Developments in privacy and access to education research data

Fred Carter

Senior Policy & Technology Advisor

Information and Privacy Commissioner of Ontario

Commissaire à l'information et à la protection de la vie privée de l'Ontario

AERO-AOCE

8 October 2024

Cont'd

Overview

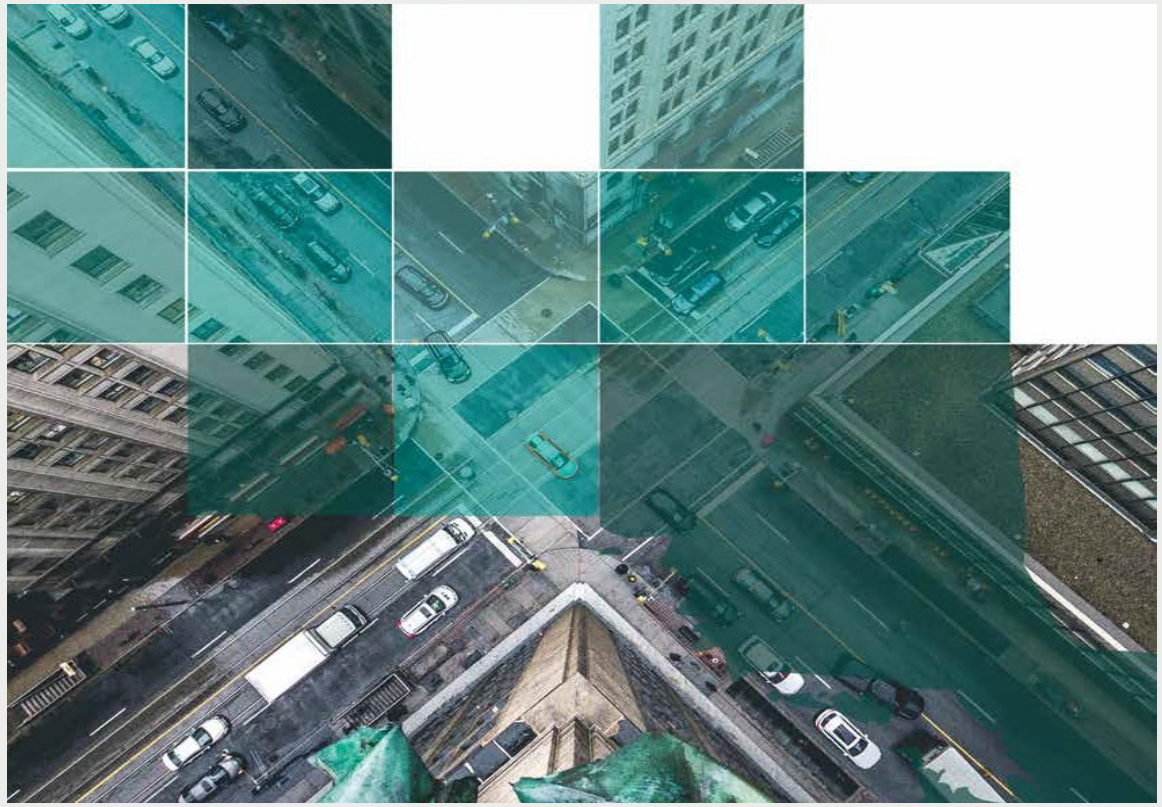
1. IPC 101
2. Open Data
3. De-identification
4. Key M/FIPPA provisions
5. Data research scenarios
6. Oversight over research consortia
7. IPC policy on Artificial Intelligence
8. Discussion

1. IPC 101

The Information and Privacy Commissioner of Ontario

- The IPC oversees Ontario's **access and privacy** laws
  - Our access and privacy laws establish the public's right to *access government-held information* and *protect their personal privacy rights*

- The IPC provides **independent** review of government decisions and practices on access and privacy



Information and Privacy Commissioner of Ontario

- **Patricia Kosseim**—appointed by Ontario Legislature (July 1, 2020)
- 5 year term
- Officer of the Legislature appointed by, and reports to, the Legislative Assembly of Ontario. Independent of the government of the day
- Assistant Commissioners:
  - Michael Maddock, Strategic Initiatives and External Relations
  - Warren Mar, Tribunal and Dispute Resolution Services

