have access to platform data is necessary to address how harms are amplified by cross posting the same content on different platforms and to establish a response system to online risks.

Providing different degrees of data access ensures that recipients have the necessary tools and skills to understand the data provided. Data about problematic trends online would partially redress the difficulties interested parties face in truly understanding the ways different groups are disproportionately harmed online.

Precedents

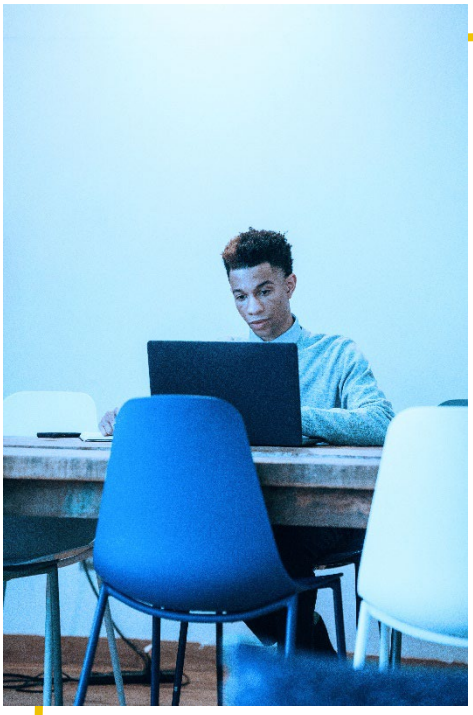
The European Union has announced a revised version of the "Code of Practice on Disinformation" including a "robust framework for access to data by researchers."[20] Article 31 of the proposed Digital Services Act (DSA) requires very large platforms (with an active user base of over 10 per cent of the European population) to provide access to data to a Digital Services Coordinator. The latter might also request the platform to provide access to vetted researchers for specific purposes.[21] Under the DSA, researchers would not be allowed to use data accessed for profit-seeking purposes or to inform political campaigns.[22]

In the United States, the proposed Platform Accountability and Transparency Act (PATA), would compel platforms to share data with "qualified researchers," defined as university-affiliated researchers and likely expanded to include civil society actors pursuing projects that have been approved by the National Science Foundation (NSF). Failing to provide data to a qualifying project would result in the platform losing the immunities provided by Section 230 of the Communications Decency Act.[23]

Alongside legislative initiatives targeting the technology sector specifically, tiered obligations have also been introduced in other areas. For example, in the proposed EU Directive on Corporate Sustainability Reporting, the EU Commission is set to establish sustainability reporting standards "proportionate to the capacities and characteristics of small and medium-sized undertakings" in contrast to those applying to large ones.[24] Similarly, in the United States, the Securities and Exchange Commission has imposed different disclosure requirements for smaller reporting companies, such as allowing a less extensive narrative disclosure.[25]

1.3. Mandate universal digital ad transparency.

Introduce a legal requirement for social media platforms to disclose regularly and archive, in a standardized format, specific information about every digital advertisement and paid content post on their platforms. Further, require that the parameters and categories used to target users be disclosed, including: the entity that paid for the advertisement; the advertising budget and overall amount spent; the intended and actual reach of the advertisement; information around personalization/micro-targeting and what voluntary codes of conduct advertiser endorses and follows. Disclosure and archiving shall be universal, meaning that platforms shall present data using a standardized machine-readable format with common minimum disclosure standards. All this data will be preserved in a single, central archive with its content available to the three tiers of users detailed in earlier recommendations. Legislators should also consider whether these provisions should be applied more broadly to other online sites beyond social media platforms given the ubiquity of targeted advertising online.



After considering the question of banning micro-targeting of advertising, the Commission concluded Canada would benefit from a more nuanced approach to the question and suggests that the Office of the Privacy Commission be given more authority to consider the question of micro-targeting.

The 2018 Elections Modernization Act[26] required social media platforms to establish an archive of all political advertising run on their site during a federal election campaign. It is now time to build on that base by extending and expanding this requirement to all advertising, all of the time, that appears on social media platforms. All advertising should be preserved in a standardized format in a single, central database available to the public.

Transparency about advertising and advertising data is essential for users, researchers/journalists and regulators to understand why and how individuals and groups are being targeted by a platform with specific advertisements and how users and groups are being profiled by platforms. This understanding will further empower users to adjust their online behaviour to better comprehend why specific content is being promoted to them. Increasing advertising transparency will also help advertisers. We heard expert testimony that advertisers are increasingly frustrated with the limited information they receive from platforms where their ads appear. For example, they would like more details from the platforms about the numbers of views of their ads, how automated bots are distorting audience numbers and assurances that their ads are not appearing on pages with content they find objectionable. Greater transparency will also enable advertisers to be held accountable when targeted advertisements are misleading or manipulating users. Advertising transparency will aid in the identification of discriminatory and biased advertising practices. On political ads, micro-targeting may enable political misinformation to spread quickly and severely impacts public debate and electoral outcomes.

Several platforms, including Twitter and Facebook, already have tools in place to allow businesses and advertisers to access some data but these tools are not shared with researchers and frequently are prohibitively expensive for researchers to use. Even advertisers would benefit from transparency that provides them with more data about how platforms manage their advertisements, details about views and clicks on the advertisement and the influence of bots rather than humans, than is currently made available to them.

Additionally, Facebook has provided users with a link “Why am I seeing this ad?” in the top right-hand corner on all advertisements on its news feed. Clicking on that link outlines the basic demographic data Facebook maintains about that individual user upon which the automated decision to display that ad to the

user is based. This is a good start to transparency and worth considering mandating for other platforms, but more specific data is required to assess the accountability of these systems in promoting online harms.

Precedents

The European Union's proposed Digital Service Act (Considerations n°52, 63, 66 (interoperability of advertisement repositories); Article 24; Article 30 demands that very large online platforms make available an ad repository through an application programming interface (API). Article 34(1)(e) sets an obligation on the Commission to develop a standardized disclosure format to ensure interoperability); Code of Practice on Disinformation II.D ("Signatories recognize that transparency should be ensured with a view to enabling users to understand why they have been targeted by a given political or issue-based advertisement"). Members of the EU Parliament have also called for a ban on targeting advertising based on sensitive data (religious beliefs, sexual orientation and racial or ethnic origin).

1.4. Introduce stronger protection for whistleblowers.

Canada needs stronger protections for whistleblowers – current or former employees who expose corporate malpractice including violations of the law, mismanagement, waste of funds, abuses of authority and dangers to health and safety. Given the economic, professional, or personal risks whistleblowers may face, any recommendations to protect individuals who report and expose internal corporate malpractice should guarantee against legal, economic and reputational retaliation by their employers. The federal government could draw from other sectors in advancing private whistleblower protection.

Whistleblower protection is central to encouraging public disclosure of harmful corporate decisions and practices (rather than exclusively technical data) as it increases legal certainty for potential whistleblowers. Whistleblowers bring public attention to an otherwise opaque issue, which can prompt policy action. In the field of data protection and cybersecurity, whistleblowers' reporting can prevent cybersecurity issues that might affect Canada's economic and social activities, as well as digital services.

In its January 2022 report, the Citizens' Assembly urged the federal government "to review and strengthen whistleblower protections to safeguard those who can demonstrate corporate actions intentionally contribute to the prevalence of disinformation."[27]

Precedents

At present, only the Canadian Revenue Agency Offshore Tax Informant Program (OTIP) and the Ontario Securities Commission Whistleblower Law protect such individuals in the private sector.[28]

EU Member States have implemented the EU Whistleblower Protection Directive. Countries such as Australia, New Zealand, Japan and the United Kingdom have employment laws protecting employees who blow the whistle in a work-related context. The United States has a number of federal laws under which

whistleblowers in the public and private sectors can receive monetary payouts if their employer is successfully prosecuted due to their disclosures.

THEME TWO: ACCOUNTABILITY

Definition

Accountability refers to the act of holding platform companies to account for their operations and business practices and their societal effects, particularly for harm to historically marginalized groups and those directly impacted by discriminatory practices. Platform practices can be evaluated in a variety of ways. Independent risk assessments on a system done in advance of implementation can assess potential impacts and performance while those done on a system's actual behaviour can determine the actual harms and threats posed to different communities and provide guidance on mitigation. Audits conducted by different parties can evaluate whether a platform company has met an objective or universal criteria.[29] Regulators, politicians, civil society representatives and the public can pressure platforms to comply with their commitments in different ways, including penalties for non-compliance, stricter regulatory measures, new laws, inquiries on misconduct and public advocacy campaigns.

Context

Accountability must begin with social media platforms telling users, advertisers, regulators, governments and the general public more about how they function.

That is an essential first step as we outlined under the theme of transparency. But simply knowing what is happening is not enough – if certain forms of data gathering, recycling and analysis compromise democratic expression, simply giving governments and researchers the ability to reproduce these effects is inadequate.

But accountable for what, to whom, and against what norms and standards? We believe the answer starts with a recommendation made in the Commission's first report in 2021. It called on the federal government to grant legislative authority to a new regulatory body that will "oversee and enforce" a new Duty to Act Responsibly on social media platforms.[30]

The recommendation added that the new body "must be constituted that is at arm's length from the government of the day yet bases judicially made findings on the rule of law and that are subject to a process of review," adding "such oversight and enforcement is imperative to mitigating harms from online content and ensuring accountability."

The report went on to note that "commercial" incentives of the digital economy and its global scale make it difficult to imagine how the current system of pure self-regulation could ever succeed. From a governance point of view, protection of individuals and identifiable groups properly falls to public authorities and institutions. By the nature of the medium, platforms will remain the first line of defence, but now answerable to official and legally sanctioned guardians of the public good.

That duty to act responsibly rejects the false premise that the sole duty borne by corporations is to maximize returns for their owners. The duty to act responsibly to be required of platforms must also include a recognition of the negative impact their activities can generate. Making money selling advertising, promulgating information, directing information to a particular segment of the population who may be vulnerable, are not done in a vacuum and are not without consequences. Our focus is on the harms to democratic expression that can result from these activities but as the Citizens' Assembly report notes, negative impacts and harms can be much broader and extend far beyond threats to democratic expression.

The duty to act responsibly is the benchmark against which the automated activities that occur within 'closed box systems' should be assessed. Thinking and acting carefully about potential negative impacts when designing algorithms to avoid bias and ensure harms are avoided; having oversight mechanisms to respond to concerns and complaints; platform consultations with complainants – all these can reduce the penalties imposed by a regulator should a platform be found liable for violating its duty to act responsibly. These and other tactics to prevent harms can all be codified within each individual platform's duty of responsibility related to its activities.

Recommendations

2.1 Increase the capacities of public agencies: Ensure existing regulators are properly empowered and equipped to operate in the 21st century digital world efficiently and effectively. In addition, create a new independent federal regulator (as noted above and proposed in the Commission's first report) vested with investigative, auditing and enforcement powers and responsibilities to ensure a new legislated duty to act responsibly is applied to platforms. Over time, the new regulator would also be entrusted with conducting systematic review of policies and regulations and suggesting proposals for reforms when needed.

The primary mandate of the new regulator is to oversee and enforce the duty to act responsibly. That duty is an integral element of all our recommendations directed at platforms. Cooperating with research efforts is part of that duty to act responsibly, as is conducting algorithmic impact assessments and human rights impact assessments as required (see recommendation 2.4). So too is compliance with the other transparency and accountability measures we propose throughout this report.

The regulator would focus on opaque platform systems and operations to promote and ensure transparency and accountability, investigate perceived harms, assess platform liability and determine and enforce remedies when platform liability is established. The regulator should be formally independent from the government, mainstream media and the platform. The Social Media Council, which the Commission proposed in its first report, with membership from platforms, civil society, citizens and other interested parties would be the forum for advice to the regulator on platform governance, policies and practices. This would ensure on the one hand, impartiality in decision-making and on the other, cooperation with all interested stakeholders. It is crucial for the regulator to be properly resourced and carefully designed, based on public consultations with a focus on fleshing out its scope and remit. The regulator should operate in a transparent and accountable fashion with legislated reporting to Parliament at regular intervals.

