Journal of Privacy and Confidentiality
Vol. 11 (2) 2021 Perspectives Submitted December 2020
Published August 2021

# RESEARCH DATA CENTRES - A REGULATOR'S PERSPECTIVE

MICHAEL HARVEY*

* Information and Privacy Commissioner for Newfoundland and Labrador, Canada

Abstract. As we continue to advance through the digital century, our governments, crown corporations, municipalities, school boards, regional health authorities, health care providers, and others are enhancing the digital element of the services that they provide. Even for those services that continue to be provided in the traditional way, there is a greater level of data collection involved. Generally speaking, this is a good thing. Many services can be delivered more broadly and efficiently when done so digitally, and having more data about all services makes it easier to tailor and improve them. However, privacy regulators such as the author of this article – the Information and Privacy Commissioner for Newfoundland and Labrador – are watching closely and with concern as our governments and public bodies collect more and more information about us. Commissioners advocate for greater openness of public bodies, and this can mean greater disclosure of information, but also the need to advocate for strong privacy protection – which can extend from the privacy principle of minimizing the collection of data in the first place, through holding it securely and disclosing it only under strict conditions and for established purposes, to destroying it as quickly as possible when no longer needed. These principles are difficult to square with the needs of researchers, who naturally want to get access to as much data as they can, as quickly as they can. It is the view of the Office of the Information and Privacy Commissioner (OIPC) for Newfoundland and Labrador (NL) that data centres can provide a way to advance these principles. This article establishes what the interest of the Commissioner is in the matter, arising from my statutory mandate, and discusses how the OIPC is able to promote such data centres while maintaining sufficient arms length from actual data centres in order to preserve its independent regulatory oversight role.

## 1. Overview of Information and Privacy Commissioners in Canada

The Information and Privacy Commissioner for NL is one of fifteen Access and Privacy Commissioners in Canada. Each province and territory has one, and at the federal level there are separate Commissioners for access to information and for privacy. Each Commissioner is provided with authority by statute, though the models differ. At the provincial/territorial

*Key words and phrases:* Research, Privacy, Access, Legislation, Oversight.

*Editorial note: This article is an edited version of the author's talk at the October 2020 Canadian Research Data Centre Network (CRDCN) conference. Information on the conference can be found at https://www.crdcn20.ca/crdcn20/program. Articles in the Perspectives series reflect the author's opinions, and do not necessarily reflect the opinions of the journal's editorial board.*


www.journalprivacyconfidentiality.org
DOI:10.29012/jpc.769

level, Commissioners are provided with authority under some version of a statute comparable to NL's *Access to Information and Protection of Privacy Act, 2015* (*ATIPPA, 2015*) which establishes the nature of access and privacy rights vis-à-vis public bodies in the province. Alberta, Quebec and British Columbian Commissioners have comparable oversight authority over privacy rights vis-à-vis private sector organizations, though the scope varies, while in other areas of the country the federal *Privacy Act* and the *Personal Information Protection and Electronic Documents Act* (*PIPEDA*) prevails with oversight responsibilities resting with the Office of the Privacy Commissioner of Canada. A number of provincial/territorial jurisdictions have some version of NL's *Personal Health Information Act* (*PHIA*), which establishes people's rights of access, correction and privacy to their own personal health information held by a custodian – a health services provider who has collected this information in the course of providing health services.[1]

Under these statutes, Commissioners have various powers. Some have what are known as "order-making powers," making them quasi-judicial administrative tribunals, issuing public decisions with reasons that take on the force of law and can be appealed to courts [10]. Others are described as the ombud model, making recommendations that public bodies may or may not choose to follow, with political rather than legal consequences. In NL, under *ATIPPA, 2015*, the Commissioner is empowered to make a wide variety of recommendations, but certain of these have greater legal force. The Commissioner may make recommendations related to: granting access to records; correcting personal information in a record; the collection, use or disclosure of personal information in contravention of the Act; or to destroy personal information collected in contravention of the Act. If a public body chooses not to follow these recommendations in whole or part, it must appeal to the court for a declaration that it does not have accept them. If it fails to do so, and yet does not implement a recommendation, then the Commissioner can file an order with the court, which gives the recommendation the force of law. This mix of ombud powers with limited order-making powers, introduced in 2015, is a hybrid model that the OIPC has found to be very appropriate for a small jurisdiction. The Office has found that, for the most part, the Commissioner can act as a flexible ombud, not as tightly bound by the strict requirements and procedures of administrative law, though in the knowledge of all parties that certain decisions are legally enforceable. *PHIA*, which came to force in 2011, provides the Commissioner with ombud authority.[2]

