Privacy and Open Government: A Balancing Act

Dr. Teresa Scassa

Canada Research Chair in Information Law, University of Ottawa, Faculty of Law

Prepared for the UN Expert Group Meeting on *Moving from commitments to results in building effective, accountable and inclusive institutions at all levels* New York, February 18, 2016

Competing Values

- Openness and transparency in government can serve the goals of greater accountability and greater public trust in government
- These principles must be balanced with other competing values, which include: the protection of privacy, national security, and the protection of confidential business information
- The protection of privacy in relation to information in the hands of government "is arguably related to wider attitudes about participation in public affairs and about trust in the authority of government agencies." (Bennet & Raab 2006)

Changing Context

- Digital information and the internet
  - Information can be quickly and easily shared & communicated, is easily searchable and can be represented in novel and interesting ways
    - Increases potential for openness to enhance transparency, citizen engagement, innovation
    - Increases risks associated with improper disclosure
- Big data environment
  - Digital government data can be harvested and combined with other data by data brokers and others; even anonymized data sets can be analyzed in combination with other data in ways that may lead to reidentification
    - Use of government data in big data analytics may increase privacy risks (e.g. profiling)
    - Difficult to assess reidentification risks or to balance potential privacy harms with transparency values

Opening Government Information

1. Publicly available personal information
   - Personal information in the hands of government which by law or policy is required to be made public (e.g.: information in public registries, court decisions, etc.)
2. Access to information/right to know legislation
   - Information in the hands of government that is released upon request according to an access to information statute
3. Proactive disclosure
   - Information that the government chooses to make public
4. Open data
   - Data sets that are made publicly available in reusable formats under an open licence

Publicly available personal information

Information published as part of a public registry	• E.g.: land titles registry information, registers of building permits, sex offender registries, registers of physicians or other professionals
Information required to be published by specific laws	• E.g.: laws requiring publication of names and salaries of public officials, campaign financing laws
Information required to be published by legal/constitutional principles	• E.g.: open courts principle, due process principles • Court decisions, court records

Publicly available personal information

- In these cases, the public policy decision has already been made that the public interest is served by making this personal information public
- BUT:
- Does the manner and form of publication have privacy implications that should be mitigated by the state?

Mode and manner of publication may affect impact on privacy

- Examples:
  - Online publication of court or tribunal decisions without redaction of names where sensitive personal information is present v. availability in law books or in for-fee databases
  - Online and publicly accessible registries v. registers available for consultation or specific requests for information
  - Open data publication of name and salary information of public servants v. publication in restrictive file formats

Laws requiring government to *protect* personal information from disclosure

Access to information/ right to know laws

- Generally apply to all information in the hands of government
- Basic principle is disclosure unless there is an exception justifying non-disclosure
- Protection of privacy is a key exception to requirement to disclose

Public sector data protection laws

- Require governments to protect personal information they collect from citizens
- Would apply to policies of proactive disclosure and open data

Laws requiring government to *protect* personal information from disclosure

- Obligation to protect privacy is generally not absolute
  - Disclosure of personal information may be permitted (or required) in some circumstances
  - Non-personal information may be disclosed
- If personal information cannot be disclosed, how do we determine what constitutes personal information?

Defining Personal Information

- Personal information = “information relating to an identified or identifiable natural person” (EU Directive)“information about an identifiable individual” or “personally identifiable information”
  - Directly identifying information (e.g.: name, identification numbers, etc.)
  - Indirectly identifying information (e.g.: information that may lead to identification when combined with *other available information*)

Other available information

- Other available information may include:
  - Other government information
  - Information from other sources:
    - Newspapers
    - Neighbors, friends, associates
    - Internet, social media
    - Data sets in the hands of private sector companies

Illustrations

- Netflix released data set of .5 million anonymized movie ratings (linked to unique identifiers) seeking public input into how to improve its ratings service
  - Reviews matched to others in IMDb to reidentify reviewers
- AOL released 20 million anonymized search strings of 658,000 people
  - Pieces of information in search strings matched to other available information including telephone directory information to reidentify individuals

Balancing privacy and transparency with open data/proactive disclosure

- Extreme concern over potential privacy impacts can lead to:
  - Decisions to not release certain data/information
  - Excessive reliance on technological barriers to reuse or restrictive licensing that can limit effective reuse of data/information
  - Overprotection of marginal privacy interests at the expense of transparency

Best practices for balancing privacy and transparency in open government

1. Pre-release processes/practices internal to government
   - Data minimization: governments collect the least amount of personal information necessary
   - Cross-government communication: knowledge-sharing about decision process; understanding what other information is being released
   - Training, polices and procedures around release of open data
   - Assessing privacy risks: perform a privacy impact assessment where appropriate before disclosing data sets

Best practices for balancing privacy and transparency in open government

2. Balancing
   - Identify transparency values/goals motivating release of data
   - Assess privacy risks
     - Likelihood of reidentification
     - Degree of sensitivity of information
   - If release is favoured consider strategies/techniques to minimize risk/harm