While the regulator would focus on the new capacities required to oversee and pass judgment on the extent and degree to which opaque and/or automated systems may cause harms, privacy issues would remain the responsibility of federal and provincial privacy commissioners. Questions about how competition policy might address closed systems would remain under the control of the Competition Bureau. The new regulator's mandate would include the approaches detailed in subsequent recommendations in this report relating to algorithmic and human rights impact assessments and algorithmic audits, platform compliance with the duty to act responsibly, the extent of platform liability for identified harms and the remedies imposed in response.

To ensure the regulator can fulfill these responsibilities, there is a duty on government to ensure all those involved in oversight and regulation of any aspect of platform activities are properly empowered and equipped. This includes the regulator we are proposing as well as those charged with enforcing competition policy and federal and provincial privacy legislation and regulations. It is not appropriate or acceptable that they do not have the required resources and authority to operate in the 21st century digital world. They must have:

- Sufficient qualified personnel who understand and can navigate through the digital world (such as data scientists, AI practitioners, social scientists, etc.).
- The financial flexibility to compete successfully with the private sector to hire the best talent.
- The flexibility and authority to communicate with each other producing greater inter-agency co-operation to address the need for coordinated responses to the fact that the cause of some harms may exceed the jurisdictional boundaries of any one regulator.

- The ability to impose remedies commensurate with the financial status of the platforms when an accountability assessment concludes with indetermination of platform liability.

This is the institutional context and environment that we believe are essential for both platforms and regulators to ensure our recommendations about oversight and accountability are effective in responding to harms.

Precedents

The Online Harms White Paper in the United Kingdom proposed providing an existing telecoms and media regulator (Ofcom) with new powers and responsibilities to oversee companies' compliance with their statutory duty of care towards their users. The UK regulator will verify whether companies are taking the necessary steps to prevent the spread of harm on their platforms and limit problematic content.[31]

Regulators have also been created in Australia (eSafety Commission) and proposed in the European Union (European Board for Digital Services). In Canada, a previous online harm proposal included the creation of the Digital Safety Commission, to provide recourse concerning specific items of content, oversee and investigate platforms' moderation systems and enable major administrative penalties to be levied against non-complying platforms. The proposal also included the creation of a recourse council and advisory council.

2.2. Mandate tiered obligations for different types of platforms and/or for services likely to be accessed by minors and adults.

All platforms regardless of size have a duty to act responsibly. However, obligations on individual platforms may differ based on the type and size of platforms according to their capability to comply with legislated requirements. There would also be differentiation between the obligations placed upon platforms based upon use by minors (under 18), and adults.

Smaller platforms do not have the same resources as bigger platforms. So, demanding the same requirements would, on one hand, pose too onerous an obligation on the former, and insufficient obligations on the latter. As the Commission noted in its first report "an identical regulatory regime put in place for huge global platforms could place an unmanageable burden on newer or smaller competitors in the space. Our purpose is not to inadvertently hobble competition or innovation. The Government and the regulator will need to take into consideration how to adjust for different levels of demands on different companies."[32] But small platforms can also be active spreaders of harmful content. However, the duty to act responsibly also must apply regardless of the platform's size. In other respects, though, matching obligations to the size of platforms would ensure that both large and small platforms have all the means to comply with the obligations imposed on them by regulators. It may also protect smaller or start-up platforms from being overwhelmed by regulations and collapsing before having an opportunity to establish themselves. Policies

applied equally to all platforms, regardless of size, might further consolidate the power of the biggest platforms, who have more means to manage the regulatory burden.[33]

Adapting requirements to the kind of platform would ensure that obligations are targeted to solving issues specific to that group of platforms. For example, regulations targeting content-related harms would apply to platforms that facilitate the sharing of user-generated content, but not to service-sharing platforms, while regulations targeting data collection and algorithmic audits would target all platforms.

Adapting platform obligations based on the age of its users ensures internationally established safeguards protect the best interest of children.

The United Nations Convention on the Rights of the Child (UNCRC) recognizes that children need special safeguards and care in all aspects of their life and requires that these should be guaranteed by appropriate legal protections. European data protection law reflects this and provides its own additional safeguards for children.

Precedents

On the broader issue of tiered obligations, the European Union's proposed Digital Services Act differentiates between platforms and very large platforms (those with an active user base of more than 10 per cent of the European population). For example, micro- or small enterprises are exempt from the obligation to provide transparency reports.

In the United States at a joint congressional hearing, Mark Zuckerberg of Meta supported a reform of Section 230 of the Communications Decency Act that would require platforms to have "adequate systems in place to address unlawful content" and that "definitions of an adequate system could be proportionate to platform size and set by a third party" – in other words, imposing different standards according to platforms' size.[34]

The United Kingdom's Online Safety Bill creates three categories of harm, each of which has different risk management requirements. The categories of harm are illegal content; services likely to be accessed by children; and content that is harmful (but not illegal) to adults. For each, companies must carry out risk assessments and adhere to "safety duties."[35]

2.3 Legislate intermediary liability protections and exceptions for platforms liability.

Clarifying liability for harmful consequences generated by platforms' algorithmic recommender systems and amplification tools would encourage platforms to better moderate user-generated content. Users should be allowed with care to express freely their views online (within the limits of what is permitted under Canadian law). At the same time, making platforms liable for all problematic user expression can result in over-removal of material and censorship. Nevertheless, the degree to which problematic content is amplified and the impact that recommender systems have on public opinion should be taken into consideration when addressing platforms' role in democratic societies. Imposing appropriate platform liability for how content is moderated would incentivize platforms to ensure that the impact of problematic content remains limited and that individuals maintain their autonomy.

Intermediary liability laws specify when a platform may be held legally liable for harms resulting from content posted by its users.

Canada has intermediary liability laws in place to deal with copyright, but it is the only G7 economy without such a law governing all other content responsibility questions. The handful of Canadian court cases that have considered the issue of whether platforms should be liable for content posted by their users have provided platforms with very limited liability protections. Yet developing such standards is important for protecting both free expression and innovation in the digital economy.

Canada needs an intermediary liability law. We believe the federal government should introduce legislation that incorporates intermediary liability protections that are consistent with Article 19.17 of the 2020 Canada-United States-Mexico Agreement (CUSMA) on trade between the three countries. Such legislation would clarify when platforms can be held liable for harms arising from content posted on the platform by users. Enacting such legislation in Canada could spur the development of new social platforms here in Canada that serve as an alternative to big platforms based in other countries.

Under such a law, platform users would remain responsible for the content they post and could be pursued through the legal system for any harms that they cause – such as injuring someone's reputation or violating their privacy.

Such legislation can and should leave open the possibility of platforms being held liable for violating their duty to act responsibly in relation to the algorithmic curation and amplification of certain kinds of content.

Under the new transparency and accountability regime we are proposing, the new regulator and members of the public would have the information they need to evaluate whether technology companies should be

pursued in court for violations of the duty to act responsibly (or of their other legal obligations). Courts can levy financial penalties for such legal violations, or issue injunctions to require companies to act more responsibly.

Parliament may consider enacting "safe harbour" provisions in such legislation that would allow technology companies to move to strike lawsuits based on alleged breaches of their duty to act responsibly, by demonstrating that they have acted in good faith to discharge this duty – such as by implementing the transparency and accountability measures we suggest in this report (for example, participating in algorithmic impact assessments or cooperating in good faith with independent researchers).

Precedents

The EU's E-Commerce Directive extends liability protections to online services when their activity is "of a mere technical, automatic and passive nature." The proposed Digital Services Act contains similar provisions but also adds a number of new obligations.

In the United States, section 230 of the Communications Decency Act provides an immunity to platforms for third-party content (except when the content at issue violates copyright, involves sexual exploitation of children or sex trafficking, or is in violation of federal criminal law). Section 512 of the Digital Millennium Copyright Act limits the liability of intermediaries if the platform responds expeditiously to remove, or disable, certain infringing material posted on its platform by another individual. The proposed U.S. Justice Against Malicious Algorithms Act would remove Section 230 immunity if an online platform knowingly or recklessly uses a personalized algorithm to recommend content to a user based on that personal information, and if that recommendation materially contributes to physical or severe emotional injury (for example hate speech or eating disorder content).[36]

2.4. Empower regulatory entities to develop and implement a rights-based algorithmic accountability framework which includes algorithmic impact assessments (AIAs), human rights impact assessments (HRIAs) and algorithmic audits.

Algorithmic systems are increasingly used as part of automated decision-making processes in both the private and public sectors, often without meaningful consent, privacy protection or recourse for those who stand to be most affected by their decisions. Relevant regulatory entities should be empowered to develop and implement a robust algorithmic accountability framework which centres rights-based approaches to algorithmic governance. Centring rights in accountability frameworks for automated decision-making systems follows international standards for addressing the heightened risks to safety and fundamental freedoms, such as the right to freedom from discrimination.[37] Rights-based approaches should be future-proof to the risks raised by AI systems and obligations placed on actors deploying high-risk AI systems should facilitate accountability to those directly impacted by them. Algorithmic impact assessments (AIAs), human rights impact assessments (HRIAs) and algorithmic audits should serve the ability to protect rights and ensure redress for those impacted by AI.

Platforms should be required to commit to respecting human rights and ensuring that any adverse impact that their services might have been adequately addressed.

Human rights impact assessments (HRIAs) have traditionally been used to evaluate the impact of business practices, public policies and technologies on human rights with the goal of anticipating compliance with human rights law and frameworks.[38] Therefore, mandating HRIAs would force platforms to undertake systematic and periodic examination of the impact their services might have on human rights prior to their implementation and how to mitigate such risks.[39] Organizations with human rights expertise would assess a platform's policies, practices, products and services to identify violations of human rights.[40]

Algorithmic impact assessments (AIAs) serve a similar purpose of understanding an algorithmic system's impacts ex ante (with some potential for ex post responses), but mainly encourage developers to identify and mitigate the potential risks of algorithmic systems by relying on a framework focused on algorithmic harms rather than human rights (e.g., bias issues, concerns around transparency and redressability of a system's impacts, environmental impacts of the system) and data ethics.[41] Given the continuous changes affecting the functioning of algorithms, AIAs should not only be conducted before their implementation but also on a regular basis, prior to engaging in new data processing. Rigorous risk assessments should determine, in high-risk or sensitive cases,[42] whether certain systems should be designed at all. AIAs serve to mitigate the risk of harm. This is especially pertinent to groups known to be disproportionately impacted by algorithmic systems in both private and public decision-making processes, including historically marginalized groups.

Proactive harm mitigation and prevention is particularly important for services likely to be accessed by children because of the special considerations offered to children given their developmental vulnerabilities and their status as early adopters of online services.

In addition to mandating HRIAs and AIAs, requiring developers (creators) and deployers (those who are procuring or putting the system into operation) to keep a record of the decision-making procedure during the designing process and the deployment of an algorithm is fundamental to determine whether it negatively impacts users. This information should be publicly presented in consistent, clear and accessible documentation, which should clarify an automated decision-making system's intended use and include a description of the policies and processes within which the algorithm is operating.

This documentation could be disclosed as part of algorithmic audit reports to be made publicly available. Overall, algorithmic audits are a method to systematically assess ex post whether algorithms and their

outputs are generating harms such as privacy violations, prevalence of hate speech, and whether its decisions are biased or discriminatory in nature. To be effective, the auditing methodology needs to be tailored to the specific technical architecture, affordances and features of different organizations.[43] Additionally, audits can serve different goals according to what entity is performing them. First- and second-party audits are performed upon initiative of the organization itself, respectively, by members of the organization or by a contracted third party. Both kinds enable proactive monitoring and evaluation by platforms of harm.[44] On the other hand, third-party audits are performed by an independent third-party auditor regardless of an organization's permission. These audits are particularly important since they enable truly independent and credible evaluation but are often impeded by lack of accessibility to the data that third parties need to conduct them.