IPC Mandate

- *Freedom of Information and Protection of Privacy Act* (*FIPPA*)
  - covers 300 provincial institutions
- *Municipal Freedom of Information and Protection of Privacy Act* (*MFIPPA*)
  - covers 1,200 municipal organizations
- *Personal Health Information Protection Act* (*PHIPA*)
  - covers individuals and organizations involved in the delivery of health care services
- *Child, Youth and Family Services Act* (Part X) (*CYFSA*)
  - children's aid societies, child/youth service providers
- *Anti-Racism Act (ARA)*
  - oversight of the privacy protective rules

IPC Roles

- **investigate privacy complaints** related to personal information
- **resolve appeals** when there is a refusal to grant access to information
- **ensure compliance** with the acts we oversee
- **review** privacy policies and information practices
- **conduct research** on access and privacy issues and **provide comment** on proposed government legislation and programs
- **reach out and educate** the public, media and other stakeholders about Ontario's access and privacy laws and current issues affecting access and privacy

2. Open Data

Open Data (FIPPA s. 63 / MFIPPA s. 50)

- **Open Government**: public has right to access records and proceedings of government; essential for democracy
- **Open by Default**: information should be made public, unless legitimate legal, privacy, security or confidentiality reasons not to do so; mirrors FIPPA and MFIPPA's overarching access principles
- **Open Data**: proactive release of data in free, accessible and machine-readable formats, to encourage use by businesses, the public and government; seek to promote research, innovation and the development of new applications and services

INFORMATION AND PRIVACY COMMISSIONER OF ONTARIO

transparency CHALLENGE

Transparency and access to government-held information is about empowerment. It equips people with the information they need to participate meaningfully in the democratic process, engage in constructive discourse, and hold their governments accountable. It's the bedrock that democracy is built on, inspiring public trust and providing trustworthy, evidence-based information to shape policies, programs, and services to improve the lives of Ontarians."

COUNCIL OF ONTARIO UNIVERSITIES

Council of Ontario Universities' Open Data Platform Revolutionizes Access to Education Data

The Ontario Universities Open Data Platform brings together and makes publicly accessible, in an easy-to-use centralized hub, rich and extensive data on university operations, finances and student outcomes.

Ontario's universities remain accountable to the students and community members they serve, as well as government and taxpayers.

The platform includes access to key resources, such as:

- the Common University Data Ontario (CUDO) tool, which offers detailed information on enrolment, degrees granted and graduation rates;
- the Council of Ontario Finance Officers (COFO) tool, which provides detailed financial data and institutional financial statements;
- the Multi-Year Data series, which includes information on applications, enrolment, funding and student outcomes;
- the Ontario Universities' Application Centre (OUAC), which compiles application and confirmation statistics for various programs; and
- the Council of Ontario Universities (COU) data reports, which leverage university, provincial and other data to inform its policy recommendations and fulfill accountability commitments.

