

InterPARES Trust Project Report



Title and code:	NA08 The Implications of Open Government, Open Data, and Big Data on the Management of Digital Records in an Online Environment
Document type:	Final report
Status:	Final
Version:	1
Research domain:	Access
Date submitted:	
Last reviewed:	
Author:	InterPARES Trust Project
Writer(s):	Jim Suderman, Kat Timms
Research team:	Grant Hurley, Valerie Léveillé, John McDonald, Kelly Rovegno, Jim Suderman, Kat Timms

Document Control

Version history			
Version	Date	By	Version notes
0.1	20160811	KT	Transformation of previous draft final report to conform to this template.
0.2	20160901	KT	Editing throughout.
0.3	20160906	KT	Reconciliation of previous comments from JS.
0.4	20160918	JS	Addressing comments from JM, GH.
1.0		KT	Final edit and circulation draft to ITrust.

Table of Contents

1. Executive Summary	4
2. Research Team	4
3. Background.....	5
4. Study Objective.....	5
5. Methodology	6
6. Findings.....	6
7. Conclusions.....	13
8. Products.....	17
9. Dissemination	17
10. References.....	Error! Bookmark not defined.
Appendix A: Interview Guide.....	19
Appendix B: Consent Form	22

1. Executive Summary

This study explored relationships between Open Data, Big Data, and Open Government with a particular focus on recordkeeping issues in the online environment. The study team sought to clarify these three concepts, which had not been well understood in a recordkeeping context, with the aims of informing the broader InterPARES Trust research objectives and providing insight and guidance regarding recordkeeping concerns.

The study was divided into three phases. The first two primarily investigated recordkeeping concerns affecting Open Government and Open Data, by conceiving of them as business processes. Big Data was addressed to some extent in the first phase, but was excluded from scope afterwards. The second phase included a broad overview of open government initiatives at all three levels of government in Canada. The third phase focused on recordkeeping concerns within citizen engagement activities of Open Government initiatives, including an analysis of both the International Association for Public Participation's (IAP2) Spectrum of Public Participation and the different contexts of recordkeeping as defined by InterPARES. This phase built on the overview of open government initiatives of the preceding phase, taking a more in-depth look at citizen engagement in particular.

The study's findings highlight trust issues with respect to Open Government and Open Data initiatives, underpinned by gaps observed in recordkeeping infrastructure and operations using the lens of the five essential levels of context identified by earlier InterPARES research. For instance, records retention may not always take into consideration the 'whole' of the record in Open Government citizen engagement initiatives that result in a fragmentation of the record between collaborating parties or among technology platforms used. Further, since many citizen engagement initiatives arguably have limited engagement or collaboration aspects (merely having the government 'Inform' the citizens), the desired effect of increasing trust by fostering more participatory or collaborative governance is brought into question.

In relation to the research objectives of the InterPARES Trust project, the study found that record keeping policy and practice was weak or absent in relation to aspects of Open Government, and particularly in relation to citizen engagement initiatives. Few measures of success were identified for engagement initiatives nor did any criteria seem to be well established within or across jurisdictions reviewed. Two challenges were identified for building and sustaining trust in government. The first being the relatively limited nature of engagement initiatives in the sense that most types of engagement are "Inform" or "Consult" rather than "Collaborate" or "Empower," and the second is the lack of clarity among all stakeholders in terms of what should be expected from Open Government initiatives generally. Conclusions of the study that could contribute to model policies and practices are set out in two publications and a records management primer for citizen engagement initiatives.

Effective GCE initiatives must be supported by a solid records management infrastructure, one that is built upon reliable standards and practices, enabling technologies, and resourceful individuals that are in turn supported by an effective accountability framework, a comprehensive policy, and a governance structure that is comprised of individuals with a high level of awareness and understanding of the importance of records to the achievement of the goals and objectives of GCE initiatives.

2. Research Team

- Jim Suderman (Lead, City of Toronto)
- John McDonald (Independent Researcher)
- Valerie Léveillé (Independent Researcher; GRA 2013-2014)
- Grant Hurley (Independent Researcher (GRA 2014-2015))

- Kat Timms (Library and Archives Canada)
- Kelly Rovegno (GRA, January 2015-present)

3. Background

This study committed to exploring the relationships between Open Data, Big Data, and Open Government with particular focus on the online environment. While it is true that these three concepts have distinct objectives and communities that support them, they nevertheless have common features that have at their root the interaction between customers/clients/citizens and organizations (private, public, academic) operating in an online environment, and have common (albeit generic) processes or workflows.

These concepts are relatively new phenomena. The first use of the concept of 'Open Government' is generally attributed to Wallace Parks in his 1957 article entitled "The Open Government Principle: Applying the Right to Know under the Constitution."¹ Both of the terms 'Big Data' and 'Open Data' are barely twenty years old, the latter having first been used in 1995 referring to the need to openly exchange "scientific information between countries [as] a prerequisite for the analysis and understanding of these global phenomena [i.e., the atmosphere, oceans and biosphere]."² The origins of the term 'Big Data' are less certain.³

These concepts are not well understood, especially from a recordkeeping perspective in terms of their evolution, their characteristics, and their future directions. Recordkeeping issues, especially as they pertain to these three themes operating in an online environment, are only just being identified and have yet to be addressed in a substantive manner. The rationale for undertaking this study was to start addressing this gap in knowledge and understanding.

4. Study Objective

The overall objective of the study was to clarify the concepts of 'Big Data', 'Open Data', and 'Open Government' and to identify records-related issues in order to support the establishment of appropriate InterPARES Trust research projects to address the issues. The study proposal noted that the three concepts are not well understood generally and in particular among the recordkeeping community. As such, a first objective was to define and clarify these concepts to support research addressing the creation and management of digital records generated in these environments. Another objective, requiring the identification of the issues in these environments, was to determine if there were common recordkeeping issues among them.