Precedents

In the European Union, Data Protection Impact Assessments are already required in the GDPR, and other forms of risk assessment are being proposed in both the Digital Services Act (DSA) and in the Artificial Intelligence Regulation.[45] Additionally, the proposed DSA also mandates audits by independent third-party auditors with technical knowledge of algorithms and other expertise, as well as gives national authorities ("Digital Services Coordinators") and, in some circumstances, the European Commission, the power to conduct on-site inspections of these companies.[46]

In Canada, the Directive on Automated Decision-Making is a working example of an algorithmic impact assessment (AIA) process in operation. It uses a questionnaire model for AIAs which requires a question-and-answer format be completed prior to public sector deployment of an automated decision-making system.[47] Ontario[48] and Quebec[49] have already introduced algorithmic audits. Both algorithmic impact assessments and audits would align with recommendations made by the Office of the Privacy Commissioner of Canada.[50]

Abroad, the UK's Information Commissioner's Office recently developed an algorithm auditing framework, which focuses on governance and accountability, and AI-specific risks.[51] Algorithmic accountability has been also heavily discussed in proposed Bills in the United States, such as in the Algorithmic Accountability Act[52] and in the Algorithmic Justice and Online Transparency Act.[53]

With regards to human rights impact assessments, the leading example is found in the UN Guiding Principles on Human Rights and Businesses, which require that companies ensure respect for international human rights through a due diligence process aimed at identifying, preventing, mitigating and accounting for how they address their impacts on human rights.[54]

2.5. Develop a Code of Practice on Disinformation.

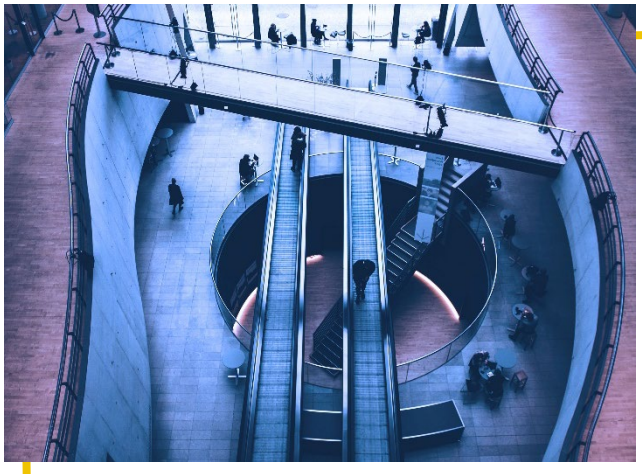
Canada should develop a Code of Practice on Disinformation in line with similar efforts underway in the European Union to establish commitments and requirements in collaboration with major online platforms. The overall aim of the Code is to promote the creation of platform-based policies and procedures to address disinformation, including demonetization of problematic content, increased transparency of political and issue-based advertising, empowering users to better control their online activities and enabling privacy-compliant access to data for fact-checking and research activities.

The Code would represent an efficient means of collaboration between public institutions and tech companies, especially considering that fighting disinformation must be a shared responsibility and goal. Soft law measures such as the Code of Practice on Disinformation are characterized by their flexibility and low pre-agreement transactions costs. Soft law mechanisms also facilitate systemic revisions to ensure that the provisions are constantly targeting contemporary societal issues.

Precedents

In the two years since its enactment, the EU Code of Practice on Disinformation has proven a valuable instrument to improve collaboration from signatory platforms and has provided a framework for a structured dialogue between relevant stakeholders to ensure greater transparency of platforms' policies against disinformation within the European Union. In the first months of its operation, the Code foresaw monthly discussions with its signatories to allow for an evaluation of its functioning and reciprocal feedback. Moreover, comparing reports over time, all signatories have shown improvements on all five commitments outlined in the Code, which could further reinforce the advantages of a collaborative tool such as the Code.

The EU Code of Conduct on Countering Illegal Hate Speech Online has served as an introductory approach to encourage platforms' efforts. Its transparency requirements have not on their own introduced accountability mechanisms, particularly because they focus on the rate and speed of content removals rather than an analysis of the type of content removed. The Code is deemed to have set the basis for stricter requirements being now imposed in the Digital Services Act however.[55]



THEME THREE: EMPOWERMENT

Definition

Empowerment refers to providing users of online platforms with the ability to manage their online presence, prioritizing the protection and mobilization of their democratic rights. It aims to redress societal power imbalances and structural inequalities that may be amplified by technological systems. As a broad framework, empowerment can be comprised of user rights (e.g., data protection, access to information, freedom of expression, etc.), platform obligations (e.g., privacy-by-design, user control tools, meaningful transparency, content moderation), regulatory measures (e.g., strong enforcement mechanisms) and public programs (e.g., civic, and digital literacy initiatives for the public to understand the choices they have online).

Context

Respect for and protection of human rights is fundamental to democratic expression to allow individuals and groups to participate freely and safely in our society without those rights being threatened or curtailed. Respect and protection of those rights are also key principles built into all our recommendations throughout this report.

Platform users are rights holders under Canadian and international law and our recommendations are designed to empower individuals to enjoy and exercise their rights while also respecting the rights of others.

In Canada, those human rights include both legislated protections and issues that touch upon the interactions between users and social media platforms such as the right to privacy, the right of ownership and control of one's own identity and the right to freedom from hate speech, abuse and harassment and the harms that come from that.

Platform activities must comply with Canadian law and international human rights standards, but beyond that we believe additional steps must be taken to address the imbalance of power mentioned earlier in our report between the platforms and the rights that users have to ensure their human rights are not threatened or violated. That starts with empowering users to decide how much control they want to exercise over the demographic information and data they provide to the platforms and how those inputs into their systems are used. Such decisions can only be made with informed consent and reasonable understanding around

opaque systems, and with freedom to leave a particular service without threat of losing existing connections.

It also involves the duty we recommend be placed on platforms to act responsibly. That includes knowing, understanding and ensuring that what is taking place inside the automated and otherwise closed systems advances human rights. Users must have the information and ability to determine whether their rights are not being protected and if so, to complain that a platform has breached its duty to act responsibly.

That duty on platforms to act responsibly also must include specific provisions for the protection of the rights of children, universally recognized as a vulnerable population, from online harassment, threats and bullying – in some cases even from other young people – which can be particularly damaging to children and teenagers. They are also a large and growing share of social media platform users. Their privacy online needs special protection perhaps through legislation. As well, children's data should neither be collected nor retained by social media platforms.

This gap in privacy legislation respecting children speaks to a broader need to strengthen and modernize all privacy legislation at both the federal and provincial levels. Much of it is out of date and needs to be overhauled for the digital age and the new challenges that has created, with a specific eye to social media platforms. Without such an overhaul, we will face a possible "chilling effect" on democratic expression, as individuals become more aware of voracious data capture and analysis but have no way to redress this.[56] Privacy officials have made recommendations in the past. Now, is the time for governments to act and pass much needed contemporary privacy legislation.

Recommendations

3.1 Support Indigenous knowledge, relationships and protocol development and Indigenous data governance for Indigenous communities.

Support meaningful participation of Indigenous peoples, and ensure Indigenous relationships and protocols are built into the development of technological and social policies, tools and mechanisms. This includes: grounding research in Indigenous epistemologies developed by and with Indigenous peoples and communities; training Indigenous data scientists and technologists; allocating seats within committees, oversight boards and other governing and regulatory bodies; ensuring that community-specific Indigenous values are built into the fundamental protocols governing how AI is developed and deployed by Indigenous peoples and communities; and prioritizing funding to support Indigenous epistemologies online. Initiatives aimed at safeguarding Indigenous data governance must also be supported. The federal government should collaborate with Indigenous peoples, communities and organizations to ensure that Indigenous data governance rights are respected and that Indigenous peoples have the means to pursue self-defined agendas. Additional support should include funding, the creation of new legislative proposals, literacy programs around data ownership and self-determination, and other needs identified in partnership with Indigenous peoples and communities.

The internet and digital technologies in general can play a role in conveying to the broader public significant knowledge of history and treaties, tactics of ongoing colonization and assimilation, and the ongoing quest to live with dignity as Indigenous people in Canada.

Ensuring meaningful participation of Indigenous peoples' representatives will safeguard Indigenous peoples' interests in addressing online democratic issues and inform broader collective values of reciprocity and self-determination.

Democratic expression via Indigenous knowledge education is also an answer to online hate, disinformation and ongoing racism targeting Indigenous communities, who represent some of Canada's most vulnerable populations. The official and substantial support for Indigenous knowledge, relationships and protocol development will address common problematic practices such as the non-citation of Indigenous authors in articles reviewing Indigenous topics, and the token participation of Indigenous academics in grant proposals involving Indigenous research.[57]

International consortiums on Indigenous AI protocols are developing AI tools and conceptual and governing approaches for AI development which centre on Indigenous peoples' experience and relationships with AI.

Concerning governance, data is a cultural, strategic and economic asset for Indigenous peoples. Guaranteeing that Indigenous people have and maintain governance over Indigenous data will ensure that information is used to fulfill Indigenous development agendas. Data governance can also be a strategic tool of decolonization. Internally, data governance promotes community self-awareness and can support projects regarding decolonization and self-determination. Externally, data governance ensures the communities are well placed to engage with the colonial entities and force the latter to rethink conventional research methodology and the associated defective data.[58]

Precedents

The Indigenous Protocol and Artificial Intelligence Working Group develops new conceptual and practical approaches to building the next generation of AI systems.[59] The International Wakashan AI Consortium is developing AI tools to represent and protect languages spoken among several First Nations communities.[60]

Canada's recently introduced Bill C-11 includes a requirement on broadcasting services to provide opportunities to Indigenous persons, and programming that reflects Indigenous cultures and is in Indigenous languages and accessible to all.[61]

In Australia, the government funds the "Indigenous Knowledge IP Hub" which provides a space for people interested in working with Indigenous knowledge.[62]

3.2 Substantially strengthen civic education respecting rights, digital literacy and access to quality information to support equity-seeking groups and community-led programs.

Public education and digital literacy initiatives should provide the public with an understanding of their rights and freedoms, how digital media works, how it can impact public opinion, and how structural biases operate within it and reinforces inequities in real life. This includes equipping citizens with skills to identify biases and assess the reliability of information, how to search, navigate, synthesize and evaluate content online, and how to meaningfully participate in communities online. Although programs should be accessible and focused on improving digital literacy throughout the entire population, including harder-to-reach communities, there are certain groups that should be targeted to minimize the effect of specific online and structural harms. For example, parents, guardians and children should be taught the risks involved with children's access to online services in addition to children's digital literacy. Underrepresented groups should be supported through targeted policies and programs that strengthen equity, including funding for the digital production of Indigenous cultures and knowledge. Programs should be offered in multiple languages including Indigenous languages.

Digital and civic initiatives should be implemented in tandem with measures to support high-quality information online, such as incentivizing quality information and harm reduction (e.g., strengthening fact-checking networks and research on authenticity online), investing in community journalism and quality journalism, and supporting certain forms of information creation and sharing (e.g., community media, local media, traditional media and digital-first outlets).

Civic literacy benefits both individual citizens and society more broadly and must now also encompass digital literacy to reflect civic engagement in the digital age. It encourages citizens to vote, promotes awareness of one's own political interests and how to advance them, diminishes the likelihood of being manipulated by negative and polarized political campaigns and improves overall community building.[63]

Digital literacy, which includes a wide variety of social and reflective practices that are embedded in work, learning, leisure and daily life, can preempt many online harms, including the spread of online illegal/harmful content.

Children are particularly vulnerable to digital harms, and research has shown that digital literacy education at a young age, results in greater resilience for both young people and their circle of caregivers.[64] Including racial literacy within digital literacy addresses the inseparable racial implications of technology.[65] Including racial literacy in disciplinary curricula will also aid those involved in the design, development and deployment of technology to consider racial issues.[66]

Strengthening public support for marginalized and equity seeking groups encourages proactive measures and community resilience in addressing online harms.