INFORMATION AND PRIVACY COMMISSIONER OF ONTARIO

transparency CHALLENGE

MINISTRY OF COLLEGES AND UNIVERSITIES

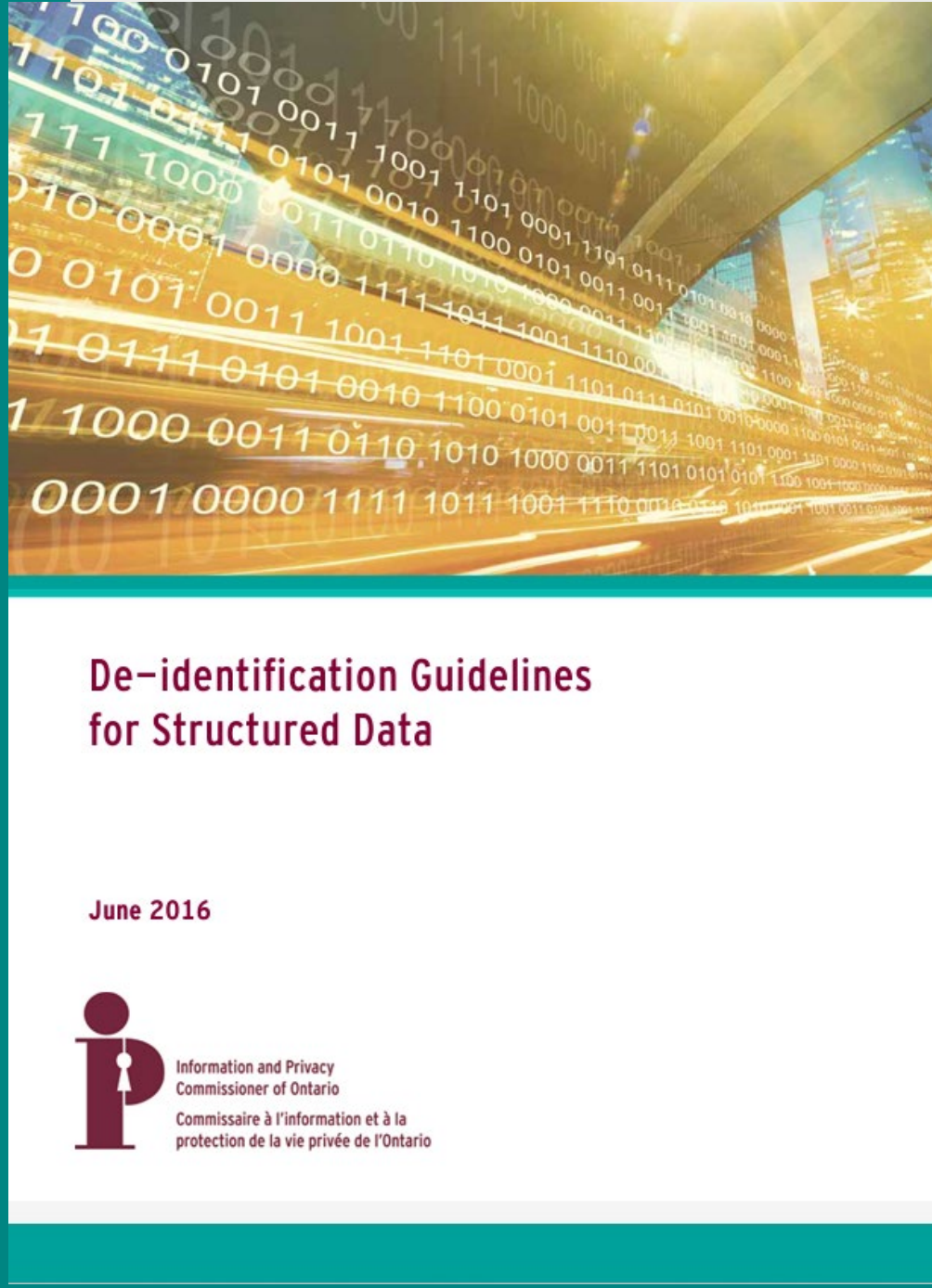
Quick Stats Tool Expands Access to Comprehensive and Detailed Postsecondary Education Data in Ontario

The Higher Education Quality Council of Ontario's Quick Stats offers detailed data on postsecondary education, helping researchers, policymakers, educators, and the public enhance policies and services and make informed decisions.

The Higher Education Quality Council of Ontario (HEQCO), an agency funded by the Ministry of Colleges and Universities, developed Quick Stats, an online tool with a rich collection of data on postsecondary education in Ontario. It includes information on applications, enrolment, student satisfaction, graduation rates, graduate outcomes, and tuition rates, with breakdowns by credential type and data on both domestic and international students.

3. De-identification

- Facilitates open data
- Useful in responding to access to information requests for structured data or data sets
- Under sections 10(2) of FIPPA and 4(2) of MFIPPA, institutions are required to "disclose as much of the record as can reasonably be severed" without disclosing any exempt information
- Preserves utility of the information
- Allows data sharing when institutions do not have authority to disclose personal information



Cont'd

De-identification

- Growing desire to break down govt "silos" and share more information within—and among—institutions. Examples:
  - information from one institution or program area may be relevant to the planning of a program or service in another institution or area
  - one institution may have expertise in data processing or software development that another institution requires, but does not have
  - an institution that funded a program or service that was delivered by another institution may want to evaluate the effectiveness of the program
- Data sets containing personal information may be shared within and among institutions only if the disclosure is permitted under privacy laws.