Given the early focus in the study on Open Government projects, which typically included an Open Data component, not much attention was given to Big Data. The latter part of the study shifted to focus on citizen engagement activities within Open Government initiatives, further deferring attention away from Big Data. Late in this study, the decision was made to permanently defer exploring recordkeeping issues in the Big Data context, leaving the possibility open for a separate InterPARES Trust initiative to investigate it. Some initial work exploring recordkeeping issues pertaining to Open Data and Big Data had been undertaken in another InterPARES Trust Study (NA09), which is further explained in the 'Findings' section below.

¹ *George Washington Law Review*, 26 (1).

² Simon Chignard, "A Brief History of Open Data," *Paris Tech Review* 29 March 2013, accessed on 31 August 2016, <http://www.paristechreview.com/2013/03/29/brief-history-open-data/>.

³ Steve Lohr, "The Origins of 'Big Data': An Etymological Detective Story," *The New York Times* (blog), 1 February 2013 (9:10am), accessed on 31 August 2016, http://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/?_r=0.

5. Methodology

The project proposal identified the following study methods:

- literature review and analysis to contribute to the NA01 Terminology project;
- analysis of evolution and relationship of the major themes of the project (Open Government, Open Data, Big Data);
- identifying recordkeeping issues related to the themes (individually or shared in common);
- developing research questions or initiatives to lead to the development of frameworks, policies, procedures and standards for recordkeeping and Open Government, Open Data and Big Data; and
- identifying InterPARES Trust partner organizations engaged in Open Government, Open Data and Big Data initiatives to explore collaboration possibilities.

These were extended during the course of the study by the following:

- semi-structured interviews with government employees actively engaged in or responsible for Open Government or Open Data programs in their jurisdictions;
- selection and development of Open Government/Open Data case studies from Canadian jurisdictions;
- cross-referencing the recordkeeping contexts defined by InterPARES against the categories of engagement set out in the Spectrum of Public Participation developed by the International Association of Public Participation (IAP2);⁴
- analysis of technology considerations in Open Government citizen engagement activities, also in relation to the IAP2 spectrum; and
- preliminary comparative analysis of the focus, scope, methods, and outcomes of the NA08 study with other InterPARES Trust projects related to Open Government or Open Data.

6. Findings

The project fell into three distinct phases.

Phase 1

The first phase was actually part of a separate InterPARES Trust study (NA09 Establishing Retention and Disposition Specifications and Schedules in a Digital Environment), but since there was cross-over in both subject matter and team members, the NA08 team has been considering the two studies as overlapping and linked. The findings from the NA09 study were published in an article by John McDonald and Valerie Léveillé entitled “Whither the Retention Schedule in the Era of Big Data and Open Data?”⁵

The article explores the conceptual evolution of Open Data and Big Data, noting some of the differences between the two, such as:

- Open Data initiatives generally focus on existing datasets while Big Data initiatives often create new datasets (mashups); and
- communities supporting each type of initiative differ, with Open Data communities more oriented to public policy-making and governmental transparency, while Big Data communities seek new knowledge or an economic edge.

The article asserts that both Big Data and Open Data initiatives share the characteristic of being based on some form of business process, which is essential to setting appropriate retention and disposition rules, a

⁴ 2007, accessed on 31 August 2016, http://c.ymcdn.com/sites/www.iap2.org/resource/resmgr/imported/IAP2%20Spectrum_vertical.pdf.

⁵ In *Records Management Journal* 24, No. 2(2014): 99-121.

task which is itself an evolving process as a result of the transition from physical to digital recordkeeping. In addition, records managers have a leading role in establishing retention and disposition rules that are implementable and relevant within the organization and to the communities served by data both big and open.

The article uses a fictitious business organization to illustrate the organizational dynamics affecting retention management of data, ranging from the priority of publishing Open Data to a 'Why not keep everything?' approach. Concluding that retaining everything was neither desirable nor technologically feasible, the illustration points to the ISO Technical Report, *Information and Documentation – Work Process Analysis for Records* (ISO/TR 26122:2008) as a guide to analyze the relevant business processes.

The article notes that a business process does not start and stop with the creation and dissemination of a dataset. It extends back into the processes generating the source data (which could come from a variety of sources) that contributed to the final data set.

It also includes the steps involved in administering the processes themselves such as the approval of data collection methods, conduct of the collection/survey, measurement of data quality, etc., all of which generate records in the form of emails, reports, etc.

The article concludes that retention and disposition specifications need:

- to account for those records that are necessary to ensure the accuracy and authenticity of the data generated to support Big Data and Open Data initiatives; these can include project management documentation, data collection procedures, metadata models and related data classification documentation, survey design documentation, data quality assessment procedures and tests, and related documentation, as well as key versions of data used to generate master files, public use data files, and other data made available through Big Data and Open Data initiatives; and
- to have been developed in an integrated fashion as part of the steps involved in designing and implementing the systems, processes and technologies that support Big Data and Open Data initiatives.

As a result, it is not seen as sufficient for the records manager to simply identify the retention period, but rather to set out the specifications for retention in a comprehensive manner that accounts for all related processes. Concluding observations emphasize the need for a more robust recordkeeping framework that accommodates data generated both internally and externally.

Phase 2

The second phase of the study began by defining the concepts of Open Government and open government data and information in the Canadian context. The study also considered the impact of these concepts on the creation, management and control of government records through business process analysis as defined by ISO/TR 26122:2008, the same standard relied upon in the first phase of the study, detailed above. Open Government initiatives active in Canadian jurisdictions were surveyed based on publicly available information sources.

The findings from the second phase of the study were published in an article by Valerie Léveillé and Katherine Timms entitled "Through a Records Management Lens: Creating a Framework for Trust in Open Government and Open Government Information."⁶ The article proposes a universal framework to ensure the accuracy, authenticity, and trustworthiness of information through the analysis of business processes, documentation, and structures supporting Open Government initiatives.