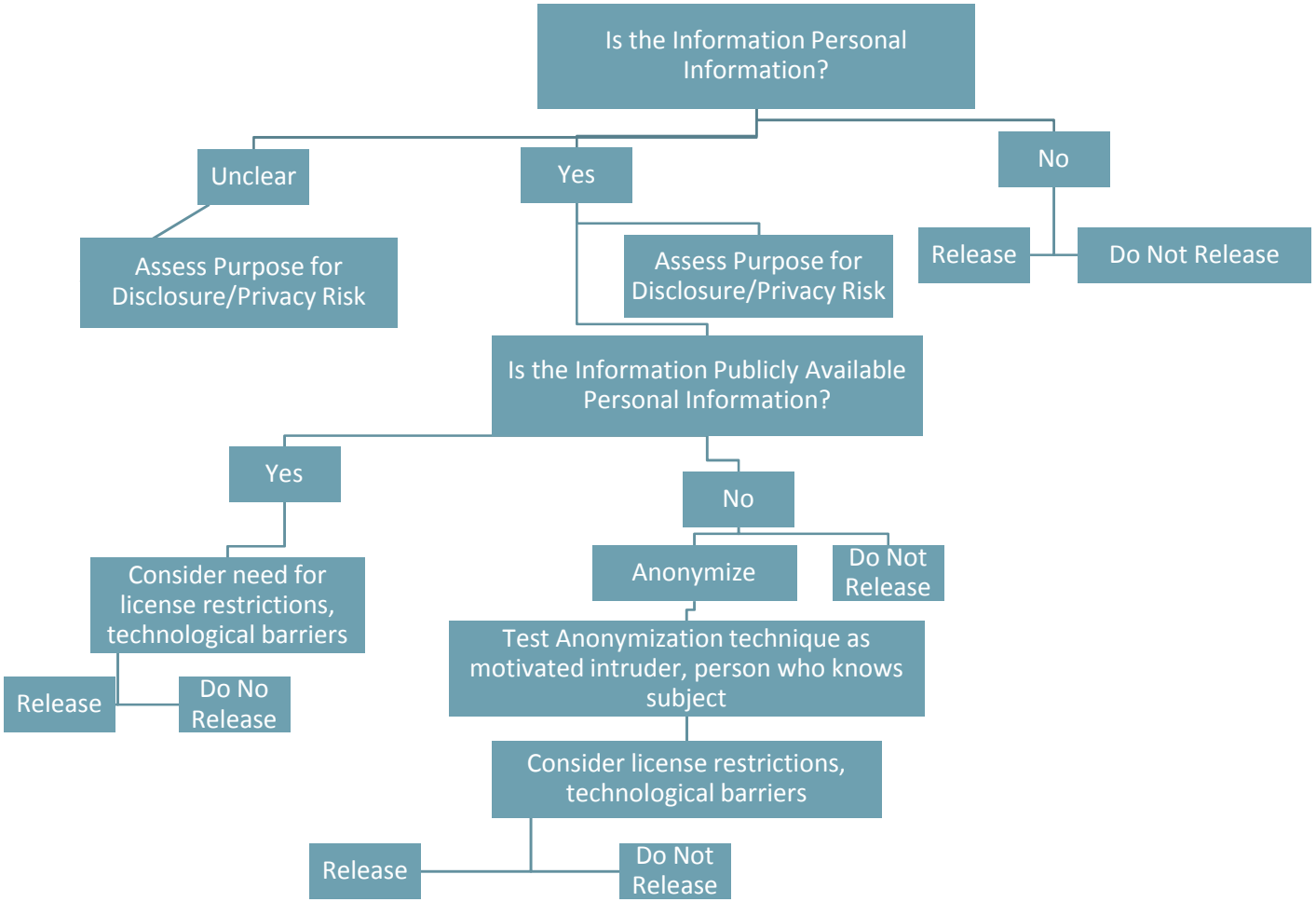
Best practices for balancing privacy and transparency in open government

3. Releasing information/data

- Anonymization: use one of several techniques including aggregation, randomization, pseudonymization, redaction, etc.
- Licence restrictions: ODL with restriction on reidentification of data; terms of use that prohibit scraping of data
- Technological barriers to reuse: restrictions on indexing of websites; choice of file format

Decision tree for digital release of personal information

(from: Scassa & Conroy, "Releasing Information in the Developing Open Government Environment: A Best Practices Guide" (forthcoming 2016)



References

- T. Scassa, “Privacy and Open Government”, *Future Internet* 2014, *6*, 397-413 (http://www.mdpi.com/1999-5903/6/2/397)
- A. Conroy & T. Scassa, “Promoting Transparency While Protecting Privacy in Open Government in Canada”, (2015) 53:1 Alberta Law Review 175-206 (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2658620)