De-identification

- If disclosure is not permitted and the institutions still wish to share data sets, then personal information must be removed.
- Even if disclosure is permitted, there may still be important privacy issues to consider, i.e., diminishing the amount of control individuals have over their personal information.
- As a best practice, institutions should always consider de-identifying data sets before sharing.
- De-identified data may be exempt from application of privacy laws *

De-identification

Process for de-identifying a dataset:

1. determine the release model
2. classify variables
3. determine an acceptable re-identification risk threshold
4. measure the data risk
5. measure the context risk
6. calculate the overall risk
7. de-identify the data
8. assess data utility
9. document the process

Cont'd

De-identification

New demands, new uses, new methods. Additional guidance needed:

- De-identification for public release
- Measuring identity and attribute disclosure risks
- Managing risk when de-identified data is used by the data custodian
- Pseudonymization
- Data transformations for de-identification
- Synthetic data generation
- Assessing utility of de-identified data
- Checklist for evaluating de-identification documentation

4. Key M/FIPPA provisions

Personal information – Collection

Authority to Collect (FIPPA s. 38 / MFIPPA s. 28)

An institution can only collect personal information if:

- expressly authorized by a statute;
- used for the purposes of law enforcement; or
- necessary to administer a lawfully authorized activity.

- Must be collected directly from the individual.
- Limited circumstances to collect personal information indirectly, e.g.
  - Authorized by legislation, statute, regulations, orders-in-council, or by-law
  - Consent from the individual
  - Documentation requirements

Personal information – Notice

FIPPA s. 39 (2), s. 39 (3) / MFIPPA s. 29 (2), s. 29 (3)

- Notice to the individual must state:
  - legal authority for collection
  - reference to specific law, section or by-law
  - principal and any secondary uses of the personal information
  - title and business contact information

Exceptions to notice requirements

- Ministerial waiver if
  - Legal authority for indirect collection
  - Impossible or difficult to provide notice
  - Administrative burden is excessive

Cont'd

Personal information – Use & Disclosure

FIPPA s. 41, s. 42 / MFIPPA s. 31, s. 32

- Conditions:
  - Only used or disclosed for the purpose collected
  - Consistent purpose (indicated in notice of collection)
    - Reasonable expectations of individual
    - Compatible purposes for indirect collection
    - Individual consent (specific and in writing)
  - Compliance with other laws
  - Performance of duties
  - Compassionate circumstances

Cont'd

Personal information – Retention & Security

FIPPA s. 40 (1), Reg. 460 (5) / MFIPPA s. 30 (1), Reg. 823 (5)

- Minimum one year retention (four exceptions)

Section 3 of Regulation 823, made pursuant to MFIPPA states, in part:

(1) Every head shall ensure that reasonable measures to prevent unauthorized access to the records in his or her institution are defined, documented and put in place, taking into account the nature of the records to be protected.

(2) Every head shall ensure that only those individuals who need a record for the performance of their duties shall have access to it.

Research exemption

- IPC has interpreted research as “the systematic investigation into and study of materials, sources, etc., in order to establish facts and reach new conclusions, and as an endeavour to discover new or to collate old facts etc., by the scientific study or by a course of critical investigation.”
- Research purposes are distinct from administrative, operational or regulatory uses of personal information in that research uses do not directly affect the individual to whom the information relates and do not relate to the usual administration of a program.
- Program audits, evaluations and operational reviews are not research

Personal information - Research exemption

- MFIPPA s.14(1)(e) / FIPPA s.21(1)(e)
- An institution can disclose personal information for a research purpose if:

  "(i) the disclosure is consistent with the conditions or reasonable expectations of disclosure under which the personal information was provided, collected or obtained,

  (ii) the research purpose for which the disclosure is to be made cannot be reasonably accomplished unless the information is provided in individually identifiable form, and

  (iii) the person who is to receive the record has agreed to comply with the conditions relating to security and confidentiality prescribed by the regulations"

  Regulations reference standard agreement available on Central Forms Repository of the Government of Ontario for both FIPPA and MFIPPA.