⁶ In *Canadian Journal of Information and Library Science* 39, No. 2(2015): 154-190.

More specifically, the article considers Open Government as a three-staged business process:

1. Initiation
2. Identification and distribution
3. Promotion and evaluation

A records management framework emerges from an Open Government policy that includes records management components, either integral to the policy or by reference to relevant records management policies. To give strength to the policy, responsibility for its implementation must be appropriately assigned organizationally or overseen by records professionals collaborating with authorities administering access to information, internal regulation, and audit functions.

An enterprise-wide inventory of government information provides a starting point for determining what information – including relevant contextual metadata – is eligible for publishing. While new technologies may have to be deployed to make this information open and accessible to the public, ‘feedback loops’ should also be established to allow the Open Government authority to determine whether citizens are aware of available information and whether or not this information is useful to them.

The article identifies six categories of ‘information objects’ that result from Open Government initiatives, specifically those that:

1. are generated to initiate and sustain the program, e.g., action plans;
2. serve as inputs to the process of releasing information, e.g., candidate data sets;
3. are outputs of the processing of releasing or publishing information;
4. are platforms for accessing released information, which may contain, for example, metadata of released information;
5. are inputs from citizens, e.g., completed feedback forms; and
6. are related to the initiative(s), e.g., reports, presentations.

Maintaining this documentation is an important records management task that provides an authoritative, complete, and accurate source to demonstrate the integrity and reliability of information published as part of the Open Government or Open Data initiative.

The article concludes with a brief assessment of Canadian jurisdictions against the framework outlined, highlighting:

- a general need for more substantive integration of Open Government with records management (looking at organizational recordkeeping from an enterprise perspective, Léveillé and Timms note that in only two of seven jurisdictions examined was the link between Open Government and information management clearly evident);⁷
- better citizen engagement mechanisms, including feedback loops; and
- meaningful metrics to evaluate Open Government initiatives.

Phase 3

The study team struggled with establishing a definition of ‘Open Government’, since it is a term that continues to evolve. In the end, the four core principles of the Open Government Partnership (OGP) were adopted as the definitional scope for the term in the context of this phase of the study:

- Accountability
- Technology and Innovation
- Civic Participation
- Transparency⁸

⁷ “Through a Records Management Lens”, p. 179.

⁸ Open Government Partnership, “What is the Open Government Partnership?”, 2015, accessed on 6 September 2016, <http://www.opengovpartnership.org>.

Partly because of the challenge of defining Open Government and partly because it appeared to be uncharted territory, the study team narrowed its focus from Open Government generally to that of citizen engagement. Citizen engagement actively involves citizens in policy or program development at any level of government. Citizen engagement processes and initiatives seek to establish a more collective basis of decision-making power by seeking input from citizens much earlier in the process of drafting legislation, for example. Though the level and opportunities for citizen input and technologies used may differ between different engagement initiatives, knowing the contexts, functions, and technologies deployed will enable more informed recordkeeping.

The five engagement categories described by the International Association for Public Participation's (IAP2) Spectrum of Public Participation are as follows:

- Inform
- Consult
- Involve
- Collaborate
- Empower

These were adopted to provide a means of categorizing citizen engagement initiatives identified through semi-structured interviews with Open Government leaders in six Canadian jurisdictions: the federal government, the provincial governments of Alberta, British Columbia, and Ontario, and the city governments of Toronto and Vancouver.

Findings were grouped under five context-based headings identified by the InterPARES research project as essential to preserve the identity and integrity of digital records.⁹

1. The juridical or administrative context, defined by InterPARES as the “legal and organizational system in which the [record-]creating body belongs”;
2. The provenancial context is that of the “creating body, its mandate, structure and functions”;
3. The procedural context refers to the processes that result in created records (the “business procedure in the course of which a record is created”);
4. The technological context comprises the “characteristics of the hardware, software, and other components of an electronic computing system in which records are created”; and
5. The documentary context is the “archival fonds to which a record belongs, and its internal structure”.

Juridical / Administrative Context

Generally speaking, there are three common juridical contexts within Canada:

- the federal or national context;
- the context of each of the provinces or territories; and
- the context of every municipality.

These contexts influence and are influenced by each other as well as by other juridical environments, e.g., international organizations. And although most Canadians act simultaneously in all three juridical contexts, the governmental entities in each case operate their own Open Government initiatives, with limited integration across jurisdictions. This is illustrated by the similar but not identical Open Government licenses prepared by the different levels of government in Canada. *Open Government Licence – Toronto* is “based on version 1.0 of the *Open Government Licence – Ontario*, which was developed through public consultation and a collaborative effort by the provincial and federal

⁹ Luciana Duranti, and Pat Franks, eds., s.v. “Context,” *Encyclopedia of Archival Science* (2015): 152. Quoted definitions are taken from: *InterPARES 2 Project Terminology Database*, accessed on 31 August 2016, http://www.interpares.org/ip2/ip2_terminology_db.cfm.

government.”¹⁰ Besides replacing references to the Province of Ontario with the City of Toronto, the latter includes “a provision for the *Ontario Personal Health Information Protection Act, 2004*.”¹¹

It was also noted that leadership at all levels of government is assigned to senior elected officials suggesting that Open Government initiatives are seen as important to legitimizing the authority of elected governments. For example, the President of the Treasury Board, an elected official, has leadership at the federal level, while the Deputy Premier, also an elected official, has leadership of the province of Ontario’s Open Government initiatives. Toronto’s City Council established “Open Government by Design” and “Engage the Public” as two strategic directions for the City, assigning leadership of the former to the City Clerk and of the latter to the City Manager, both central officers of the City administration.¹²

The relevance of the juridical / administrative context of records to citizen engagement initiatives lies in the meaning of the records created through the initiative. In cases where a government “informs” citizens of a new law or policy, the authority of the records held by government can be assumed to be strong. Where there are cooperative or collaborative initiatives involving non-citizen organizations, e.g., international scientific or human rights organizations regarding climate change or poverty reduction, the authority of the records held by any of the parties needs to be understood in their proper context.