Research has shown that access to more Indigenous knowledge education online would not only increase the viability of such knowledge but would facilitate better relationships between settlers and Indigenous communities.[67] Reliable quality information shared through a robust journalism and media ecosystem is essential to empowering the public to make informed decisions and ensures independent oversight of powerful actors.

As well, the January 2022 final report of the 2021-22 Citizens' Assembly on Democratic Expression presented a series of recommendations on public education and awareness including one that the federal government create and fund a Centre for the Preventing of Disinformation that would play a major role in educating Canadians about all aspects of disinformation.[68]

3.3 Mandate interoperability and data mobility.

Information systems should be able to regularly interact and exchange information with one another, which would allow alternatives such as start-ups and platform co-ops to connect with existing services. Canada should ensure the interoperability of digital services to empower individuals with greater choice and control over their interactions online. Additionally, Canada should introduce the right to data portability – giving individuals the right to have their personal data transmitted directly from one platform to another, without hindrance.

Guaranteeing the interoperability of information systems by mandating the existence of an infrastructure that can enable information systems to communicate with each other and exchange information is necessary to empower users. Interoperability may further competition and innovation by allowing alternatives such as start-ups and platform co-ops to connect with existing services, thus hindering dominant platform corporations' tactics of leveraging one service to its benefit by "locking in" users. The main benefit of interoperability is that it can provide users more agency over their data and empower unsatisfied users to leave information systems while retaining connections to other users, including families, communities and customers. While interoperability can introduce new risks to user privacy and data security, if adequately developed, interoperability can be a net gain for user privacy rights.[69]

Interoperability has also been addressed at the level of antitrust solutions and regulatory controls. The proposed ACCESS Act in the United States requires applicable platforms to not make changes to their interoperability interfaces without approval from the Federal Trade Commission.[70] The European Union has

also expressed the need to strengthen interoperability in the proposed Digital Markets Act which would require full interoperability in core and ancillary services.[71]

Mandating interoperability will enhance the efficacy of the right to data portability, which provides users with a choice and control over their data in both the democratic and the market sense.

Data portability can enhance privacy through increased user control. Indeed, whenever users can decide to transfer their data to another platform automatically and without additional burdens, platforms will be forced to compete among each other to provide robust safeguards for data processing to attract users.

The European Union's General Data Protection Regulation codifies the right to data portability in Article 20. Analogously, data portability is also safeguarded in the Brazilian Data Protection Law and in the California Consumer Privacy Act. In the United States, the right to data portability was proposed in the draft of the ACCESS Act.

3.4 Modernize Canada's Privacy Legislation.

Privacy protections are fundamental to human rights and democratic expression. Systematic data collection and targeting online not only interfere with democratic expression, they can also threaten human rights and civil liberties and disempower users. Canada should update its privacy legislation to a rights-based framework for current and future technological developments. The Privacy Commissioner of Canada should be given greater authority to modernize Canada's current privacy legislative framework and decide how private platform companies can collect, process and target individuals' data.

Canada's current privacy regime for the private sector is 20 years old and does not reflect the significant technological, social and legal changes that have occurred over the last two decades.

During their testimonies to the Commission, Canadian and international experts alike encouraged Canada to modernize its current privacy legislation to a rights-based framework. In parallel, the 2021-22 Citizens' Assembly on Democratic Expression convened several experts who also advanced strengthening privacy laws to curb the spread of disinformation and prohibit the use of personal data for micro-targeting. The Citizens' Assembly recognized that, by strengthening user privacy and individuals' right to control who uses and accesses their data, vulnerable and marginalized groups would be further empowered to hold platform companies accountable for undue online harm.[72]

Without a robust rights-based privacy regime, Canada cannot adequately protect the democratic expression rights of children, Indigenous peoples and other marginalized groups.

Increased public awareness of data capture and targeting online without adequate means of redress may additionally pose a "chilling effect" on democratic expression by limiting individuals – especially vulnerable groups – from participating online.

In modernizing Canada's privacy legislation, the Privacy Commissioner of Canada should be given greater authority over making decisions on the practice of micro-targeting – advertising messages aimed at specific communities – particularly the importance of gaining insight into who saw an ad based on certain criteria of interest rather than what categories were used to target the ad in the first place. While some experts who briefed the Commission supported proposals made in other jurisdictions to ban micro-targeting, the Commission believes Canada would benefit from a more nuanced approach to the question of micro-targeting than a ban and suggests that the Office of the Privacy Commission be given more authority to decide how to regulate micro-targeting.

To date, efforts to protect and promote democratic expression and to protect privacy have been treated as separate issues in Canada. Empowering privacy rights of individuals is a primary lever through which central questions surrounding democratic expression can be addressed, including transparency and accountability of what and to whom. As such, the Commission finds privacy protections to be central to, and fundamentally interconnected with, protecting and empowering democratic expression in Canada.

Precedents

In the European Union, the General Data Protection Regulation has been a landmark legislation. It imposes stricter obligations onto organizations that target or collect data related to individuals in the EU, by limiting what data may be collected, how it may be used and under what circumstances.[73]

Similarly, the California Consumer Privacy Act has enhanced individuals' right to know information collected about them, to delete such information, to opt out of the sale of their personal data and to non-discrimination for exercising these rights.[74]

Both legislations have guided reforms in other U.S. states such as Virginia and Colorado, and around the world, such as in Brazil, Japan, Uruguay, Nigeria, South Africa and South Korea. All these reflect most obligations present in the GDPR on how organizations can collect and process data.

CONCLUSION

We don't pretend to know what may emerge in the coming years. Technology innovation is fast-paced and technology forms evolve rapidly. But we do believe that the fundamental rights, principles and values that have long formed the basis of Canadian society remain as relevant today and into the future as they have in the past.

Our purpose is to advocate on behalf of the public interest in an era in which the giant digital platforms have come to dominate not just distribution of information but public discussion around it. Our hope is that implementing our recommendations will start to rebuild the sense of confidence and safety that individuals, groups and communities must have to participate fully and equally in our democracy. Without this our democracy is at risk.

To prevent that happening, we all have a duty to act responsibly in all our communications and that applies equally to users and social media platforms. Both must always act with care ensuring that everything that is posted, shared and circulated goes beyond the minimum standard of not being illegal. As a Commission, we have consulted experts, researched and debated in a total of 15 study and deliberative sessions both online and in person to reach the conclusions and recommendations contained in this report that will allow public agencies to exercise oversight in a way we believe can best preserve public rights and freedoms of democratic expression.

APPENDICES

APPENDIX ONE

FREEDOM OF EXPRESSION IN THE CANADIAN CONTEXT

As noted in our report much of the debate about democratic expression and the platforms has centred on the potential for and risks of constraints on freedom of expression and free speech. Here, Canada is also different from the United States and those difference need to be understood.

For that reason, we believe it is worth providing considerable detail in this report to explain the context within which freedom of expression has been applied and interpreted by Canada's courts.

In the United States proposed constraints on free speech quickly turn into debates about the breadth of protection offered by its constitution's First Amendment.

THE DEBATE OVER FREEDOM OF EXPRESSION

The Right Honourable Beverley McLachlin PC, CC

Freedom of expression enjoys constitutional protection in Canada as well, but that protection is not absolute. Section 2(b) of the Charter of Rights and Freedoms provides:

2. Everyone has the following fundamental freedoms:

b. freedom of thought, belief, opinion and expression, including freedom of the press and other media of communication.

Purpose

The protection of freedom of expression is premised upon fundamental principles and values – the value of the search for and attainment of truth, participation in social and political decision-making and the opportunity for individual self-fulfillment through expression: *Irwin Toy Ltd. v. Quebec (Attorney-General),* [1989] 1 S.C.R. 927 and 976.

Interpretation

The Courts have interpreted Section 2(b) broadly to apply to anything that has expressive content not removed by the method or location of the expression – i.e., expression that takes the form of violence or threats of violence: *Canadian Broadcasting Corp. v. Canada (Attorney-General),* 2011 SCC 2.

Physical violence is not protected, nor are threats of violence: *Irwin Toy, supra; Suresh v. Canada (Minister of Citizenship and Immigration),* [2002] 1 S.C.R. 3 at paragraphs 107-108. In other respects, the form or medium used to convey a message is considered part and parcel of the message and protected by Section 2(b): *Weisfeld* (F.C.A.). Otherwise, harmful speech is protected – hate speech, child pornography and misinformation enjoy Section 2(b) protection.

The reference in the guarantee to "other media of expression" makes it clear that it applies to the internet. Online messages of all types (possibly except for threats of violence) are presumptively protected by the constitutional guarantee of freedom of expression. The content of the expression does not remove the Section 2(b) protection; it covers even odious and hateful expression.

Section 1 of the Charter

Notwithstanding the broad reach of the guarantee of freedom of expression, the state may impose limits on free expression under Section 1 of the Charter, which provides that the rights guaranteed are "subject only to such reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society."

Section 1 recognizes that the rights guaranteed by the Charter, including freedom of expression are not absolute. In a civilized society rights – including free expression – must sometimes be limited to prevent harms to others. Most modern guarantees of rights follow this model and expressly recognize that freedom of speech can be limited considering conflicting values and concerns. (By contrast, the U.S. Bill of Rights is cast in absolute language – "no person shall" abridge freedom of speech. However, even U.S. courts have upheld limits on freedom of expression to prevent harms. No one, to quote Justice Oliver Wendell Holmes, is permitted to falsely cry fire in a crowded theatre.)

Section 1 allows freedom of expression to be balanced against other rights protected by the Charter, like life, liberty and equality (protection from discrimination). To justify a limit the government must show a pressing objective. It must also show that the limit is no broader than necessary. At the final stage of the Section 1

(Oakes) analysis the government must show that the measure is proportionate in its impact; the question is whether the benefits of the measure that limits freedom of expression are proportionate to the harm that would be caused by permitting the expression without restriction.

Through the application of these criteria, the Charter permits the government to place "reasonable" limits on expression. Hate speech, obscenity, pornography and defamation are common categories of restricted speech in Canada. On the other hand, a vague and overbroad prohibition on spreading "false news" (misinformation) was struck down by the Supreme Court in *R. v. Zundel.*

Zundel reflects the fact that courts are acutely aware of the chilling effect that restricting speech may have on free expression. Laws restricting expression must be clear, concise and targeted in order to justify restricting expression. People must know with some degree of precision what they can say and what they can't.

Who Gets to Limit Expression and How Can This be Done?

Under Section 1 of the Charter any limits on freedom of expression must be imposed by law – that is by an enactment of Parliament or the provincial legislatures.

These laws take two forms. First, Parliament can criminalize certain types of speech, as it has for hate speech, pornography, conspiracy and terrorism. These provisions obviously apply to digital communications and social media. But they may be less than effective because enforcement requires criminal prosecution and entails all the delays inherent in criminal trials.

The second approach is for Parliament and the legislatures to set up regulatory schemes, like those providing for human rights commissions, that: (a) restrict certain types of speech; and (b) delegate enforcement to boards of appointed members. Similar schemes could be imposed with respect to speech on social media and other internet communications. But enforcement has also been difficult on this approach. The commissions have been much criticized for delay and their cumbersome processes and, by some, for interfering improperly with freedom of expression.

A third approach might be to appoint a regulator and delegate a large role to that regulator in defining when digital speech should be limited as well as administering the scheme. This might be challenged as unconstitutional, depending on how it is structured and set up.

The only "law" the regulator could make would be secondary law, or regulations. Generally, the power to make regulations must be closely tied to the primary law made by Parliament.

Parliament can't simply delegate its power to unelected officials in a way that leaves it up to them to decide what's allowed and what's prohibited. In other words, any restrictions on speech must be formulated by

Parliament with considerable precision in the law they pass; they can't just pass the duty on to someone else. Nor can the provincial legislatures do this. To do so would be unconstitutional.

