Brian Beamish
Information and Privacy Commissioner of Ontario

January 26, 2017

Government and Big Data:
Privacy Risks and Solutions

Information and Privacy Commissioner of Ontario
Commissaire à l'information et à la protection de la vie privée de l'Ontario

Ontario's Access and Privacy Laws

- The *Freedom of Information and Protection of Privacy Act (FIPPA)*
  - applies to over 300 provincial institutions such as ministries, provincial agencies, boards and commissions, as well as community colleges and universities
- The *Municipal Freedom of Information and Protection of Privacy Act (MFIPPA)*
  - applies to over 1,200 municipal institutions such as municipalities, police services boards, school boards, conservation authorities and transit commissions
- The *Personal Health Information Protection Act (PHIPA)*
  - covers individuals and organizations in Ontario that are involved in the delivery of health care services, including hospitals, pharmacies, laboratories and health care providers such as doctors, dentists and nurses

The Historical Perspective

- Concerns about the privacy implications of data integration existed before *FIPPA* and *MFIPPA* were proclaimed in force
- 1980 Williams Commission Report on *Freedom of Information and Individual Privacy* stated:

  "*The prospect of greater integration of databases raises, in turn, a number of privacy issues…*

  …it is feared that the use of such dossiers may constitute a form of data surveillance which might operate against the legitimate interests of the individual"

What Is Big Data?

- Equal parts buzzword and concept
- Gartner's three V's: high-**volume**, high-**velocity**, high-**variety**
- McKinsey Global Institute:
  - "datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze"
- Represents a shift in how we think about and use data:
  - New combinations of data may contain useful, but hidden patterns and insights
  - Advanced analytics can discover these insights
- Ever-evolving term used to denote any data-intensive technology or analysis

Big Data for Government

- The sharing, linking and analysis of data across government to provide new insights for the purposes of supporting:
  - policy development
  - system planning
  - resource allocation
  - performance monitoring
- Sometimes referred to as “data integration”

Privacy and Big Data

- Fundamental tension between big data and some basic principles of privacy:
  - personal information (PI) should be collected directly from the individual
  - PI should only be used for the purpose for which it was collected
- Big data involves information that has been:
  - collected indirectly
  - used for a purpose which may not have been intended at the time of collection
- Additional set of privacy measures needed to allow for big data

Privacy Risks of Big Data

- Generation of new PI not collected directly from the individual
- Use of poorly selected data sets that:
  - lack information/are incomplete
  - contain incorrect or outdated information
  - disproportionately represent certain populations
- Incorporation of implicit or explicit biases
- Generation of pseudo-scientific insights that assume correlation equals causation
- Lack of knowledge/transparency regarding the inner "logic" of the system
- *If not designed properly, can result in uses of PI that may be unexpected, invasive and discriminatory*

Current Model of Data Protection

- *FIPPA* and *MFIPPA* reflect the needs and expectations of a different time:
  - information technology was less prevalent
  - types of data and analytics were less complex
  - uses of personal information were discrete and determinate
- The result is a model of data protection where government institutions are treated as "silos":
  - collection of personal information must be "necessary"
  - secondary uses are generally prohibited
  - information sharing is restricted

The Big Data Challenge

- Big data represents an era where:
  - information technology is ubiquitous
  - types of data and analytics are complex
  - uses of personal information are less discrete and less determinate
- Striking the right balance between data use and the protection of privacy is challenging
- How can we ensure data protection while enabling the personal and societal benefits that come from the use of big data?

Reform of *FIPPA* and *MFIPPA*

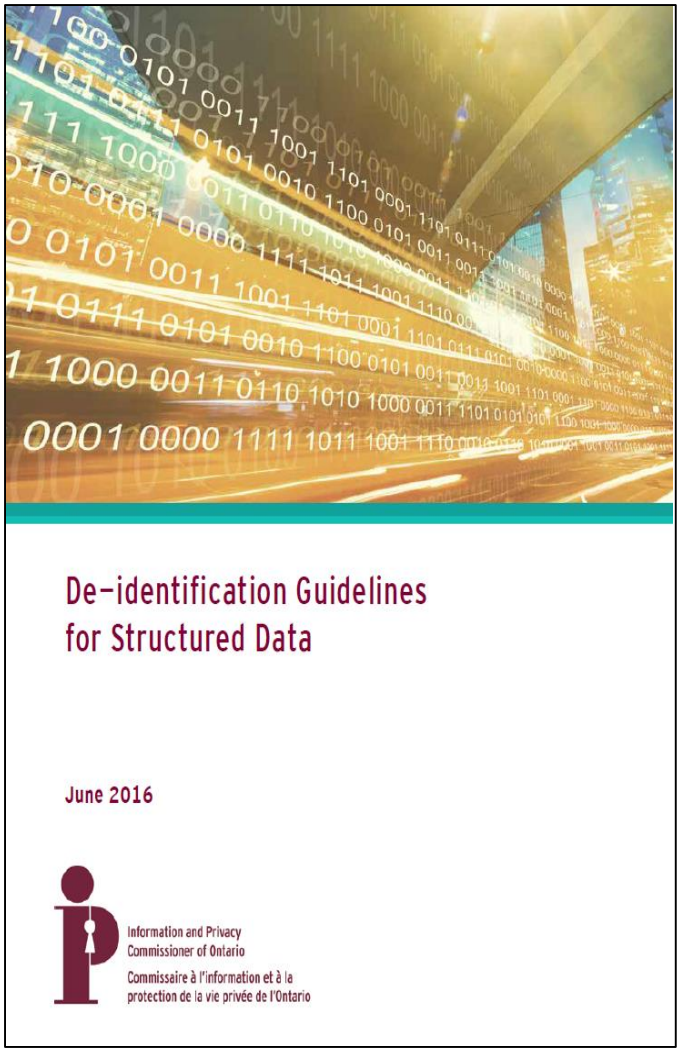
- Need principled-based legislation governing data linking and big data analytics which could include the following safeguards:
  - creation of a data institute or institutes with expertise in privacy, human rights and ethical issues involved with data integration and analytics
  - requirements for data minimization
  - privacy impact assessments and threat risk assessments
  - mandatory breach notification and reporting to the IPC and the affected individuals
  - order-making and audit powers for the IPC