Provenancial Context

Interviewees from the federal government observed that that most engagement initiatives are conducted by individual departments. They were aware of few multi-departmental, collaborative consultations, commenting that there is no culture of sharing between departments, partly as a result of privacy concerns and partly because doing so requires planning and guidance which is absent. One result of this situation is that separate engagement initiatives conducted by different departments may go to the same stakeholders with similar requests. Another example are organizations which have a central office for leading engagement initiatives even though execution of such initiatives is a distributed responsibility, e.g., individual government ministries. Responsibilities of each must be clear to records managers so that the records are maintained in their proper context and the correct authorizations are in place to schedule them and implement disposition. Many Canadians are not aware of the roles and responsibilities of different levels of government. One interviewee commented that web renewal plans include ‘scoping down’ online content to be more citizen-centric and less organization dependent. This has advantages for the citizen but may obscure provenance.

It also appears that provenancial context is complex at least for engagements falling into the ‘Collaborate’ and ‘Empower’ categories of the IAP2 spectrum. In these cases there is more than one creating body, each with its own mandate, structure and functions presumably engaging as more or less equal partners and where recordkeeping requirements and practices may vary substantially. One interviewee asked us who the steward was for information collected by both government and non-government groups. Establishing clear stewardship for such records is crucial to determining what legal obligations and rights pertain to them.

The relevance of the provenancial context lies in the authority of the records as representative of the obligations and competence of the participating organizations and individuals. For example, in a cooperative engagement initiative, the records held by one of the parties can be presumed to represent that organization’s scope of authority and competence and set out the organization’s obligations in the context of the initiative.

¹⁰ Toronto, *Open Government Licence – Toronto*, updated 13 August 2013, accessed on 31 August 2016, <http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=4a37e03bb8d1e310VgnVCM10000071d60f89RCRD>.

¹¹ *Ibid.*

¹² Toronto, *Strategic Actions 2013-2018*, accessed on 31 August 2016, https://www1.toronto.ca/City%20Of%20Toronto/City%20Manager's%20Office/City%20Manager%20Profile/City%20Initiatives/Strategic%20Actions%20for%202013%20to%202018_FINAL.pdf.

Procedural Context

Study findings indicate a recent trend towards centralization of organizational policy, performance measures, and practices, very likely linked to the centralization of responsibility for Open Government initiatives, including their citizen engagement components, noted previously. Nevertheless, citizen engagement policies and procedures appear to be still very much a work in progress.

- The liquor policy review in British Columbia was the only instance encountered that had a clearly established engagement process, although it appears that process was unique to this consultation and did not follow a well-established, common process.
- “Growing Conversations,” a Toronto engagement initiative to enhance citizen engagement practices:
 - instituted a “Conversation Corps” (drawing from practice in Austin, Texas);¹³
 - produced a guide prepared by the City Planning division.
- Communication and consultation guides exist in some departments in the Canadian government, and may be sufficient to govern simpler types of engagement.
- In Vancouver there is currently no suite of policies for Open Government, nor are there formal, mapped-out workflows for Open Data. The development of policy pieces, guidelines, etc. – including those for citizen engagement – will be addressed in the next phase of the program.
- The absence of whole-of-government guidance on conducting consultations inhibits coherence, consistency and the development of standards. The absence of standards, guides, etc. was recognized as a major issue by several interviewees (e.g. Treasury Board Secretariat and Privy Council Office).

Interviewees identified even fewer performance measures.

- One jurisdiction, having conducted an environmental scan, concluded that there were not many Open Government jurisdictions with mature performance measurement frameworks against which to benchmark.
- It is hard to establish effectiveness measures even for the simplest type of engagement on the IAP2 spectrum, ‘Inform’. Measures may be increasingly difficult to establish for the more complex types of engagement.

In terms of record keeping, the procedural context sets out the normal course of business of the participating organizations. The procedural context provides essential authentication of the records purported to represent that activities, obligations, and position of the organization or individual involved. That authentication function of this aspect of records’ context is weakened where processes are only partially described or non-existent. Procedural contexts are, of course, dependent on the provenancial context.

Technological Context

All citizen engagement initiatives identified used some kind of web-based technology to inform citizens, collect input, or offer resources related to Open Government, such as Open Data. Besides conventional technologies, such as meeting rooms and snail mail, citizen engagement technologies ranged from:

- Use of commonly available web technologies, such as blogs, e-mail and surveys;
- Third-party social media and collaboration platforms, such as Twitter, LinkedIn, and Google Docs;
- Adapted third-party social media platforms built for citizen engagement, such as MindMixer (now mySidewalk) which supported Toronto’s IdeaSpaceTO – a place where “you can apply your ideas,

¹³ This program allows individuals (e.g. private citizens) to sign-up as facilitators to lead consultations on topics pertaining to government business.

- knowledge & expertise to help shape the future of Toronto [and] explore and comment on the ideas of others within the community.”¹⁴
- Data access platforms built in-house, such as the CKAN-based (open source) portal planned by Alberta, where people will be able to search for information and data over many databases.

What constitutes an effective technology depends on the type of engagement initiative. A blog may provide better support for a collaborative engagement than an email blast, for example. However, engagement may depend more on a suite of technologies, as is evident from the statistics for the consultative review of BC's Liquor Policy, which used e-mails, Twitter, and blog posts to encourage feedback. The trend appears to be moving toward more substantial uses of technologies to encourage feedback and conversation from many different sources, or as in the case of Toronto's immersive engagement platform, the deployment of technologies specifically intended for citizen engagement processes.