This legal restriction on the scope of regulation mirrors the common sense concern one often hears about such proposals – it's wrong to arm an unelected official with broad powers to restrict freedom of expression. If anyone is to restrict this fundamental right, it should be elected representatives of the people.

A further wrinkle is that any law passed by Parliament would need to consider provincial powers over, for example, property and civil rights and the administration of justice. Such concerns, as well as the over-broad delegation of power to a regulator, led the Supreme Court of Canada to strike out much of Canada's reproductive rights scheme a few years ago.

So, who gets to impose restrictions on internet speech?

The answer is Parliament or the legislatures. They must define what's in and what's out. And who enforces Parliament's laws? This might well be some sort of commission or regulator. But the powers of this body would need to be closely tied to the primary law enacted by Parliament, and ways would have to be found to avoid the problems that have beset human rights commissions.

So, what is the best way to proceed in the Canadian context?

We favour starting modestly and growing the scheme over the years. If Parliament tries to do too much too fast the whole scheme may implode in an agony of court challenges.

A starting point might be the restrictions on speech crimes already established by Parliament – hate speech, pornography, terrorist speech, etc. Parliament could enact a scheme for enforcing these prohibitions in the internet setting, with a regulated body empowered to monitor and take appropriate enforcement action, like take-down orders. The body would also have the power to impose fine and other penalties after a hearing into the merits. It would also have the power to receive complaints, which could be mediated and, if mediation is not successful, adjudicated. To avoid the problems of delay we have encountered with the human rights tribunals, we would look to a new, streamlined tribunal along the lines of the British Columbia Civil Resolution Tribunal (CRT) which works well.

The foregoing framework could be accompanied by a council of citizens and experts to monitor and advise the commission and Parliament on changes.

Starting with categories of restricted speech that have been accepted by the courts will eliminate court challenges based on the content of the speech. If the accompanying scheme for enforcement and monitoring is well crafted, it would go some way to rein in abuses. It would also have the merit of providing clear guidance for platforms and users.

Conclusion

The answer to regulation of the internet is not easy or obvious, given the constitutional constraints that flow from the free expression guarantee in the Charter, the constitutional restrictions on delegation of powers to regulators, and the division of powers between the federal and provincial governments.

Harmful speech can be restricted in Canada – the law makes that clear. The bigger question is how that can successfully be done.

APPENDIX TWO

WHY WE FOCUS ON CLOSED SYSTEMS

The closed box has been at the centre of academic, policy and industry discussions about proprietary opaque decision-making systems and algorithms worldwide. Advanced machine-learning-based technology has brought many societal efficiencies, from economics to health, from behavioural trends to scientific discoveries. However, it has also negatively impacted democratic processes and human rights more broadly.[75] The opacity that characterizes some machine learning (ML) applications has been highly criticized, accused and found guilty of increasing and perpetuating inequality and biased decision making.[76] Today, scholars and policymakers are trying to avoid harmful impacts of ML applications by focusing on alternative ways to audit the algorithms that govern some of the most fundamental societal decisions, such as access to financing, education, employment and even policing.[77] To do so, understanding what makes machine learning (ML) and the closed box so complex to explain is essential.

ML models are the result of a training process instructed by the programmers based on a given dataset. There are several ways in which a model could output unfair and biased decisions. While discussing the technical aspects is out of the scope of this memo, it is important to remark that ML-based decisions risk reflecting human biases and prejudices, either consciously or unconsciously.[78] To complicate this further, receiving an explanation of the model that was used by the algorithm to reach a certain decision (i.e., the logic of the closed box) is particularly difficult. That is because, the most advanced kinds of ML applications, such as deep learning and artificial neural networks, use big data to discern patterns and make decisions in ways that do not necessarily follow human intuitive logics.[79] Doing so, closed boxes risk using personal data in distortive ways. It must also be mentioned that trade secret protections further prevent companies from making their algorithms more transparent to the extent possible.[80]

In light of this complexity, the phenomenon has been compared to a closed box, of which "we can observe its inputs and outputs, but we cannot tell how one becomes the other."[81] The closed box opaquely generates new knowledge about individuals and society.[82] However, the opacity in which such knowledge is obtained is problematic from a variety of perspectives, including ethics, accountability principles,[83] finance,[84] health,[85] industrial liability[86] and scientific research.[87]

Further, the closed box is particularly relevant when addressing online harm. Deep learning applications are, nowadays, widely used across social media platforms and search engines to analyze users' behaviour and offer ad hoc targeted services and recommendations. Recent events[88] have shown the impact of social media platforms on democratic societies worldwide through the spread of misinformation, increased polarization and radicalization, and the toxicity enabled by anonymous online environments. Recommender systems have proven to severely affect the quantity and quality of information individuals have access to, thus shaping their opinions and their interpersonal decisions.[89] When the algorithms that govern these

systems cannot be explained, freedom of speech, access to information, politics and the rule of law are undermined.

At present, the most promising approach to open the closed box is through "counterfactual explanations." These explain how a decision has been reached by pointing to which characteristics of the input data would need to change to reach a different decision, thus indicating which features influenced it.[90] Although counterfactual explanations do not require an understanding of the internal process of the algorithms and they overcome the issue of excessive disclosure of information potentially infringing IP and privacy rights, counterfactual explanations still present a major limitation. They provide all possible changes that would lead to a different outcome. In this way, the affected individual still has no means of knowing exactly which parameter has been decisive and whether that parameter was biased. Therefore, it is important to devote the necessary resources to further study the closed box and use this knowledge to strengthen democracies rather than weaken them.

APPENDIX THREE

COMMISSIONER BIOGRAPHIES

Rick Anderson

Principal, Earnscliffe Strategy Group

Rick Anderson brings decades of senior-level experience in business and government to Earnscliffe, with a focus on providing strategic advice and counsel on corporate strategy and public issues management.

Rick works with senior executives in the world's largest and most successful organizations, as well as assisting early-stage and high-growth entrepreneurs. He has deep experience working with C-Suite leaders, and familiarity with public policy, governance, political and regulatory affairs, mergers and acquisitions, and communications and marketing.

Prior to joining Earnscliffe, Rick spent 15 years in Canada, the United States and the United Kingdom with a foremost strategic communications firm and ran his own professional consulting practice. He currently divides his time between Vancouver and Ottawa, working out of Earnscliffe's offices in both cities.

Highly active in politics and public affairs commentary throughout his life, Rick has served in senior advisory positions to prime ministers, party leaders and leadership candidates. He is a frequent political affairs commentator on Canada's leading news organizations.

Wendy Chun

Canada 150 Research Chair in New Media, Simon Fraser University

Wendy Hui Kyong Chun is Simon Fraser University's Canada 150 Research Chair in New Media and leads the Digital Democracies Institute. She is the author of several works including *Discriminating Data* (forthcoming from MIT 2021), as well as three other books from MIT: *Updating to Remain the Same: Habitual New Media*, *Programmed Visions: Software and Memory* and *Control and Freedom: Power and Paranoia in the Age of Fiber Optics*.

Wendy has been Professor and Chair of the Department of Modern Culture and Media at Brown University, where she worked for almost two decades. She has held numerous visiting chairs and fellowships, from institutions such as Harvard, the Annenberg School at the University of Pennsylvania, the Institute for Advanced Study (Princeton), the Guggenheim, ACLS and American Academy of Berlin.

Nathalie Des Rosiers

Principal, Massey College, Full Professor, Faculty of Law (Common Law) University of Ottawa, Distinguished Visitor, Faculty of Law, University of Toronto

Nathalie Des Rosiers is the Principal of Massey College. From 2016 to 2019, she was MPP representing the riding of Ottawa-Vanier. She was Minister of Natural Resources and Forestry from January to June 2018. Prior to politics, she had been the Dean of Law, Common Law, University of Ottawa (2013-2016), General Counsel of the Canadian Civil Liberties Association (2009-2013), Vice-President, Governance, University of Ottawa (2008-2009), Dean of Law, Civil Law (2004-2008) and President of the LAW Commission of Canada (2000-2004).

With Peter Oliver and Patrick Macklem, Nathalie co-edited the *Oxford Handbook of Canadian Constitutional Law* (2017). She also wrote with Louise Langevin and Marie-Pier Nadeau, *L'indemnisation des victimes de violence sexuelle et conjugale* (Prix Walter Owen, 2014). She has received the Order of Canada, the Order of Ontario, honourary doctorates from Université UCL (Belgium) and the Law Society of Ontario, le Prix Christine Tourigny (Barreau du Québec) and is a Fellow of the Royal Society of Canada.

Amira Elghawaby

Director of Programming and Outreach, Canadian Race Relations Foundation

Amira Elghawaby is a journalist and human rights advocate.

She currently serves as the Director of Programming and Outreach at the Canadian Race Relations Foundation.

Previously, Amira worked in Canada's labour movement and additionally spent five years promoting the civil liberties of Canadian Muslims at the National Council of Canadian Muslims between 2012 and 2017. She has supported several national initiatives to counter hate and to promote inclusion, including as founding board member of the Canadian Anti-Hate Network and past board member at the Silk Road Institute.

Amira obtained an honours degree in Journalism and Law from Carleton University in 2001.

Merelda Fiddler-Potter

Vanier Scholar, PhD Candidate and Executive in Residence, Johnson Shoyama Graduate School of Public Policy

Merelda Fiddler-Potter is currently a PHD candidate at the Johnson Shoyama Graduate School of Public Policy in Regina. Awarded a Vanier Canada Graduate Scholarship in 2019, her research explores the media's role in helping Canadians learn the truth of our colonial policies, the impact it has on Indigenous people and how the media can keep Indigenous issues high on the public agenda.

Merelda Fiddler-Potter is a former journalist and documentary filmmaker, who spent 16 years working for the Canadian Broadcasting Corporation (CBC) in radio, television and online. She also launched her own documentary film company, making numerous films for national Canadian broadcasters. Merelda has a Master of Arts in Canadian Plains Studies and a Bachelor of Journalism and Communications, both from the University of Regina.

In addition to her doctoral studies, Merelda is a sessional lecturer at First Nations University of Canada, where she teaches in Indigenous Studies, Indigenous Communication Arts, Indigenous Business and the Reconciliation Certificate. She was also the Dallas W. Smythe Chair at the University of Regina School of Journalism from 2017 to 2018.

As a Métis woman committed to creating space in all institutions for Indigenous peoples, Merelda consults with organizations looking to learn about Indigenous Reconciliation and how to employ it effectively in the workplace.

Philip Howard

Director, Programme on Democracy and Technology and Professor of Internet Studies, Balliol College, University of Oxford

As Director of Oxford University's Programme on Democracy and Technology, Phil Howard oversees a large research team working on the use of new information technologies in politics, with the aim of raising civic engagement and improving public life around the world. In addition to his position as Director, Howard is a Professor and Fellow of Balliol College.

Howard, a scholar of political communication and an authority on global media, has long been immersed in the study of elections, conflict and international affairs. He has done field work in 16 countries – democracies and authoritarian regimes – and even worked as an election observer.

Ground-breaking investigations by Howard and his team have changed the global conversation about the role of social media in public life. Since 2014, Howard has led the study of misinformation around the world, through public writing and lectures, and has advised world governments, the technology industry and key civil society groups on the best responses to election interference, fake news and misinformation.

As an academic, Howard has taught courses on political communication, globalization, comparative media systems, international relations and social science research methods. He has published 10 books, edited volumes and authored over 130 scholarly articles, book chapters and working papers. He has won best book prizes from multiple professional organizations across the social sciences.

He was recently named a "Global Thinker" by Foreign Policy, and the National Democratic Institute gave him their "Democracy Prize" for pioneering the social science of fake news.