It is relevant to note that, beyond recommendation-making or order-making powers, the Commissioner enjoys a broad mandate for advocacy under *ATIPPA, 2015*. Section 3(1)(f)(i) establishes that the OIPC should be "an advocate for access to information and protection of privacy." The Act does not limit this to jurisdictional scope, i.e. to access and privacy of personal information held only by public bodies rather than private entities, but references these topics in general. In practice, the OIPC has interpreted this advocacy purpose to

[1]All provinces and territories with the exception of British Columbia and Nunavut have health specific access and privacy legislation, and access and privacy rights would be addressed by the legislation that governs public bodies (for those health care providers that are public bodies, like regional health authorities) or private sector organizations (for those that are private businesses, such as physicians in private practice).

[2]*PHIA* contains a mandatory statutory review provision at five years, and a review was launched in 2016. The Review Committee recommended that PHIA be amended to provide a hybrid oversight model akin to *ATIPPA, 2015* (Final Report of the *Personal Health Information Act* Statutory Review Committee, May 2017, amended September 2017). The provincial government has not yet introduced amendments to *PHIA* at the time of writing.

include personal information and personal health information held by public bodies, private companies and custodians.

## 2. Research and ATIPPA, 2015 and PHIA

Neither *ATIPPA, 2015* nor *PHIA* identifies the disclosure of personal information for the purposes of research as a purpose of the statute, though section 3(1)(a) of ATIPPA, 2015 establishes the first purpose as "ensuring that citizens have the information required to participate meaningfully in the democratic process" and a broad understanding of this purpose can conclude that access to information for secondary purposes, to improve public policy, can be an element. *ATIPPA, 2015* establishes that disclosure of information for the purposes of research is a legitimate purpose, and one that must lead to public benefit, by establishing, in section 70, the circumstances under which disclosure can occur: the research must be unable to be conducted without personal information being disclosed; linkages that occur between records are not harmful to individuals and benefits arising from disclosure are in the public interest; it can be established that the information will be held with security and confidentiality; information will be de-identified at the earliest possible occasion; and no unauthorized subsequent disclosure is permitted. This latter set of conditions needs to be formalized in an agreement between the public body and the person to whom it is disclosed. Moreover, section 40(2)(e) establishes that disclosure is not an unreasonable invasion of privacy if it is for research or a statistical purpose and is in accordance with section 70. Section 111 details the requirements for Publication Schemes. Though to date such schemes have not been implemented by the provincial government, the section envisions "a standard template for the publication of information by public bodies to assist in identifying and locating records in the custody or under the control of public bodies."

As for *PHIA*, beyond establishing that one of the purposes of the statute includes establishing rules for disclosure of personal health information, there is nothing to suggest that *PHIA* is preoccupied with the secondary use of data. The body of the Act, however, suggests that it is a broader preoccupation than with *ATIPPA, 2015*. Indeed, PHIA provides a legal framework for the disclosure of personal health information for the purposes of research. Section 44 is the key section that establishes that a "custodian may disclose personal health information without the consent of the individual who is the subject of the information for research purposes ... where the research project has been approved by a research ethics board or research ethics body under the *Health Research Ethics Authority Act*."[3] Research is a specifically defined term:

> *Section 2(1)(v): "research" means a systematic investigation designed to develop or establish principles or facts or to generate knowledge, or any*

[3]In NL, unique among provinces, the *Health Research Ethics Authority Act* (*HREA Act*) establishes in law that all health research involving human subjects conducted in the province must receive approval by the Health Research Ethics Board (HREB), which is constituted by the *HREA Act* and governed by the Health Research Ethics Authority (HREA), or approved by another health ethics board approved by the HREA. Regulations under the *HREA Act* establish that all clinical trials and research projects involving genetics must be approved by the HREB. The *HREA Act* establishes that a health ethics board must be guided by the Tri Council Policy Statement (now TCPS2), the International Conference on Harmonization of Technical Requirements for the Registration of Pharmaceuticals for Human Use Guidance E6: Good Clinical Practice: Consolidated Guideline, or some other HREA approved guidelines. In practice, TCPS2 guides almost all, if not all, reviews.

