CARL ABRC
CANADIAN ASSOCIATION OF RESEARCH LIBRARIES
ASSOCIATION DES BIBLIOTHÈQUES DE RECHERCHE DU CANADA

November 13, 2020

Mr. John Roberts
Chief Privacy Officer and Archivist of Ontario, and Chief Information Security Officer (A)
Information, Privacy and Archives Division and Cyber Security Division
Ministry of Government and Consumer Services
134 Ian MacDonald Blvd
Toronto, ON, M7A 2C5
Sent via email: John.Roberts@ontario.ca

RE: Consultation: strengthening privacy protections in Ontario

Dear John,

Personal information protection is a core value of the library and archival community. Libraries and archives preserve the public record, defend access to information and advocate on behalf of both research libraries and public interest for fair privacy laws in Canada. CARL would like to thank the Ontario government for consulting with Canadians on ways to protect citizens while encouraging innovation and advancement in Canada.

As a first and prefatory comment, we urge that privacy legislative reform in Canada be undertaken first and foremost at the federal level. Privacy protections and appropriate uses of information should apply evenly across Canada to reduce confusion and disparity that can occur through multiple, varying, provincial laws. Digital information and the commercial interests of internet companies do not recognize borders and we would encourage all provinces across the country to work with the federal government to achieve a common and forward-looking national legislation.

We trust that Canada (and Ontario) have studied whether the General Data Protection Regulation (GDPR) implemented in the European Union in 2018, and the California Consumer Privacy Act (CCPA) also from 2018, have been successful. However, if there are issues with these types of approaches, we would hope Canada would forge its own legislative approach.

For this consultation, we would like to focus our comments in four areas: the right to be forgotten, artificial intelligence, open by default, and the role of data management.

The Right to be Forgotten

Our position is that there are important rights and freedoms to be weighed, respected, and judiciously balanced in any legislative or regulatory approach to the Right to be Forgotten (RTBF).

In 1987, CARL adopted a freedom of expression statement which confers responsibility on Canadian research libraries to “facilitate access to all expressions of knowledge, opinion, intellectual activity, and creativity, from all periods of history to the current era, including those which some may consider unconventional, unpopular, unorthodox, or unacceptable.” This statement echoes the fundamental right to expressions of knowledge, creativity, and intellectual activities, as embodied in the Canadian Charter of Rights and Freedoms.

The RTBF is a complex emerging ethical and technological issue, which demands a careful balancing of fundamental rights that can at times appear to be in conflict.

Libraries are, by their very mission, upholders of the public interest and are sensitive to concerns around personal privacy on the Internet. Information on the internet can cause harm, particularly in cases where the information is false or defamatory. The RTBF can be a legitimate means for individuals to address such situations.

As preservers of the public record and defenders of freedom of speech and access to information, the research library community favours a legislative or regulatory approach to the RTBF that:

1. Aims to balance an individual's right to privacy with others’ freedom of expression. Any RTBF approach must ensure that the privacy rights of an individual who is the subject of content on the Internet does not unduly impinge on the expression rights of creators of the content, such as authors or publishers.

2. Protects from over-removal of content. If RTBF is encoded in legislation, lawmakers and/or regulators must be proactive in reducing the incentives of platforms like Google or Facebook to simply de-list information upon any request. It is important to remember that, for each time that an individual’s privacy is protected through a RTBF request, it may muffle the speech of those whose content is being delisted, thereby increasing the spectre of censorship.

3. Respects the integrity of the historical record. Information on the Internet may have future value for both the public and for researchers. We believe an expert assessment of the impact on the historical record, preserved for future generations of Canadians, and ways to mitigate that impact, should form part of every decision to remove information. In recommending this, research librarians recognize that the digital age has increased both the ephemeral nature and the accessibility of historical records that might otherwise have persisted only in physical library or archival repositories.

With this in mind, any approach to the RTBF that downplays visibility, such as suppressing access through search engines, seems marginally more acceptable than outright removal. In effect, de-listing removes information from the public view obtained through a simple keyword search, but does not actually remove it from the reach of a more skilled and persistent researcher who may also search repositories that are not indexed by search engines.