Vivek Krishnamurthy

Samuelson-Glushko Professor of Law at the University of Ottawa

Vivek Krishnamurthy is the Samuelson-Glushko Professor of Law at the University of Ottawa and Director of CIPPIC – the Samuelson-Glushko Canadian Internet Policy and Public Interest Clinic.

Vivek's teaching, scholarship and clinical legal practice focus on the complex regulatory and human rights-related challenges that arise in cyberspace. He advises governments, activists and companies on the human rights impacts of new technologies and is a frequent public commentator on emerging technology and public policy issues.

Vivek was previously the Assistant Director of Harvard Law School's Cyberlaw Clinic and Counsel in the Corporate Social Responsibility Practice at Foley Hoag LLP. He is a Rhodes Scholar and clerked for the Hon. Morris J. Fish of the Supreme Court of Canada upon his graduation from Yale Law School. Vivek is currently a Fellow of the Carr Center for Human Rights Policy at the Harvard Kennedy School, a Faculty Associate of the Berkman Klein Center for Internet & Society at Harvard University, and a Senior Associate of the Human Rights Initiative at the Center for Strategic and International Studies in Washington, D.C.

The Right Honourable Beverley McLachlin, PC, CC

Beverley McLachlin served as a Justice of the Supreme Court of Canada from 1989 to 2000 and as Chief Justice of the Court from 2000 to 2017.

Ms. McLachlin received her post-secondary education at the University of Alberta: B.A. (Hon.) 1965; M.A. 1968; LL.B. 1968. She practised law in Alberta and British Columbia and taught law at the University of British Columbia, before being named to the bench in British Columbia, where she served as a trial and appellate judge before being named to the Supreme Court of Canada.

Since retiring from the Supreme Court of Canada, Ms. McLachlin has pursued her interest in dispute resolution as an arbitrator and mediator, as a member of the Hong Kong Court of Appeal, the Singapore International Commercial Court and the Hong Kong International Arbitration Centre. She continues to work for access to justice, and to write and speak on legal and other matters in Canada and abroad.

Ms. McLachlin is a Companion of the Order of Canada and the recipient of numerous awards and honours.

Taylor Owen

Beaverbrook Chair in Media, Ethics and Communications and Associate Professor, Max Bell School of Public Policy, McGill University

Taylor Owen is the Beaverbrook Chair in Media, Ethics and Communications, the founding director of The Center for Media, Technology and Democracy, and an Associate Professor in the Max Bell School of Public Policy at McGill University. He is the host of the Big Tech podcast, a Senior Fellow at the Center for International Governance Innovation, a Fellow at the Public Policy Forum, and sits on the Governing Council of the Social Sciences and Humanities Research Council (SSHRC). He was previously an Assistant Professor of Digital Media and Global Affairs at the University of British Columbia and the Research Director of Tow Center for Digital Journalism at the Columbia School of Journalism. His Doctorate is from the University of Oxford and he has been a Trudeau and Banting scholar, an Action Canada Fellow and received the 2016 Public Policy Forum Emerging Leader Award.

He is the author of Disruptive Power: The Crisis of the State in the Digital Age (Oxford University Press, 2015) and the co-editor of The World Won't Wait: Why Canada Needs to Rethink its Foreign Policies (University of Toronto Press, 2015) and Journalism After Snowden: The Future of the Free Press in the Surveillance State (Columbia University Press, 2016). His forthcoming book with Emily Bell will be published by Yale University Press in 2021. His work focuses on the intersection of media, technology and public policy and can be found at www.taylorowen.com and @taylor_owen.

APPENDIX FOUR

TIMELINE OF COMMISSION CONVENINGS

DATE	ACTIVITY
Sept 28. 2021	**Orientation Session**
Oct 7. 2021	**Study Session** Included testimony from: **J. Nathan Matias**, Assistant Professor, Cornell University Department of Communication and Founder of the Citizens and Technology Lab **Rebekah Tromble**, Director of the Institute for Data, Democracy, and Politics, Associate Professor School of Media and Public Affairs at George Washington University
Oct. 14. 2021	**Study Session** Included testimony from: **Catherine Armitage**, Advisor at AWO Agency **Laura Edelson**, PhD Candidate, NYU Tandon School of Engineering **Ethan Zuckerman**, Associate Professor of Public Policy, Communication, and Information at the University of Massachusetts at Amherst and Founder of the Institute for Digital Public Infrastructure
Oct. 28, 2021	**Study Session** Included testimony from: **Seeta Peña Gangadharan**, Associate Professor at the Department of Media and Communications at the London School of Economics (LSE) **Laura Murphy**, Civil Liberties and Civil Rights Leader, Policy Strategist
Nov. 4, 2021	**Study Session** Included testimony from: **Kate Klonick**, Assistant Professor of Law, St. John's University Law School and Affiliate Fellow, Information Society Project, Yale Law School **Ravi Naik**, Legal Director, AWO Agency **Emily Laidlaw**, Associate Professor, Faculty of Law and Canada Research Chair – Cybersecurity Law, University of Calgary

Nov. 18, 2021	**Study Session** Included testimony from: **Divij Joshi**, Doctoral Researcher, University College London **Andrew Strait**, Associate Director, Ada Lovelace Institute **Jennifer Wemigwans**, Anishnaabekwe (Ojibwe/Potawatomi) from Wikwemikong First Nation and Assistant Professor, OISE University of Toronto
Nov. 21, 2021	**Study and Deliberative Sessions** Included Testimony from: The Citizen's Assembly on Democratic Expression 2021 **Meetali Jain**, Deputy Director, Reset **Marietje Schaake**, International Policy Director, Stanford University Cyber Policy Center and International Policy Fellow, Stanford's Institute for Human-Centered Artificial Intelligence **Mark Scott**, Chief Technology Correspondent, POLITICO
Nov. 22, 2021	**Study and Deliberative Sessions** Included testimony from: **Evan Balgord**, Executive Director, Canadian Anti-Hate Network **Cory Doctorow**, Journalist, Writer, Digital Rights Activist **Willie Ermine**, Assistant Professor with the First Nations University of Canada in Regina and from Sturgeon Lake First Nation located in north central Saskatchewan **Sue Gardner**, Founder and CEO, Tiny Ventures **Michael Geist**, Professor of Law, University of Ottawa **Mohammed Hashim**, Executive Director, Canadian Race Relations Foundation **Cynthia Khoo**, Technology and Human Rights Lawyer and Researcher **Brenda McPhail**, Director, Privacy, Technology & Surveillance Program, Canadian Civil Liberties Association **Ryan Merkley**, Managing Director, Aspen Digital, Aspen Institute
Jan. 5, 2022	**Study and Deliberative Sessions** Included testimony from: **Matthew Boswell**, Commissioner of Competition, Competition Bureau of Canada
Jan. 6, 2022	**Study and Deliberative Sessions** Included testimony from:

	Kevin Chan, Senior Global Director and Head of Public Policy Canada, Facebook **Rachel Curran**, Public Policy Manager Canada, Facebook **Colin McKay**, Head, Canada Government Affairs and Public Policy, Google
Jan. 27, 2022	**Study Session** Included testimony from: **Stéphane Perrault**, Chief Electoral Officer of Canada **Daniel Therrien**, Privacy Commissioner of Canada
Feb. 2, 2022	**Deliberative Session**
Feb. 9, 2022	**Deliberative Session**
Feb. 16, 2022	**Deliberative Session**
Feb. 23, 2022	**Deliberative Session**

APPENDIX FIVE

SUPPORTING MATERIAL

The Commission is grateful to the below persons for having prepared written materials to inform the Commission's study and deliberation.

Catherine Armitage, Advisor at AWO Agency

Laura Edelson, PhD Candidate, NYU Tandon School of Engineering

Divij Joshi, Doctoral Researcher, University College London

Kate Klonick, Assistant Professor of Law, St. John's University Law School and Affiliate Fellow, Information Society Project, Yale Law School

Emily Laidlaw, Associate Professor, Faculty of Law and Canada Research Chair – Cybersecurity Law, University of Calgary

J. Nathan Matias, Assistant Professor, Cornell University Department of Communication and Founder of the Citizens and Technology Lab

Ryan Merkley, Managing Director, Aspen Digital, Aspen Institute

Laura Murphy, Civil Liberties and Civil Rights Leader, Policy Strategist

Ravi Naik, Legal Director, AWO Agency

Seeta Peña Gangadharan, Associate Professor at the Department of Media and Communications at the London School of Economics (LSE)

Andrew Strait, Associate Director, Ada Lovelace Institute

Rebekah Tromble, Director of the Institute for Data, Democracy, and Politics, Associate Professor School of Media, and Public Affairs at George Washington University

Jennifer Wemigwans, Assistant Professor, OISE University of Toronto

Ethan Zuckerman, Associate Professor of Public Policy, Communication, and Information at the University of Massachusetts at Amherst and Founder of the Institute for Digital Public Infrastructure

The memos are available at www.ppforum.ca/demx

APPENDIX SIX

ACKNOWLEDGMENT

The Public Policy Forum would like to thank the members of the Commission Secretariat for their hard work and dedication in supporting the Commissioners in their deliberations.

In alphabetical order:

Gareth Chappell, Project Manager

Heba Elhalees, Event Coordinator

Peter MacLeod, Lead Facilitator

Adelina Petit-Vouriot, Editing and Project Support

Lisa Semchuk, Project and Research Associate

Sonja Solomun, Lead Researcher and Policy Analyst

Chris Waddell, Lead Writer

Sabrina Wilkinson, Research Fellow

Alessia Zornetta, Researcher

ENDNOTES

[1] For more detailed information about data access for researchers, including potential trade-offs see Caitlin Vogus & Emma Llansó, "Making Transparency Meaningful" (Centre for Technology and Democracy, December 2021), https://cdt.org/wp-content/uploads/2021/12/12132021-CDT-Making-Transparency-Meaningful-A-Framework-for-Policymakers-final.pdf

[2] For instance, over 120 civil society organizations have called on the European Union to adopt an Artificial Intelligence Act (AIA) which centres on fundamental rights. See "An EU Artificial Intelligence Act for Fundamental Rights" (30 November 2021), https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf and https://cdt.org/insights/eu-tech-policy-brief-january-2022-recap

[3] Page 11. CanadianCommissionOnDemocraticExpression-PPF-JAN2021-EN.pdf (ppforum.ca)

[4] DemX-RecommendationsToStrengthenCanadasResponseToDisinformationOnline-PPF-Jan2022-EN.pdf (ppforum.ca)

[5] See Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, *Report on disinformation,* OHCHROR, UN Doc A/HRC/47/35 (13 April 2021), https://documents-dds-ny.un.org/doc/UNDOC/GEN/G21/085/64/PDF/G2108564.pdf?OpenElement ; Privacy International, "The UN Report on Disinformation: A Role for Privacy," Privacy International, May 17, 2021, http://privacyinternational.org/news-analysis/4515/un-report-disinformation-role-privacy

[6] Yuan Stevens & Sonja Solomun. Facing the Realities of Facial Recognition Technology: Recommendations for Canada's Privacy Act, *Cybersecure Policy Exchange*, Feb. 2021. See also, Penney, Jonathon, Chilling Effects: Online Surveillance and Wikipedia Use (2016). Berkeley Technology Law Journal, Vol. 31, No. 1, p. 117, 2016, Available at SSRN: https://ssrn.com/abstract=2769645; Miles Kenyon, "Jon Penney on the Chilling Effects of Online Surveillance", (11 July 2017), *Citizen Lab*: https://citizenlab.ca/2017/07/jon-penney-on-the-chilling-effects-of-online-surveillance/#:~:text=Jon%20Penney%2C%20research%20fellow%20at,be%20hesitant%20to%20share%20content

[7] Willie Ermine, "The Ethical Space of Engagement" (2007) 6:1 Indigenous Law Journal 193 at 202

[8] Canadian Citizens' Assembly on Democratic Expression. (2022) "Canadian Citizens' Assembly on Democratic Expression: Recommendations to strengthen Canada's response to the spread of disinformation online." Ottawa, Public Policy Forum. p. 32