> *combination of principles, facts and knowledge, and includes the development, testing and evaluation of research.*

Other sections provide for indirect collection of personal health information for the purposes of research, clarify that it can be used or disclosed for research in the case of a deceased person, and establish a procedure for obtaining consent for a researcher to notify an individual about unauthorized disclosure.

## 3. OIPC and Research Data Centres

With the statutory framework thus established, I will now move on to explain how this mandate relates to research data centres. The starting point is the broad mandate for advocacy for access and privacy found in *ATIPPA, 2015* and the recognition that first, these two principles are based in constitutional rights and second, that there is a complex relationship between them.

The OIPC starts from the position, expressed by former Chief Justice of the Supreme Court of Canada Beverly McLaughlin, that access to information and privacy are "vital rights" and that *ATIPPA, 2015* and *PHIA* are akin to the federal statutes, which she has described as "twin laws of quasi-constitutional status," citing the Supreme Court of Canada in *Lavigne v Canada* [4] [6]. They are inherent to the operation of our liberal democracy, but just as liberal democracy embodies an inherent tension between the concepts of individual freedom and collective rule, so too is there an inherent tension between the notion that by default, information held by the state should be broadly accessible to the public and yet that the state is bound to hold information about individuals secure.

Many other commissioners (e.g. [1]) and academics (e.g. [3, 2]) write about the need to "balance" the rights of access and privacy. I, however, am of the view that the "balance" concept suggests a too-linear opposition between the two rights. Understood thusly, the relationship between the two may be a zero-sum game, in which for each question there is some point on a spectrum between the poles of openness and secrecy that must be found. A better conceptualization of the relationship between the two rights may be understood as an "inherent tension," wherein each right is at once diametrically opposed to the other yet dependent upon it for its own existence. Conceived in this way, we can imagine positive-sum approaches to the management of information that simultaneously enhance both access and privacy. I am of the view that research data centres are one such example of a positive-sum approach to the management of access and privacy. The case that I will focus on to illustrate this is the DataLab that is currently being developed by the Newfoundland and Labrador Centre for Health Information (NLCHI).

## 4. The Newfoundland and Labrador Centre for Health Information DataLab

Some background on NLCHI is required. The Newfoundland and Labrador health system has one of the most consolidated and interconnected set of databases of health information in Canada. The system has four regional health authorities (RHAs), supported by a single organization – NLCHI – which describes its mandate as "Developing, managing and operating a comprehensive and aligned information system; Developing data and technical standards; Managing provincial health data and information assets; Preparing health reports and conducting research and evaluation; [and] Providing analytics and decision support services" [7]. To advance this mandate, it holds and maintains numerous databases of

health information that it has been working for years to consolidate into a "data warehouse" with considerable investment of resources. This ongoing effort has created one of the most comprehensive and consolidated electronic health records in the country. It is further benefited by NLCHI's management of: a health information system with one vendor – Meditech - and only three instances;[4] an Electronic Medical Record with a single vendor that is being used by an ever-growing proportion of primary care providers in the community and within the RHAs; and a Pharmacy Network which connects every pharmacy in the province. The primary purposes of this effort have been to support enhanced clinical services and health system decision support through analytics, but the existence of these databases provides an extraordinary resource for secondary use as research.

NLCHI and the four RHAs have long released health information to health researchers, both academic and commercial, for the purposes of health research. While NLCHI has not provided record-level data to commercial researchers, it has provided record-level data to academic researchers. Data used by heath researchers is commonly de-identified, which in principle makes it no longer personal health information for the purposes of *PHIA*. However, NLCHI takes the position, with which the OIPC agrees, that the contemporary risk of re-identification, given advances in big data analytics, is such that it must assume that de-identification cannot be relied upon and that the data must be assumed to be personal health information [5, 9]. As noted above, NLCHI, as a custodian under *PHIA*, indeed has the legal authority to disclose personal health information, without consent in the case of a study approved by the HREB, for the purposes of research.

From an access and openness standpoint, OIPC supports the availability of information for research purposes. While the access rights referenced above are primarily tied directly to efforts of the public to hold public bodies accountable, it is not a significant extension to suggest that research about the delivery of public services such as health services that is intended to evaluate and improve those services is entirely consistent with that right. To the extent that the Data Warehouse project provides for the interconnection of the numerous datasets and allows for broader access, the OIPC is supportive of the extension of the access right.