Regulations (s.10 of Reg 460 (FIPPA) and Reg 823 (MFIPPA))

Regs set requirements for security and confidentiality that must be agreed upon before an institution may disclose personal information for a research purpose:

1. The person shall use the information **only for a research purpose** set out in the agreement or for which the person has written authorization from the institution.
2. The person shall name in the agreement any other persons **who will be given access** to personal information in a form in which the individual to whom it relates can be identified.
3. Before disclosing personal information to other persons under paragraph 2, the person shall **enter into an agreement** with those persons to ensure that they will not disclose it to any other person.

Regulations (cont'd)

4. The person shall keep the information in a **physically secure location** to which access is given only to the person and to the persons given access under paragraph 2.
5. The person shall **destroy all individual identifiers** in the information by the date specified in the agreement.
6. The person shall **not contact any individual** to whom personal information relates, directly or indirectly, without the prior written authority of the institution.
7. The person shall ensure that **no personal information will be used or disclosed** in a form in which the individual to whom it relates can be identified without the written authority of the institution.
8. The person shall **notify the institution** in writing immediately if the person becomes aware that any of the conditions set out in this section have been breached.

5. Data Research Scenarios

Scenarios

Student surveys
Linkage to internal data
Linkage to third-party data
Consortia
Case study

Student Surveys

- 2008 Privacy Investigation report MC06-53
- Parent complaint re: voluntary board-wide student survey
- IPC found collection and use were permitted for purposes under ss. 169.1(1), 170.1(1), and 171(1) of the *Education Act*
  - Necessary for planning, management, resource allocation and policy development purposes (i.e. to address “achievement gap”)
- Other issues discussed:
  - Appropriate notice
  - Linking of survey data with other data sources
  - Security measures, including de-identification

Linkage to Administrative/Achievement Data

- 2011 Privacy Investigation report MI10-5 similar to MC06-53
- Authorized and “necessary” for the proper administrative of a lawfully-authorized activity. Notice also satisfactory.
- Linkage of survey responses with identifiable administrative/achievement data → need to reduce risk of reidentification via new identifiers.
- Linkage deemed necessary to meet consistent purposes
- Append, disclosure (to government) meet reasonable expectations.
- Contracts in place with third-party service provider
- Other security controls

Linkage with third-party data

- 2017 joint research project: “Tracking student achievement from high school to University”
- Project linked school cohort data with university data to create a new research dataset containing identifiable information
- Collection and disclosure deemed to be authorized under M/FIPPA
- Linking and analysis (use) deemed consistent with s.31, i.e. “for the purpose for which it was obtained or compiled or for a consistent purpose.”
- Retention of dataset a concern: recommendations to destroy or de-identify dataset upon project termination.

Linkage with multiple third-party data sets

- 2020 education-related research project involving multiple high schools, post-secondary institutions, and HEQCO over 20 year period
- Proposal to create an infrastructure to collect and link identifiable data that will be de-identified prior to research use
- OEN used to link records; authorized by s.266(3) of *Education Act*
- Data flows among institutions require each party to have authority to collect, use and disclose personal information
- Reliance on 14(1)(e) of MFIPPA and section 21(1)(e) of FIPPA for disclosures
- Data minimization and de-identification methods applied; future notices to be updated to reflect new uses

PHIPA Decision 243 (21 May 2024)

IPC investigation / decision re: research consortium practices under PHIPA s.44

- Doctors = schools; Patients = students; UofT = lead on behalf of research consortium acting as agents of 600 primary care practitioners
- Ongoing extraction and retention of identifiable EMR data retained in longitudinal databases since 2014 (600,000 records)
- Research plan evolved to include identifiers, new data, new uses and disclosures
- No patient consent; invalid research plans and research agreement, inadequate notices, no meaningful ability of doctors to opt-out.
- Research database effectively shut down; de-identification practices scrutinized

Trends

- Growth of collection, use and disclosure of identifiable data
- Increased data matching and linkage, uses, longer retention
- Involvement of multiple parties
- Growing reliance on legal authority and consistent purposes
- Weak notices of collection or ability to opt-out
- Need for specialized expertise (necessity, de-identification, security)
- Involvement and continued oversight of Research Ethics Boards
- Consultations with IPC Policy