[9] Ibid p. 31

[10] "The Commission on Information Disorder Final Report." The Aspen Institute, Nov. 2021. CC BY-NC. https://creativecommons.org/licenses/ by-nc/4.0/

[11] Ibid. p .8

[12] European Commission, "The Digital Services Act package | Shaping Europe's digital future," (4 March 2022), online: *digital-strategyeceuropaeu*, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

[13] Government of Canada, "Technical Paper", (29 July 2017), https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/technical-paper.html

[14] See Caitlin Vogus & Emma Llansó, "Making Transparency Meaningful" (Centre for Technology and Democracy, December 2021), https://cdt.org/wp-content/uploads/2021/12/12132021-CDT-Making-Transparency-Meaningful-A-Framework-for-Policymakers-final.pdf Meaningful transparency also often includes user notifications, which falls outside the scope of this report

[15] Sonja Solomun, Maryna Polataiko, Helen A. Haye, "Platform Responsibility And Regulation In Canada: Considerations On Transparency, Legislative Clarity, And Design" (2021) 34 Harvard Journal of Law & Technology, https://jolt.law.harvard.edu/assets/digestImages/Solomun-Polataiko-Hayes.pdf

[16] For limitations of existing transparency reporting mechanisms, see Chris Tenove & Heidi Tworek, *Processes, People, and Public Accountability: How to Understand and Address Harmful Communication Online*, Research Report – Canadian Commission on Democratic Expression (2020), https://www.mediatechdemocracy.com/work/processes-people-and-public-accountability-how-to-understand-and-address-harmful-communication-online, at 14; Amélie Heldt, *Reading Between the Lines and the Numbers: An Analysis of the First NetzDG Reports*, 8 Internet Policy Review 2 (2019), https://policyreview.info/articles/analysis/reading-between-lines-and-numbers-analysis-first-netzdg-reports

[17] Composed of the Canadian Institutes of Health Research (CIHR), the Natural Science and Engineering Research Council (NSERC) and the Social Sciences and Humanities Research Council (SSHRC)

[18] Nate Persily, "U.S. Proposed Platform Transparency and Accountability Act" (2021), online (pdf): https://techpolicy.press/wp-content/uploads/2021/10/Persily-proposed-legislation-10-5-21.docx.pdf

[19] For more detailed information about data access for researchers, including potential trade-offs see Caitlin Vogus & Emma Llansó, "Making Transparency Meaningful" (Centre for Technology and Democracy, December 2021), https://cdt.org/wp-content/uploads/2021/12/12132021-CDT-Making-Transparency-Meaningful-A-Framework-for-Policymakers-final.pdf

[20] European Commission, "Commission Presents Guidance to Strengthen the Code of Practice on Disinformation" (2021), online: *European Commission* https://ec.europa.eu/commission/presscorner/detail/en/ip_21_2585

[21] EU, *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*, [2020], online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en

[22] Due to surveillance and due process concerns, the DSA may additionally exclude law enforcement agencies from data access regimes

[23] US, *Proposed Platform Accountability and Transparency Act*, online https://www.coons.senate.gov/imo/media/doc/text_pata_117.pdf

[24] EU, Proposal for a Directive of the European Parliament and of the Council amending Directive 2013/34/EU, Directive 2004/109/EC, Directive 2006/43/EC and Regulation (EU) No 537/2014, as regards corporate sustainability reporting, [2021], art 19c, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021PC0189&from=EN

[25] US, Code of Federal Regulations, 47 FR 11401, Part 229, as amended on 31 January 2022, https://www.ecfr.gov/current/title-17/chapter-II/part-229

[26] https://laws-lois.justice.gc.ca/eng/annualstatutes/2018_31/page-1.html

[27]Canadian Citizens' Assembly on Democratic Expression, "Canadian Citizens' Assembly on Democratic Expression: Recommendations to strengthen Canada's response to the spread of disinformation online" (Public Policy Forum, Ottawa, 2022), https://static1.squarespace.com/static/5f8ee1ed6216f64197dc541b/t/61f4701c5e04c053565d6c30/1643409442281/DemX-RecommendationsToStrengthenCanadasResponseToDisinformationOnline-PPF-Jan2022-EN.pdf pp 37-38

[28] Jamie Linde, "The Importance of EHR Interoperability" (30 September 2020), online: Wheel https://www.wheel.com/companies-blog

[29] Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, Madeleine Clare Elish & Jacob Metcalf, "Assembling Accountability: Algorithmic Impact Assessment for the Public Interest" (Data & Society, 19 June 2021), https://datasociety.net/library/assembling-accountability-algorithmic-impact-assessment-for-the-public-interest

[30] Page. 32. CanadianCommissionOnDemocraticExpression-PPF-JAN2021-EN.pdf (ppforum.ca)

[31] UK, Department for Digital, Culture, Media & Sport, "Online Harms White Paper: Full Government Response to the Consultation" (15 December 2020), online: *Gov.UK* https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response

[32] Page 33. CanadianCommissionOnDemocraticExpression-PPF-JAN2021-EN.pdf (ppforum.ca)

[33] Ysabel Gerard, "'Too good to be true': the challenges of regulating social media start-ups" in Tarleton Gillespie et al., "Expanding the debate about content moderation: scholarly research agenda in the coming policy debates" (2020) 9:4 Internet Policy Review

[34] Mark Zuckerberg, Testimony at the Hearing Before the United States House of Representatives Committee on Energy and Commerce Subcommittees on Consumer Protection 25 March 2015), online (pdf): https://docs.house.gov/meetings/IF/IF16/20210325/111407/HHRG-117-IF16-Wstate-ZuckerbergM-20210325-U1.pdf

[35] The safety duties for services likely to be accessed by children are to (Clause 10): "1. Take proportionate steps to mitigate and effectively manage the risk and impact of harms to children in different age groups 3. Prevent children of any age from encountering certain content 4. Protect children in age groups judged to be at risk of harm from encountering harmful content." See Reset, Online Written Evidence Submitted regarding Online Safety Bill (September 2021), https://committees.parliament.uk/writtenevidence/39851/pdf

[36] U.S., Justice Against Malicious Algorithms Act of 2021, [introduced] online: https://www.congress.gov/bill/117th-congress/house-bill/5596

[37] For instance, over 120 civil society organizations have called on the European Union to adopt an Artificial Intelligence Act (AIA) which centres on fundamental rights. See https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf and https://cdt.org/insights/eu-tech-policy-brief-january-2022-recap

[38] Many HRIAs are also conducted *after* harms have occurred

[39] Human rights impact assessments are the first step in identifying corporate misconduct and breach of human rights. This identification can be then used to pressure companies to correct procedural and operational mistakes. Such impact assessments can identify legal risks under human rights law which might not have been identified in algorithmic-impact-assessments.

[40] Individuals or organizations conducting human rights impact assessments should also be familiar with any cultural issues associated with the communities involved and be guaranteed adequate resources to conduct the audit.

[41] An AIA framework could combine features of existing impact assessments including data protection impact assessments (DPIA), human rights impact assessments and equality assessments. Some AIA models also differentiate themselves in that they involve the active participation and engagement with affected stakeholders All of them encourage reflexive consideration by developers of a technology to think deeply about the potential impacts before they take place. See e.g., Ada Lovelace Institute, "Algorithmic impact assessment: a case study in healthcare" (8 February 2022), online: https://www.adalovelaceinstitute.org/report/algorithmic-impact-assessment-case-study-healthcare

[42] The EU Data Protection Board has defined as "high-risk" those cases regarding: i) evaluation or scoring (e.g., profiling and predicting); ii) automated-decision making with legal or similar significant effect; iii) systematic monitoring; iv) sensitive data or data of a highly personal nature; v) data processed on a large scale; vi) matching or combining datasets; vii) data concerning vulnerable data subjects; viii) innovative use or applying new technological or solutions; ix) processing which in itself prevents data subjects from exercising a right or using a service or a contract. See EU Data Protection Board, Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, (2017), https://ec.europa.eu/newsroom/article29/items/611236

[43] There is a wide array of auditing methodologies that are used for different purposes, from auditing for potential bias in a dataset to auditing for the prevalence of hate speech on a specific platform. See Ada Lovelace Institute, "Technical methods for regulatory inspection of algorithmic systems in social media platforms: A survey of auditing methods for use in regulatory inspections of online harms" (December 2021), https://www.adalovelaceinstitute.org/report/technical-methods-regulatory-inspection for a survey of some relevant to the online harms space.

[44] Including a 'good Samaritan' safe harbour could encourage organizations to conduct these audits in order to have minor fines in case harmful outcomes are identified.

[45] The European Union's General Data Protection Regulation (GDPR) also requires that data controllers undertake an impact assessment (Article 35) when processing data in specific circumstances, including automated processing (Article 35(3)(a)). Article 26 of the European Union's Digital Services Act would require certain ICT companies to engage in yearly risk assessments that consider certain specified risks, including their services' impact on particular human rights. The proposed EU Artificial Intelligence Regulation aims at harmonizing the legal framework on artificial intelligence among EU Member States. Article 29 imposes on users of high-risk AI systems (i.e., systems that pose significant risks to the health and safety or fundamental rights of person) the obligation to carry out a data protection impact assessment (DPIAs) according to Article 35 EU GDPR.

[46] European Commission, "Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and Amending Directive 2000/31/EC," European Commission, December 15, 2020, Article 28, https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-parliament-and-council-single-market-digital-services-digital-services; European Commission, "Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and Amending Directive 2000/31/EC," European Commission, December 15, 2020, Articles 41 & 54, https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-parliament-and-council-single-market-digital-services-digital-services. In some cases, the Commission would also be empowered to request access to and information around platforms' databases and algorithms.

[47] The Directive requires providers of automated decision-making systems to document decisions; test and monitor outcomes; validate the quality of data; conduct security risk assessments; and report information on the effectiveness and efficiency of the system to the public. Described as a "light" approach, the Canadian questionnaire model includes promising features such as tiered requirements as risk increases, peer review and incorporating other assessments required by law. Other current AIA models include the questionnaire format, the data protection impact assessment (DPIA) and the public agency model. See Building a systematic framework of accountability for algorithmic decision making https://www.ifow.org/publications/policy-briefing-building-a-systematic-framework-of-accountability-for-algorithmic-decision-making

[48] Ontario's PHIPA requires that organizations maintain an electronic audit log in the context of personal health information to be provided to the Ontario Information and Privacy Commissioner on request. See Bill 188, *Economic and Fiscal Update Act*, 2020, s 10.1.

[49] Quebec's Bill 64 demands traceability in automated decision-making for individuals upon request. See Bill 64, *An Act to Modernize Legislative Provisions As Regards the Protection of Personal Information*, 1st Sess, 42nd Leg, Quebec, 2020, clause 102 (s 12.1(2) of the Act)

[50] Particularly around traceability and privacy impact assessments See Office of the Privacy Commissioner of Canada, "A Regulatory Framework for AI: Recommendations for PIPEDA Reform", (12 November 2020), online: *wwwprivgcca* https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/reg-fw_202011

[51] The former, guides companies through the process of re-assessing whether their governance and risk management are fit for purpose, also in light of the EU GDPR. The latter proposed eight risk areas specific to AI: fairness and transparency in profiling, accuracy, fully automated decision-making models, security and cyber, trade-offs (e.g., accuracy v privacy), data minimization and purpose limitation, exercising of individual rights, and impact on broader public interests and rights. See UK Information Commissioner's Office, "AI Auditing Framework" (2021), online: https://ico.org.uk/about-the-ico/news-and-events/ai-auditing-framework. Additionally, together with audit authorities in Norway, Germany, Finland and the Netherlands, the UK published the first international white paper on auditing machine-learning and AI algorithms in the public sector. See Supreme Audit Institutions of Finland, Germany, the Netherlands, Norway and the UK, "Auditing machine learning algorithms" (24 November 2020), online: https://www.auditingalgorithms.net

[52] The proposed Algorithmic Accountability Act would entrust the Federal Trade Commission with the task of issuing and enforcing regulations that would require certain entities using personal information to conduct impact assessments and "reasonably address in a timely manner" any identified biases or security issues.