The study also noted that while a range of technologies may play a key role in engaging citizens, initiatives that encourage inputs from many sources may have the effect of fragmenting the record of engagement, where records documenting inputs may be spread across many platforms, some of them outside of the control of record keepers. Furthermore, one or more technologies used may cease to exist as was the case with the IdeaSpace platform used by Toronto. The proprietor substantially over-hauled the technology with the resulting platform no longer meeting the needs of the City. Toronto confirmed that the data has been extracted from the obsolete platform but no details of how it can be accessed or what context survived the extraction were available. Where knowledge of the technology(ies) used is not available, whether it is used as a public-facing platform to enable citizens to have their say or used at the back end to analyze the data collected and represent it to decision-makers, it may not be clear how citizen feedback is actually used to inform the end result. Records attesting to how technologies were used as part of the engagement process, e.g., algorithms used to create data visualizations, help create greater transparency about the initiatives and any resulting decisions or actions taken.

Documentary Context

That records are essential to effective citizen engagement initiatives is hardly a surprising premise. Using technologies that enable citizens to create records themselves seems to contribute positively to engagement.

As noted above, the record of an engagement initiative may be fragmented as a result of different provenancial or technological contexts. For example, free text comments on a blog posting may be difficult to easily compare against tweets or re-tweets on Twitter. Detailing that contextual information in such a way that is intelligible to individuals examining the records subsequent to the initiative will be a significant challenge to record keepers. They may not even be aware of related contexts, which may create uncertainty regarding where the complete record resides and whether fragments held in different places can be understood as a cohesive whole.

A further consideration may be the existence of records that are not relevant to the engagement initiative but that come into existence as a result of a technological platform that allows 'free' posting. For example, even if the technological context is that of the government, not all records held on that technological platform may be addressed to or authored by it; two participants may conduct a sidebar conversation in the same discussion forum intended for a particular issue.

¹⁴ Toronto, "Ideaspace", currently available only via the Internet Archive, accessed on 31 August 2016, <https://web.archive.org/web/20160317024450/http://toronto.mindmixer.com/>.

Even basic records like a report resulting from an engagement initiative may have significant variations. One interviewee advised that the composition of reports on consultations will range from verbatim to highly synthesized – even sanitized – content.

7. Conclusions

The InterPARES Trust Research objectives are:

1. to discover how **current policies and practices** regarding the handling of digital records by institutions and professionals affect the public's trust in them, in light of the exponential growth of and reliance on Internet services;
2. to anticipate problems in maintaining any trust in digital records under the control of entities suffering a **waning level of confidence** from the public (including legal, law enforcement, financial, medical, broadcasting, "hacktivist," and governmental organizations and professionals);
3. to establish what **significance national/cultural contexts** have with regard to the level of trust digital records on the Internet enjoy;
4. to **articulate model policies, procedures, and practices** for creating, managing, accessing, and/or storing records on the Internet, especially in social media and cloud computing environments and through mobile technologies, and test them in a variety of contexts so that, from them, international standards, guidelines and best practices can be developed, and
5. to formulate proposals and models for law reform, and functional requirements for the systems in which Internet providers store and manage digital records.¹⁵

What follows are the findings of this study in relation to the above, with exception to the fifth objective, for which there are none.

Current Policies and Practices

The first phase of this study showed that current records retention scheduling practices appear to be increasingly inadequate for government records, particularly where those records take the form of structured data which may itself be published by the government in whole or in part as Open Data. Also, data relied on by government may be purchased and as such are subject to contractual terms which may not be consistent with other statutory, financial, or administrative considerations when establishing retention.

Conversely, one of the purposes of establishing and implementing records retention periods is to limit organizational liability by not retaining records longer than they are needed by the organization. If governmental records are also published as Open Data, that liability mitigation may be limited or eliminated entirely. Open government, as represented by the OGP's four guiding principles (accountability; citizen participation; technology and innovation; transparency), may mean that setting retention requirements to meet statutory and operational needs may not be sufficient in an Open Government environment. No longer is governmental accountability the pre-eminent consideration for recordkeeping. What are the recordkeeping requirements for the other three principles? Will meeting the recordkeeping requirements for one of the principles be at odds with those of another principle? For example, recordkeeping practices to meet accountability requirements may require the government to ensure it is able to manage all of the records it needs to demonstrate accountability. However, doing so may conflict with the priorities of collaborative and empowering citizen participation, where key records may be held outside of the government, e.g., by collaborating organizations. The need to balance priorities when setting and implementing retention rules may change the risk profile of government. What are its obligations for full accountability in an environment of shared responsibility? Should the records manager's responsibilities remain defined by the boundaries of the organization that employs

¹⁵ InterPARES Trust, *Goal and Project Description*, p. 3.

them or are collaborative or partnering agreements among records managers needed with external organizations which are engaged to ensure that records are represented in their proper contexts?

In fulfilling Open Government objectives, governments are also utilizing on-line, third party products and social media services, such as Twitter, particularly in pursuit of encouraging and enabling citizen participation. There are ways to copy such content and retention rules could be established for such copies but this study did not discover any governments that were doing among those studied. And while copying such content has the benefit of enabling the government to reproduce it, the government must take care to include consideration of how the original context of the content may have affected how it was understood and the weight it received as a result. Perhaps more importantly, the government may have disposed of such copied content through conventional retention scheduling practices and yet it may continue to be available through the originating service or through collection practices of other organizations. For example, the Library of Congress has pledged to acquire all public tweets posted on Twitter. Clarifying record keeping expectations may benefit from appraisal practices in terms of the value of copies of records when removed from the technological context of their creation and use.