The IPC and Big Data

- The IPC is committed to ensuring the privacy of Ontarians is protected
- The IPC has been involved in addressing issues raised by big data in numerous ways, including:
  - releasing guidance materials
  - consulting with government institutions
  - providing comments on legislative amendments
- Overall position: It is possible to use big data in a privacy-protective manner

IPC Guidance on De-identification

- "De-identification" – the removal of personal information from a record or data set
- Provides a step-by-step process for de-identifying data sets
- Discusses key issues of:
  - direct and indirect (or "quasi-") identifiers
  - types of re-identification attacks
  - common de-identification techniques
  - disclosures for open data and research
- The privacy protections of *FIPPA* and *MFIPPA* do not apply to de-identified information



Information and Privacy
Commissioner of Ontario
Commissaire à l'information et à la
protection de la vie privée de l'Ontario

Example - Amendments to *PHIPA*

- IPC worked with the Ministry of Health and Long-Term Care (MOHLTC) to enable beneficial "big data"-type uses of personal health information (PHI) while protecting privacy
- S. 55.9 of *PHIPA* provides MOHLTC with the authority to collect PHI indirectly and link it for the purposes of:
  - funding, planning or delivering health services
  - detecting, monitoring or preventing fraud
- However, MOHLTC must:
  - unit designated by regulation
  - put in place practices and procedures approved by the IPC to protect the privacy of individuals
  - de-identify the PHI

IPC Fact Sheet on Big Data for the Public

- Released for Data Privacy Day
- Developed to help members of the public understand what big data is, and how it can have an impact on an individual's privacy
- Discusses key issues, such as:
  - proportionality
  - accuracy of results
  - bias in data sets
  - individual rights

JANUARY 2017

PRIVACY
FACT SHEET

Big Data and Your Privacy Rights

New tools for combining and analyzing information have made it possible for researchers to uncover hidden patterns and connections in large data sets that would have previously been unknown. Collectively, these large data sets and the analytical tools and practices used to identify trends are known as 'big data.' While private sector companies often use big data analyses to support marketing and product development, public organizations are attracted to it as a way to improve policy and program development and ensure it is supported by better evidence.

This fact sheet has been developed to help members of the public understand what big data is, and how it can have an impact on an individual's privacy.

Big data has the potential to provide governments with greater insights into the quality and effectiveness of services and programs such as healthcare, social services, public safety and transportation. However, it also raises concerns regarding privacy and the protection of individuals' personal information.

The Office of the Information and Privacy Commissioner of Ontario (IPC) is responsible for oversight of the ***Freedom of Information and Protection of Privacy Act***, the ***Municipal Freedom of Information and Protection of Privacy Act*** and the ***Personal Health Information Protection Act***. Organizations governed by these acts, such as government ministries, municipalities, police services, health care providers and school boards, must comply with these acts when collecting, using and disclosing personal information.

This fact sheet has been developed to help members of the public understand what big data is, and how it can have an impact on an individual's privacy.

Information and Privacy Commissioner of Ontario
Commissaire à l'information et à la protection de la vie privée de l'Ontario

Information and Privacy Commissioner of Ontario
Commissaire à l'information et à la protection de la vie privée de l'Ontario

Forthcoming: Big Data Guidelines

- To be released in Spring 2017
- Developed to inform institutions of key issues to consider and best practices to follow when conducting big data projects involving personal information
- Topics include:
  - data linking protocols
  - ethics review boards
  - public notification
  - profiling
- Discussion panel at International Association of Privacy Professionals (IAPP) Canada Privacy Symposium 2017

Conclusion

- The bigger the data, the greater the responsibility to protect the privacy of individuals
- *FIPPA* and *MFIPPA* were not designed with big data in mind
- It is possible to use big data in a privacy-protective manner
- Striking the right balance in Ontario requires input from all stakeholders, including:
  - government
  - public
  - regulators

How to Contact Us

Information and Privacy Commissioner of Ontario
2 Bloor Street East, Suite 1400
Toronto, Ontario, Canada
M4W 1A8

(416) 326-3333 / 1-800-387-0073
TDD/TTY: 416-325-7539
www.ipc.on.ca
info@ipc.on.ca