[53] The Act takes a systematic approach to establishing a safety and effectiveness standard for algorithms, such that online platforms may not employ automated processes that harm users.

[54] UN Human Rights Office of the High Commissioner, *Guiding principles on business and human rights – Implementing the United Nations "Protect, Respect and Remedy" Framework,* [2011], online (pdf), UN http://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf. Similarly, the Section A.2 of the Organization for Economic Co-operation and Development (OECD)'s *Due Diligence Guidance for Responsible Business Conduct* requires companies to identify and assess actual and potential adverse impacts associated with their operations, products or services (2.2. Mentions human

rights impact assessments specifically). See OECD, *Due Diligence Guidance for Responsible Business Conduct*, [2008], online (pdf): http://mneguidelines.oecd.org/OECD-Due-Diligence-Guidance-for-Responsible-Business-Conduct.pdf

[55] Natasha Lomas, "On illegal hate speech, EU lawmakers eye binding transparency for platforms" (23 June 2020), online: *TechCrunch* https://techcrunch.com/2020/06/23/on-illegal-hate-speech-eu-lawmakers-eye-binding-transparency-for-platforms/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAADSpS383iySoMG0zqq-psDXcC2iT9TLSr_fGXHUWn4aUmehXKdbTqo_r1gaJi8bWbuAEyUUROO42tD0sY43J5xK4ly1LAo6VOk72CLyr2xVxAoL6T2b9kRRR30rMqHo4Q7eec4tv0Ht4c2yZzphTRKis0AeGXesGOsY4IJ5DTMjm

[56] Yuan Stevens & Sonja Solomun. Facing the Realities of Facial Recognition Technology: Recommendations for Canada's Privacy Act, *Cybersecure Policy Exchange*, Feb. 2021. See also, Penney, Jonathon, Chilling Effects: Online Surveillance and Wikipedia Use (2016). Berkeley Technology Law Journal, Vol. 31, No. 1, p. 117, 2016, Available at SSRN: https://ssrn.com/abstract=2769645; Miles Kenyon, "Jon Penney on the Chilling Effects of Online Surveillance", (11 July 2017), *Citizen Lab*: https://citizenlab.ca/2017/07/jon-penney-on-the-chilling-effects-of-online-surveillance/#:~:text=Jon%20Penney%2C%20research%20fellow%20at,be%20hesitant%20to%20share%20content

[57] Kayla Hilstob, "Karrmen Crey – Indigenous Epistemologies", (7 December 2021), online: Digital Democracies Institute https://digitaldemocracies.org/karrmen-crey-indigenous-epistemologies

[58] Chidi Oguamanam, "Indigenous Data Sovereignty: Retooling Indigenous Resurgence for Development" (December 2019) CIGI Papers No. 234 at 15

[59] Indigenous AI, "Indigenous Protocol and Artificial Intelligence Working Group", https://www.indigenous-ai.net

[60] Michael Running Wolf, "International Wakashan AI Consortium", online: *MIT SOLVE* https://solve.mit.edu/challenges/2020-indigenous-communities-fellowship/solutions/33358

[61] House of Commons of Canada, *Bill C-11 An Act to amend the Broadcasting Act and to make related and consequential amendments to other Acts*, (2 February 2022), online: https://www.parl.ca/DocumentViewer/en/44-1/bill/C-11/first-reading

[62] Australian Indigenous Knowledge IP Hub https://www.ipaustralia.gov.au/indigenous-knowledge-ip-hub

[63] Michael Morden et al., "Investing in Canadians' civic literacy: An answer to fake news and disinformation." (Toronto: The Samara Centre for Democracy, 2019), online (pdf): https://www.samaracanada.com/docs/default-source/reports/investing-in-canadians-civic-literacy-by-the-samara-centre-for-democracy.pdf?sfvrsn=66f2072f_4

[64] Philip N. Howard, Lisa-Maria Neudert, Nayana Prakash & Steven Vosloo, "Digital misinformation/ disinformation and children" (UNICEF, August 2021) https://www.unicef.org/globalinsight/reports/digital-misinformation-disinformation-and-children & Sonia Livingstone, Mariya Stoilova & Rishita Nandagiri, "*Children's data and privacy online Growing up in a digital age: An evidence review*" (2018) https://www.lse.ac.uk/media-and-communications/assets/documents/research/projects/childrens-privacy-online/Evidence-review-final.pdf

[65] Jessie Daniels, Mutale Nkonde & Darakhshan Mir, "Advancing Racial Literacy in Tech. Why Ethics, Diversity in Hiring & Implicit Bias Trainigns Aren't Enough" (Data & Society, 2019)

[66] Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. Journal of Personality and Social Psychology, 81(5), 856–868. https://doi.org/10.1037/0022-3514.81.5.856

[67] Jennifer Wemigwans policy memo session 5 Nov 16

[68] Canadian Citizens' Assembly on Democratic Expression, "Canadian Citizens' Assembly on Democratic Expression: Recommendations to strengthen Canada's response to the spread of disinformation online" (Public Policy Forum, Ottawa, 2022), https://static1.squarespace.com/static/5f8ee1ed6216f64197dc541b/t/61f4701c5e04c053565d6c30/1643409442281/DemX-RecommendationsToStrengthenCanadasResponseToDisinformationOnline-PPF-Jan2022-EN.pdf pp 37-38

[69] Cyphers, Bennett and Cory Doctorow (2021). Privacy Without Monopoly: Data Protection and Interoperability. Electronic Frontier Foundation. https://www.eff.org/wp/interoperability-and-privacy

[70] U.S., Bill HR 3894, *Augmenting Compatibility and Competition by Enabling Service Switching Act of 2021*, 117th Cong, 2021 (proposed), https://www.congress.gov/bill/117th-congress/house-bill/3849/text

[71] EU, Proposal for a Regulation of the European Parliament and of the Council on Contestable and Fair Markets in the Digital Sector (Digital Markets Act) [2020] com/2020/842 final, https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A842%3AFIN

[72] Canadian Citizens' Assembly on Democratic Expression, "Canadian Citizens' Assembly on Democratic Expression: Recommendations to strengthen Canada's response to the spread of disinformation online" (*Public Policy Forum*, 2022), online https://static1.squarespace.com/static/5f8ee1ed6216f64197dc541b/t/61f4701c5e04c053565d6c30/1643409442281/DemX-RecommendationsToStrengthenCanadasResponseToDisinformationOnline-PPF-Jan2022-EN.pdf

[73] EU, Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L 119, https://eur-lex.europa.eu/eli/reg/2016/679/oj

[74] California Consumer Privacy Act of 2018, S 1. Section 1798.100 of the Civil Code, as added by Section 3 of Chapter 55 of the Statutes of 201, https://leginfo.legislature.ca.gov/faces/billCompareClient.xhtml?bill_id=201720180SB1121&showamends=false

[75] Frank Pasquale, *Black box society: the secret algorithms that control money and information* (Cambridge, Massachusetts: Harvard University Press, 2016) at 14

[76] See e.g., Julia Angwin et al., "Machine Bias" *ProPublica* (23 May 2016), online: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; Anthony Flores, Kristin Bechtel & Christopher Lowenkamp, "False Positives, False Negatives, and False Analyses: A Rejoinder to 'Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks.'" (2016) 80:2 Federal Probation Journal, online: https://www.uscourts.gov/sites/default/files/80_2_6_0.pdf; Sonia K. Katyal, "Private Accountability in the Age of Artificial Intelligence" (2019) 66:1 UCLA L Rev 54

[77] See e.g., Sandra Wachter, Brent Mittelstadt & Chris Russell, "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR" (2018) 31:2 Harvard Journal of Law & Technology 842 at 843; Lilian Edwards & Michael Veale, "Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?" (2018) 16:3 IEEE Security & Privacy; Marco Almada, *Human intervention in automated decision-making: Toward the construction of contestable systems* (New York, NY, USA: Association for Computing Machinery, 2019)

[78] Sandra G. Mayson, "Bias In, Bias Out" (2019) 128:8 Yale L J 2218 at 2224; Solon Barocas & Andrew D. Selbst, "Big Data's Disparate Impact" (2016) 104:3 Calif L Rev 671

[79] Jeff Ward, "Black Box Artificial Intelligence and the Rule of Law" (2021) 84:3 Law & Contemporary Problems

[80] Frank Pasquale, *Black box society: the secret algorithms that control money and information* (Cambridge, Massachusetts: Harvard University Press, 2016) at 4

[81] Frank Pasquale, *Black box society: the secret algorithms that control money and information* (Cambridge, Massachusetts: Harvard University Press, 2016) at 3

[82] Allison Trites,"How Algorithmic Decision-Making is Changing How We View Society and People: Advocating fort he Right for Explanation and the Right to be Forgotten in Canada" (2019) 11:2 Global Media J 18 at 19

[83] Jeff Ward, "Black Box Artificial Intelligence and the Rule of Law" (2021) 84:3 Law & Contemporary Problems i at ii; Joshua A. Kroll et al., "Accountable Algorithms" (2017) 165 U Pa L Rev 633; Rashida Richardson et al., "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice" (2019) 94 NYU L Rev Online 15

[84] See DeltaFreq, Comment to Barry Ritholtz on blog post, "Where's the Note? Leads BAC to Ding Credit Score," The Big Picture (blog), December 14, 2010, 11:03 a.m., https://ritholtz.com/2010/12/note-bac-credit-score

[85] Jeff Harrington, "2010 Adds Its Own Terminology to Business Lexicon," Tampa Bay Times, December 23, 2010, https://www.tampabay.com/news/business

[86] Alan F.T. Winfield & Marina Jirotka, "The Case for an Ethical Black Box", in Yang Gao et al., *Towards Autonomous Robotics Systems* (Springer, 2017), at 265-266

[87] Dino Pedreschi et al., "Open the Black Box Data-Driven Explanation of Black Box Decision Systems" (2018) 1:1 Association for Computer Machinery 1 at 2

[88] Concerning electoral manipulation, see: Mark Scott, "Cambridge Analytica helped 'cheat' Brexit vote and US election, claims whistleblower", (27 March 2018), online: POLITICO https://www.politico.eu/article/cambridge-analytica-chris-wylie-brexit-trump-britain-data-protection-privacy-facebook (on the UK Brexit Referendum and the U.S. Presidential Elections of 2016); Bruna Martins dos Santos & Joana Varon, "Analysis of the playing field for the influence industry in preparation for the Brazilian general elections" (*Coding Rights for Tactical Technology Collective*, 2018), online: https://cdn.ttc.io/s/ourdataourselves.tacticaltech.org/ttc-data-and-politics-brazil.pdf (on the Brazilian Federal Elections of 2018). On misinformation, see Mieiam Fernández, Alejando Bellogín & Iván Cantador "Analysing the Effect of Recommendation Algorithms on the Amplification of Misinformation" (2021), online: https://arxiv.org/abs/2103.14748. On teenager and children harm, see Georgia Wells, Jeff Horwitz & Deepa Seetharaman, "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show" *Wall Street Journal* (14 September 2021), online: https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739

[89] Robyn Caplan & Danah Boyd, "Who Controls the Public Sphere in an Era of Algorithms? Mediation, Automation, Power" (*Data and* Society, 13 May 2016) at 8

[90] Riccardo Guidotti et al., "Factual and Counterfactual Explanations for Black Box Decision Making" (2019) 34:6 IEEE Intelligent Systems 14; Sandra Wachter, Brent Mittelstadt & Chris Russell, "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR" (2018) 31:2 Harvard Journal of Law & Technology 842 at 849-852, 860-872

DEMOCRATICEXPRESSIONDÉMOCRATIQUE

DEMOCRATICEXPRESSION.CA