Waning Level of Confidence

This study focused on Open Government initiatives as a particular type of recordkeeping context and whether citizens trusted engagement processes and technologies in order to engage in establishing government policy. The information it gathered came primarily from Canadian government employees, many of whom were pioneering or at least tasked with leading Open Government initiatives in their jurisdictions. With this in mind, the study team sought information about measures that governments were relying on to determine whether engagement activities enjoyed the confidence and trust of the participants.

The case studies outline instances where those conducting the engagements were confident about the value of the initiatives based on the levels of response from the public. While this study did not find any reason to doubt such confidence, it also did not find any independent confirmation of the success of the initiatives. Furthermore, even assuming that trust and confidence increased in the context of the specific initiatives, there is no way to know if trust and confidence levels were raised overall, that is, whether other actions of government conducted at the same time reduced or increased trust and confidence.

A clear gap was identified in terms of an absence of policies and procedures for conducting citizen engagement activities. While some consistency was provided by communication policies, communication is only one element of citizen engagement initiatives, although it is an important one. A particular challenge identified by interviewees was coordinating engagement initiatives by two or more departments within a government or by two or more levels of government. This challenge goes beyond recordkeeping concerns, of course, but the study has prepared a primer or guide to recordkeeping considerations in citizen engagement initiatives with the intention of reducing the gap identified.

One major challenge to building and sustaining trust in government in Canada is that Open Government practice is driven by limited initiatives, usually selected by government with limited, if any, input from the public on what matters the government may seek to engage the public, and what degree of engagement the government will enable. The study relied on the IAP2 Spectrum of Public Participation which emphasizes that the level of impact on the public increases from a minimum ('Inform') to a maximum ('Empower'). One interviewee from the federal government stated that final decisions still reside with government and that 'Inform' engagements remain 'classical' for communicating government messages.

A second major challenge is effectively communicating what should be expected from Open Government. Antti Halonen notes that there are both internal and external benefits to Open Government – although his primary focus is on Open Data. The former include “better efficiency, increased internal understanding of the organisation’s work and objectives and releasing of resources into more productive tasks” while the

latter include “supporting entrepreneurship, empowering citizens, and democratic accountability and participation.”¹⁶ It will be a tall order to satisfy the expectations of all for each of these benefits. Using the consistent format and definitions required by the OGP, Mary Franco and Amanda Clarke compared the action plans of seven nations, all members of the OGP, concluding that access to information, which they argued was the core concept of Open Government when the term was coined in the late 1950s, remains a dominant benefit; however benefits derived or expected from increased public participation and improved public services are growing in importance over time.¹⁷ And while they concluded that there is little variation based on geography, they noted that action plans of North American and European countries referenced reuse of government information, economic growth, and alternative service delivery while plans of other countries did not. Conversely, the action plans of North American and European countries did not reference prevention of corruption, which was included in the action plans of Azerbaijan, Brazil, and Kenya.

Part of the challenge is figuring out who gets to determine what the focus of Open Government will be and what benefits are most desirable. If widely available information is the focus of Open Government, will that reduce government corruption or make government officials more vulnerable (or willing) to taking bribes? Is the focus of Open Government something that government should determine and just advise citizens accordingly? What are the checks and balances on setting openness priorities? For example, does or should technology enable more direct participation in setting governmental priorities – possibly limiting the role of elected representatives and the reliance of the public service on expert staff? More specifically in terms of record keeping, what records should be created in the planning, execution, and retained for reference or action following engagement initiatives? Without a clearer understanding of what constitutes the ‘normal course of business’ of an engagement initiative, records managers have few grounds on which to understand the records lifecycle or how records should be maintained to sustain whatever growth in public trust was achieved through the engagement.

Significance of National and Cultural Contexts

The information gathered in this study was limited to the Canadian context, albeit from all three levels of government, and as such its conclusions regarding national and cultural contexts is limited. Even within the Canadian context, no information was gathered from the province of Quebec or from the governments of First Nations.

Some variation in cultural contexts were detected relating to the scope of Open Government, which is broad enough that different jurisdictions set different priorities for Open Government initiatives. For example, the city governments of Vancouver and Toronto appeared to emphasize more collaborative and empowering citizen engagement initiatives than the federal government, for instance, about which one interviewee observed that public opinion research is tightly controlled and that there are in fact communication policies in place that have the effect of inhibiting government departments from asking citizens for their opinions. This reinforced an earlier observation to the effect that the ‘Inform’ type of engagement activity predominates. One explanation for this distinction between different levels of government may be how directly municipal governments deliver services to citizens compared to the federal government.

The various jurisdictions contacted through this study had several things in common, including the development of action plans for opening up government. Action plans are a central obligation of

¹⁶ Antti Halonen, “Being Open About Data: Analysis of the UK Open Data Policies and Applicability of Open Data,” (London: Finnish Institute in London, 2012), p. 22, accessed on 31 August 2016, <http://finnishinstitute.cdn.coucouapp.com/media/W1siZiZlsljwMTMvMDEvMzEvMTNfMDNfMTNfNjM1X2JlaW5nXzlw b3B1b8yMGFib3V0XzlwZGF0YS5wZGYiXV0/being%2520open%2520about%2520data.pdf?sha=32c50d21c5cabb01>.

¹⁷ Amanda Clarke and Mary Franco, “What’s in a Name? A Comparison of ‘Open Government’ Definitions across Seven Open Government Partnership Members,” *eJournal of eDemocracy* 6, No. 3 (2014): 248-266. Accessed on 31 August 2016 <http://www.jedem.org/index.php/jedem/article/view/227>.

members of the OGP, which is limited to national governments,¹⁸ and it is likely that the provincial and municipal action plans are at least notionally comparable. Action plans of member states of the OGP are independently reviewed to assess the implementation and effectiveness of the commitments stated in the plans. Governments at all three levels are seeking to establish metrics to demonstrate openness but these are still in early stages. Likewise, it appears that leadership of Open Government is being centralized at all three levels of government and this has made visible the absence of a policy and procedural framework within each jurisdiction for how Open Government initiatives are conducted, communicated, and reported.