6. Oversight of research consortia

PHIPA s. 39 and s. 45

Prescribed persons / entities are permitted to collect PHI without consent for purposes of planning, management and analysis of the health system

- Corresponding authority of HICs to disclose
- Huge longitudinal databases created from many sources
- Prescribed persons and entities must meet certain conditions / restrictions
- Trienniel IPC review and approval of practices and procedures
- Extensive IPC review process, based on Manual for the Review and Approval of Prescribed Persons and Prescribed Entities (updated Jan 2024)
- IPC has power to shut down operations

Data Integration Units (FIPPA)

- Part III.1 enables prescribed data integration units (DIU) to collect PII for linking to create / enable access to de-identified datasets for analysis of:
  - management or allocation of resources
  - planning for the delivery of programs and services provided or funded by the Government
  - evaluation of those programs and services
- The legislation, as well as related data standards, set out specific requirements and restrictions that data integration units must follow related to, among other things, the collection, use (including linking and de-identification), disclosure, security, retention, and destruction of personal information.
- IPC mandated to review and approve DIU policies and procedures

Interpretations

- What is research?
- Data minimization and “necessary”
- Lawfully authorized / consistent purposes / uses
- De-identified data and risks of re-identification
- Governance, oversight and enforcement / Role of REBs
- Terms of research agreements / recipient rule
- Effective notices / consent / subject access rights
- Interplay between operation of different Acts (e.g. ARA)
- Status and role of research consortia

7. IPC Policy on use of AI

IPC Policy on AI

- IPC acknowledges potential of AI technologies to significantly enhance Ontarians' lives. Public sector use of AI can accelerate the delivery of government services, enhance government decision-making, improve public engagement, and help solve complex societal problems.
- However, AI systems are not infallible. Setting up AI systems often depends on vast amounts of personal information that may be highly sensitive and may be inappropriately shared with others. They sometimes return inaccurate results for reasons that are nearly impossible to explain and account for.
- Automated decisions based on information or inferences resulting from AI systems may significantly impact people's lives. They might perpetuate discrimination and bias against historically marginalized groups.

IPC AI Policy

Fundamental AI principles and guardrails should codified in statute.

The IPC strongly recommends that the proposed legislation [Bill 194] be amended to include explicit statutory language that sets out the basic parameters within which eventual regulations must be established.

By codifying strong normative principles in the statute itself, the public can be assured that a robust, transparent, and principles-based approach will help enable the potential benefits of these powerful technologies while protecting individuals and groups from potential harms.

Cont'd

IPC AI Policy

Public sector entities developing or deploying AI systems must ensure that such systems are:

- Valid and reliable
- Safe
- Privacy Protective
- Transparent
- Accountable
- Human Rights Affirming

Cont'd

IPC AI Policy

Bill 194 recommendations:

- Add statement of interpretive principles and AI principles
- Risk-based regulatory approach
- No-go zones
- Recommendations specific to digital technologies affecting individuals under 18
- The third portion of Schedule 1 entitled Digital Technology Affecting Individuals Under the Age of 18, would establish a regulatory regime governing the digital information that school boards and children's aid societies may collect, use, retain or disclose relating to individuals under 18 and the types of digital technologies they may make available to children and youth.

Cont'd

Access / Consent

- Any individual may make an access request – no minimum age requirement
- Duty to assist, provide information in timely and useful manner, and to respond to requests for corrections and annotation
- Issues of capacity and consent
- *MFIPPA* and *Education Act* have different age requirements
- *MFIPPA*: parent or guardian of child may consent on child's behalf only if under 16
- Education Act: parent or guardian of student may provide written consent for use or disclosure of information in child's OSR only if under 18
- In general, *Education Act* applies to OSR, MFIPPA to info not part of the OSR

8. Discussion

HOW TO CONTACT US

Information and Privacy Commissioner of Ontario

2 Bloor Street East, Suite 1400

Toronto, Ontario, Canada M4W 1A8

Phone: (416) 326-3333 / 1-800-387-0073

TDD/TTY: 416-325-7539

Web: www.ipc.on.ca

E-mail: info@ipc.on.ca

Media: media@ipc.on.ca / 416-326-3965