However, there is a privacy concern that emerges from the increase in the number of individuals in possession of personal health information or de-identified information which can potentially be re-identified. This is where the DataLab comes in. NLCHI was provided a mandate by the provincial Minister of Health and Community Services to build a DataLab to provide access to data for authorized users, including health researchers, to protect privacy and confidentiality of individuals [8]. The DataLab draws data from the Data Warehouse, automatically de-identifies the datasets, and provides access to researchers. Access is virtual – researchers do not have to come to a physical site – and flexible – there are different interfaces depending upon the type of researcher and researchers can bring in their own third-party software applications and datasets of their own to work with the data. The critical aspect of the DataLab is that researchers do not take record level data into their custody. While they leave with the product of their work, they do not leave with the personal health information or data that could potentially be re-identified.

NLCHI's DataLab is an example of a solution developed in consideration of mandates for both access and privacy. NLCHI is a custodian under *PHIA* and of course has all of the authorities and responsibilities under that statute. Its governing statute is the *Centre for*

[4]Eastern Health, the largest of the regional health authorities, and Labrador-Grenfell Health share a single Meditech instance while each of Central and Western Health have their own instance.

*Health Information Act, 2018* which has, among its objects, section 4(1)(a)(i) "developing, operating and managing a comprehensive and aligned information system that fully integrates and uses data and health information from all components of the health and community services system for the delivery of health care and health system planning," which would presumably include health research, and under section 4(1)(b) "to protect the privacy of individuals whose personal information or personal health information is collected, used, disclosed, stored or disposed of by the centre." The DataLab is a solution that allows NLCHI to advance both of these mandates simultaneously, rather than one at the expense of the other. And to the extent that it does it properly, the OIPC supports the effort as consistent with the quasi-constitutional principles that undergird our mandate.

While the OIPC is supportive of the DataLab project in particular, and the concept of research data centres in general, we need to maintain our regulatory independence. The public must have the confidence that, should they have an access or privacy complaint about the DataLab, that we must retain the ability to investigate that complaint without having prejudiced ourselves through implication in the project's development. While this can be a delicate line to walk, it is not an unusual one. The OIPC has been briefed on numerous occasions by NLCHI on its DataLab project, dating back to the early days of its conception, and we have been provided with and provided comments on its Privacy Impact Assessment. However, our support of the concept, consistent with our advocacy mandate, is well short of pre-certification that would prejudice our complaint investigation mandate.

## 5. Conclusion

In summary, NLCHI's DataLab is one example of how access and privacy can be advanced not as a zero-sum balance between two competing principles but rather as positive-sum improvement to two principles that exist in an inherent co-dependent tension. In principle, the same would generally be true of other research data centres designed to simultaneously maximize access and privacy. It is my view that advocating for research data centres to be designed in such a way is consistent with my legal mandate and while other Commissioners are the best judges of their own mandates, I would suggest the same is true in jurisdictions across this country. We can advocate for such initiatives without prejudicing our independent oversight roles.

## References


[1] B. Beamish. *Protecting and Balancing Access and Privacy Rights: The Role of the Public's Right to Know*. University of Ontario Institute of Technology Distinguished Visitors Lecture Series, 2017.

[2] G. T. Duncan and S. F. Roehrig. Mediating the tension between information privacy and information access: The role of digital government. In G. Garson, editor, *Public Information Technology: Policy and Management Issues*, page 94–119. IGI Global, Hershey, PA, 2003.

[3] J. Lane and C. Shur. Balancing access to health data and privacy: A review of the issues and approaches for the future. *Health Services Research*, 45(2):1456–1467, 2010.

[4] R. Lavigne. v. Canada (Office of the Commissioner of Official Languages). 2002, 2 S.C.R. 773, at para. 24.

[5] B. Lubarsky. Re-identification of 'anonymized data.'. *Georgetown Law Technology Review*, 1(1):202–213, 2017.

[6] B. McLachlin. Access to information and protection of privacy in canadian democracy: Remarks of the Right Honourable Beverley McLachlin, P.C., Chief Justice of Canada, 2009. Text of speech delivered May 5, 2009. Supreme Court of Canada: Judges: Speeches.

[7] Newfoundland and L. C. Health Information, 2020. accessed December 6,.

[8] G. Newfoundland and Labrador. News release: Supporting healthcare innovation to benefit residents and create economic opportunities", 2018. May 4, 2018,.

[9] C. Porter. De-identified data and third party data mining: The risk of re-identification of personal information. *Shidler Journal of Law, Commerce & Technology*, 5(1):1–8, 2008.

[10] A. Roberts. New strategies for enforcement of the access to information act. *Queen's Law Journal*, 27:647–683, 2002.