It seems likely that different levels of government, e.g., local versus national, are subject to different drivers and challenges in terms of public trust. For example, public trust in local government may be based primarily on citizen satisfaction with how services are delivered while public trust in national government might be based more on broader policy issues, like immigration. If so, then this reflects some of the observations made above relating to juridical / administrative and provenancial contexts as well as influencing what records need to be created and maintained at each level of government.

Articulation of Model Policies, Procedures and Practices

The study's conclusions regarding model policies, procedures and practices are detailed in two publications and a third product, a records management primer for citizen engagement initiatives. The challenges of managing retention of Open Data and Big Data were addressed in the previously mentioned *Records Management Journal* article by John McDonald and Valerie Léveillé.¹⁹ This article emphasized the importance of the business process in resolving issues relating to data retention and disposition. The role of the business process was explored in depth in a subsequent article, the aforementioned "Through a Records Management Lens: Creating a Framework for Trust in Open Government and Open Government Information" by Valerie Léveillé and Katherine Timms²⁰ that explored establishing a universal framework supporting the distribution of accurate, authentic, and trustworthy information in the context of Open Government initiatives. Building on these articles and further research into Open Government initiatives, with a particular focus on citizen engagement, McDonald, Léveillé, and Grant Hurley developed a records management primer, *Managing Records of Citizen Engagement Initiatives*, intended for both records management professionals and governmental citizen engagement leaders.

This primer also benefited from previous papers prepared by study group members:

- "The Contexts of Records and the Spectrum of Public Participation" (2016) presented an analysis of the relationship between the five InterPARES records contexts and the five engagement categories of the IAP2 Spectrum.
- "Contextualizing Technologies for Citizen Engagement: Seeking the Records and Supporting Transparency" (2016) sought to address the current gap in knowledge on the dynamic relationships between technologies, governments and citizens that create records. It combines as its framework for analysis the IAP2 Spectrum for public participation and archival diplomatics. It concludes with a discussion of the implications of technologies in managing the records created by citizens and their governments as they interact, communicate and collaborate.

Limitations and Next Steps for Further Research

The study identified record keeping issues and policy gaps in the domains of record keeping as it relates to Open Government initiatives and Open Government policies and practices as they relate to building trust and engaging citizens. The ways in which identified issues and gaps might be resolved are found in the *Managing Records of Citizen Engagement Initiatives: A Primer*, referenced in section 8, below. The study

¹⁸ The OGP has recently instituted a subnational government pilot program (accessed on 31 August 2016, <http://www.opengovpartnership.org/how-it-works/subnational-government-pilot-program>).

¹⁹ "Whither the Retention Schedule in the Era of Big Data and Open Data?" 24, No. 2 (2014): 99-121.

²⁰ *Canadian Journal of Information and Library Science* 39, No. 2 (2015): 154-190.

did not conduct workshops or consultations to resolve them and this is clearly an opportunity for further research. It also means that the conclusions stated in this report are necessarily tentative.

As mentioned, Big Data was not explored during the NA08 study, although it was included in the *Records Management Journal* article published as an outcome of the related NA09 study. Perhaps a future InterPARES Trust study will tackle Big Data, building on this initial foundation as well as relevant related NA08 study outcomes.

Other potential avenues of further research:

- Comparison of this study's findings with other InterPARES Trust studies related to Open Government and Open Data;
- Comparable investigation of other facets of Open Government including characteristics of recordkeeping that relate to transparency (as distinct from accountability) and innovation;
- A more international focus on Open Government citizen engagement initiatives and their implications for recordkeeping.

8. Products

Hurley, Grant. "Contextualizing Technologies for Citizen Engagement: Seeking Records and Supporting Transparency." Forthcoming paper at the Trust and Transparency in E-government international conference, Oslo, Norway, October 26, 2016.

Hurley, Grant, Valerie Léveillé, and John McDonald. *Managing Records of Citizen Engagement Initiatives: A Primer*. InterPARES Trust, 2016.

McDonald, John. "The Contexts of Records and the Spectrum of Public Participation." Research Paper. InterPARES Trust, 2016.

Rovegno, Kelly. "BC Digital Services Consultation Case Study." Case Study. InterPARES Trust, 2016.

----- "BC Liquor Policy Review Case Study." Case Study. InterPARES Trust, 2016.

9. Dissemination

McDonald, John and Valerie Léveillé. "Whither the Retention Schedule in the Era of Big Data and Open Data?" *Records Management Journal* 24, No. 2 (2014): 99-121.

McDonald, John. "Open Data and Open Government: Implications for the Management of Records." Paper presented at the 1st InterPARES Trust International Symposium, Stockholm, Sweden, May 14, 2014.

Léveillé, Valerie and Katherine Timms. "Through a Records Management Lens: Creating a Framework for Trust in Open Government and Open Government Information." *Canadian Journal of Information and Library Science* 39, No. 2 (2015): 154-190.

Léveillé, Valerie. "Open Government: an Ideal or a Bad Deal." Paper presented at the Association of Canadian Archivists' Annual Conference, Victoria, B.C., June 2014.

Suderman, Jim. "Contexts of Open Government: Cloud Computing and Recordkeeping." Paper presented at the 4th InterPARES Trust International Symposium, Marburg, Germany, June 8, 2016.

Suderman, Jim. "How Open is Open? Recordkeeping in the Era of Open Data and Open Government. Citizen Engagement Initiatives in Canada." Paper presented at the Society of American Archivists' Annual Conference, Cleveland, Ohio, August 2015.

Appendix A: Interview Guide

Interview Guide | Revised 9 February 2015

A. OPEN GOVERNMENT: Context and Culture

Scope: Questions that address the contexts, rationale, policies, and governance that guide open government programs and initiatives.