The Right to be Forgotten should not be able to be too casually invoked by individuals, or their requests too readily acceded to by search engines. If implemented, RTBF must have limited application, with clarity as to the conditions under which it may apply. There are complex considerations to be weighed and rights to be balanced, very likely requiring case-by-case assessment. In most cases, a review by an informed but impartial party is essential. A Right to be Forgotten regime that requires a judicial order for any information or data removal seems merited, rather than leaving companies like Google or in fact, research libraries with the task of deciding sensitive ethical situations pertaining to individual Canadians.

Artificial Intelligence

Big data benefits companies but can pose risk for consumers in terms of over surveillance, unregulated sharing and misuse of information, exposure to hacking, and data breaches. The larger the amount of consumer information that is being stored, the greater the risk to individuals.

Regulating personal data collection, storage, and usage to data that is relevant, adequate, and absolutely necessary for carrying out the purpose for which the data is processed, would provide protection for individual privacy and human rights both by reducing the risk of exposure of consumers' data, but also by minimizing the exploitation by industry of consumers' data for the purposes of financial gain.

The increasing use of artificial intelligence (AI) brings with it a myriad of concerns related to privacy, some of which are human rights (by reinforcing bias and systemic racism), transparency, an educated society, and a right to explanation. Democracy, laws, and human rights form the basis of our constitution and it is therefore important these principles are enveloped into this technology.

By implementing the principles of Privacy by Design[1] and Human Rights by Design into the very basics of AI design and operation, the act of "checking against privacy, human rights, and the basic tenets of constitutional democracy" is effectively embedded into the process. Under the EU General Data Protection Regulation (GDPR), businesses are required to build in Privacy by Design to all operations. Having already undertaken this process in this context, expanding design elements to include the foundational components of human rights would be the logical next step.

[1] Cavoukian, A., Privacy by Design "The 7 Foundational Principles) https://www.ipc.on.c /wp-content/uploads/resources/7foundationalprinciples.pdf

Increased transparency provides individuals with more comprehensive information on how their data is being and will be used which allows individuals to make better informed decisions.

The pervasiveness of AI and the lack of understanding by Canadians of its depth, reach and authority[2] is what has created the need for explainable AI (XAI) or the right to explanation. According to M. Ridley, author of the 2019 article entitled *Explainable Artificial Intelligence,* "Algorithmic decision-making, enabled by machine learning, is ubiquitous, powerful, often opaque, sometimes invisible, and, most importantly, consequential."[3] The right to explanation is seen as a method for transparency and accountability for artificial intelligence and other automated systems "that (have) for too long been defined by its opacity."[4]

Protections such as the right to object are relied upon by businesses, but they are meaningless if a person does not understand how the decision was taken and therefore the right to explanation is critical. The OECD AI Principles provide an example of some of the questions that should be considered by industry to be transparent and explainable[5]

We see there is a great amount of misunderstanding by the general public of AI and its pervasiveness. As such, businesses should be obligated to provide the protections mentioned above for consumers.

Open by default

CARL believes in the direction of open by default/open data/open science as a means to make available information to the public, for the public good--especially that which has been captured or created with public funds. World events in 2020 have demonstrated to us the importance of open, unfettered access to research, data, and public policy. For the most part, governments have adopted eight principles of open data as defined by Opengovdata.org[6]: We commend Canada's governments for the movement toward

[2] A Deloitte study surveyed 1,000 Canadians and found only 4 percent of survey participants reported they were able to explain what AI is and how it works.The Deloitte study also uncovered that 65 percent reported privacy concerns over the use of the data, 86 percent of Canadians said they don't think they currently use AI technology and 50 percent said they don't believe they will use them in the five years. Nearly a third of those surveyed in the Deloitte research don't think they will ever use AI in their lives. However, according to Statistics Canada 76 percent of Canadians own a smartphone. Deloitte. Canada's AI imperative Overcoming risks, building trust. P. 4, https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/deloitte-analytics/ca-overcoming-risks- building-trust-aoda-en.pdf
Statistics Canada. Life in the fast lane: How are Canadians managing?, 2016. https://www150.statcan.gc.ca/n1/daily-quotidien/171114/dq171114a-eng.htm?HPA=1

[3] Ridley, M. *Explainable Artificial Intelligence* https://www.google.com/url?q=https://publications.arl.org/18nm1df/&sa=D&ust=1583765950963000&usg= AFQjCNH8V0DL3Stfjfb0kmBnP0GVD3ffNw

[4] Ibid

[5] OECD. Transparency and explainability (Principle 1.3), https://oecd.ai/dashboards/ai-principles/P7

[6] The 8 Principles of Open Government Data, https://opengovdata.org/

transparency, civic engagement, and e-governance by adopting an open by default perspective.