1. What is the rationale for your open government initiative?
 - a. What specific objectives or goals have emerged from this rationale?
 - b. Why has your jurisdiction embraced open government as a strategy?
 - c. What policy instruments currently inform or provide guidance on your open government initiative?
2. What is the scope of open government in *[interviewee's jurisdiction]*?
 - a. Which elements of open government currently make up your open government initiative?
 - b. Do you have any additional elements that you are planning for?
3. What is the current governance structure of your open government initiative, including roles and responsibilities?
 - a. How are citizens involved in the governance structure of your open government initiative?
 - b. Is there an agency with responsibility for coordinating open government initiatives within your jurisdiction?
 - i. If so, what is the name of the agency and where is it located?
 - c. Who are the leaders/key people involved in your jurisdiction's open government initiative?
 - i. What is their background?
 - ii. What are their key skills?
4. Who is the intended audience of the initiative?
 - a. Is there a primary audience? A secondary audience?
 - b. Who might be other possible stakeholders, both internal and external to your organization?
5. Are you engaged in any collaborative arrangements or partnerships with external groups?

B. OPEN DATA / OPEN INFORMATION: Process and Procedure

Scope: Questions that address the nuts and bolts of open government initiatives, including open data/open information components, data alteration and data quality.

1. What kinds of data or information are being opened in your jurisdiction?
 - a. In what formats are data/information released?
 - b. How long do they remain open?
 - c. By what means are they released?
2. Who decides what gets open?
 - a. Is there a formal process or workflow?
 - b. How are citizens involved in determining what data or information is opened?
3. How is data and information altered as part of preparation for release?
 - a. Are these changes/processes made known (disclosed) along with the object?
 - b. What kinds of additional metadata is added to the object for release?
4. Will the information/data released as part of an open government initiative be made available in perpetuity online?
 - a. If so, which mechanisms are put in place to support long-term preservation/access?
5. Has data quality been identified as an issue in the release/opening of data?
 - a. Have users demonstrated concern about data quality?
 - b. Have any formal citizen feedback mechanisms been set up to address data quality?
6. What measures are used to track and assess progress towards and achievement of open government goals?
 - a. How is the use of open data information being tracked?

C. OPEN GOVERNMENT: Citizen Engagement

Scope: Questions that address how citizens fit into an organization's open government initiatives, including citizen engagement initiatives, assessment, and trust.

1. Is citizen engagement an aspect of your open government initiative?
 - a. If so, what initiatives exist?
 - b. When and how do citizens become involved in the process?
 - c. What purposes does citizen engagement serve?

2. How do tools and technologies support the goals of the initiative?
 - a. What information objects are created from citizen engagement processes?
 - b. How are these objects used and for what purpose?
 - c. How are these information objects managed (organized, protected, etc.)?
 - d. What information objects do citizens see as important for engagement? The organization? Do these differ, and if so, why?
3. What tutorials, guides, workshops or other methods available to support citizen use of open data or open information?
4. How is trust established, supported and maintained by the open government initiative between information objects and citizens?
 - a. Is trust a stated goal?
 - b. What policies are used to support trust?
 - c. What procedures and processes support trust?
 - d. What tools or technologies used to support trust?
 - e. What information is generated to support trust?
 - f. How do citizen engagement initiatives or processes support trust?
5. What are some key issues or challenges going forward regarding open government initiatives and connections with citizens?



a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA

School of Library, Archival and Information
Studies
University of British Columbia
Suite 470, 1961 East Mall, Vancouver, BC
Website: <http://www.slais.ubc.ca>

Consent Form²¹

Big Data-Open Data-Open Government

Principal Investigator: Dr. Luciana Duranti
Chair and Professor, Archival Studies
School of Library, Archival, and Information Studies
Phone [604 822 2587](tel:6048222587)

Purpose:

The purpose of this project is to clarify the concepts of 'big data', 'open data', and 'open government' and to identify records-related issues in order to support research conducted by InterPARES Trust into issues of trust in digital records in online environments. The interview is intended to gather information about your jurisdiction's open government initiatives generally but it does have a more specific focus on citizen engagement initiatives and the technologies and records resulting from them.

Procedure:

If you are receiving this consent form, you have agreed to be interviewed for this project.

An interview will be conducted between you and the researchers at a time mutually agreed, in person or over the telephone. The interview will last approximately 60 minutes. The interview questions will explore the concepts of open data, big data, and open government, and your experience with them. The questions will be sent to you ahead of the interview, however no preparation is necessary. Your participation is voluntary, and you may end the interview at any time.

If you have any questions or concerns, please ask the researcher, or follow up with the Principal Investigator.

Confidentiality:

Your identity will be kept confidential. There will be no personal information requested prior to, during the course of, or following the interview. All material will be stored in a secure location to which only the researchers will have access, and computer files will be encrypted according to UBC policy, and password-protected. No identifying information will be used in any publication or presentation of research results, unless authorized by your express consent.

Potential risk:

There are no known risks or potential risks from participating in this interview.

Contact for information about the study:

If you have any questions or desire further information with respect to this study, please contact the Principal Investigator, Luciana Duranti at 604-822-2587 or luciana.duranti@ubc.ca.

If you have any concerns about your rights as a research subject and/or your experiences while participating in this study, you may contact the Research Subject Information Line in the UBC Office of

²¹ Version 2: April 23, 2014

Research Services at 604-822-8598 or if long distance e-mail RSIL@ors.ubc.ca or call toll free 1-877-822-8598

Consent:

Your participation in this study is entirely voluntary. Your signature below indicates that you have received a copy of this consent form for your own records.

Your signature indicates that you consent to be interviewed for this study, and to have the interview recorded.

Subject Signature: _____

Printed name of subject: _____

Date: _____