Governments collect a tremendous amount of information and much of it is personal information about individuals. This can create large datasets of information that can be anonymized and released as open data. However, privacy must be at the forefront of these decisions.

Anonymization is unfortunately not failsafe. One of the main ways of anonymizing data is deidentification, but 2010 study *Broken Promises Of Privacy: Responding To The Surprising Failure Of Anonymization*[7] discusses a variety of techniques that re-identified personal information in several large datasets that had been made publicly available. According to CIPPIC, "First, when combined with other datasets, anonymous data can be re-identified. Second, personal data can be directly released in "anonymized" datasets, though often accidently. Finally, third parties can use their own data to re-identify anonymous open data."[8]

A public data trust would be acceptable if it is developed with rules of consent, transparency, and the right to explanation. Sharing data through a data trust would need to be disclosed as a purpose and would need consent from individuals when personal data is collected.[9]

This leads to the provision of justification for the sharing of data in order to obtain proper consent. While there are privacy risks associated with open data and data sharing, there is great benefit for governments to continue moving towards open by default. The privacy risks can be managed by implementing 'privacy by design" and its attendant policy and protocols.

Data Management

The library and archival community sees research data management as key to ensuring appropriate protection of individual privacy while, at the same time, enabling more data to be openly accessible and allowing technology-based research that mines anonymized or aggregated data sets. Research libraries play an increasing role in research data management and are very engaged in defining, practicing, and supporting ethical management of sensitive data, especially personal data.

CARL Portage Network's Sensitive Data Expert Group is composed of a broad membership from research communities – including research ethics professionals, representatives of funding agencies, and members of Indigenous organizations with

[7] 57 UCLA LAW REVIEW 1701 (2010), https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1016&context=hightechevents

[8] Ibid.

[9] Data trusts: legal and governance considerations, April 2019 https://theodi.org/wp-content/uploads/2019/04/General-legal-report-on-data-trust.pdf

direct interests in this subject, and is working to develop practical guides and tools for the management of sensitive research data in the Canadian landscape. The group recently released three guides as part of its Sensitive Data Toolkit for Researchers:[10] Glossary of Terms for Sensitive Data used for Research Purposes, Human Participant Research Data Risk Matrix, and Research Data Management Language for Ethics Approval and Informed Consent. Such tools will support researchers in adhering to current and future privacy laws. The development of technology and policy-based tools to facilitate ethical management of, and appropriate access to, sensitive data supports Canadian research participants by creating new opportunities to contribute data for the public good.

Research libraries are embedded in the research ecosystem. With sensitive data in particular, research libraries work on the front-line -- at the intersection of researchers, institutional policies, data repository requirements, publisher specifications, funder policies, research ethics principles, privacy law, and other competing and overlapping policy frameworks. As technology continues to increase research capacity, the inclusion of research libraries as stakeholders in the development of privacy law and policy will ensure strong understanding and representation of research data management considerations, improve Canadian research quality, protect Canadian research participants, and expand appropriate discovery and use of sensitive data.

CARL is the voice of Canada's research libraries. Our members include Canada's 29 largest university libraries and two federal institutions. CARL enhances its members' capacity to advance research and higher education; promotes effective and sustainable knowledge creation, dissemination, and preservation; and advocates for public policy that enables broad access to scholarly information. CARL's two federal member institutions contribute to Canada's research enterprise and collaborate in coordinated efforts with the academic library community, but do not engage in CARL's federal advocacy.

On behalf of CARL, thank you for the opportunity to participate in this consultation.

Sincerely,

CShepstone

Jonathan Bengtson
President, CARL

Carol Shepstone
Chair, CARL Public Policy Committee

cc. Susan Haigh, Executive Director, CARL

[10] Sensitive Data Expert Group publications https://portagenetwork.ca/network-of-experts/sensitive-data-expert-group/