CHAPTER 1

Looking Back toward a "Smarter" Open Data Future

Tracey P. Lauriault

Abstract

Open data is a relatively new practice when compared to the history of data sharing. The idea of sharing government records may have started with the Domesday Book of 1086, or more officially with the 1766 Swedish Freedom of the Press Act (Government of Sweden, 1766), which argued for access to government records; or possibly with the data-sharing principles of the International Meteorological Organization (IMO) in 1873; or perhaps with one of the first international agreements on data sharing, the Antarctic Treaty of 1959 (Secretariat for the Antarctic Treaty, 1959). For Canada, the genealogy of data sharing has its own particularities. What is clear is that open data did not just appear out of the ether; it has a history, and I suggest that it starts with the natural and social sciences. This chapter tells a Canadian open data story from a critical data-studies approach. It conceptualizes open data as a social and technical data assemblage, and traces the genealogies of open access to data and open data in Canada. It argues that open data, and how it is technologically conceptualized, might be too narrow a focus, and instead calls for the adoption of a broader and more integrated openness approach, especially as open data are being subsumed by smart systems or digital twins. The chapter concludes by suggesting that the future of open data requires looking back at the

epistemic groups involved in its creation, overcoming its technological legacy to ensure that when smart systems and digital twins come online, they do not suffer the same fate in terms of quality and a lack of systems thinking. It also suggests that a broader concept of openness be adopted, especially if there is to be an integrated and systems-based approach to smart systems, as seen in the case of the emerging open smart city.

Acknowledgements

This chapter is informed by work and research conducted at the Geomatics and Cartographic Research Centre at Carleton University, funded by the Social Sciences and Humanities Research Council (SSHRC) and the Canada Foundation for Innovation; the SSHRC-funded InterPARES 2 Project at the University of British Columbia; the European Research Council–funded Programmable City project at Maynooth University; and the Open Smart Cities project funded by the GeoConnections (2008, 2012) program of the Government of Canada.

In this chapter, I suggest that open data is a discursive regime, and to better understand it I apply a critical data studies perspective and frame the discourse within a socio-technological assemblage framework (Kitchin, 2014). I then proceed to briefly describe how open data in Canada evolved by tracing its genealogy (Cosgrove, 2001; Foucault, 2003) to demonstrate that epistemic groups, institutions, materialities, and legalities have uniquely shaped this discursive space. This is part of the social-shaping thesis to data and technology, whereby it is understood that data do not exist independently from the context within which they were created, and the systems and processes that produce them (Dalton & Thatcher, 2014; Kitchin & Lauriault, 2014; Iliadis & Russo, 2016). I then suggest that, as large "smart" social and technological systems (Hughes, 1987) are built, such as smart cities, smart grids, or digital twins, for data to remain open it is critical to move beyond the narrow technological conceptions of open data seen in most definitions. Also, I argue that greater attention should be paid to epistemic groups and their subjectivities so as to avoid the past mistakes made with open data, and I imagine the future of open data by situating it in the context of the emerging open smart city (Lauriault et al., 2019). I propose that if we

want open data-driven decision making, we will have to think about openness more broadly; to govern data as more than simply technical objects and, instead, reconceptualize them as open social and technical processes. To conclude, I call for a more political form of citizen engagement, known as technological citizenship, to better govern open data systems (Barney, 2007; Feenberg, 2011). This, I suggest, is especially important to avoid data and technological colonialism (FNIGC, 1998 & 2019; Thatcher et al, 2016; Couldry & Mejias, 2018), which is increasingly being normalized within smart systems, digital twins, and is not addressed in digital strategies.

1. Open Data Social and Technological Assemblage

An assemblage is a theoretical framing of data as a constellation of co-functioning, loosely coupled, heterogeneous elements (DeLanda, 2016), as seen in Figure 1.1. Open data thought of as an assemblage implies that context frames how such data are socially understood in their environment, while technologies, processes, and materialities are the content that perform the tasks of making data open. Open data, because of the component parts of their assemblages, differ from place to place, but as an assemblage they are consistent and known. For example, open data is commonly understood by data formats, licences, standards, and dissemination portals but, it is argued here, should also be about systems and forms of thought. For example, open datasets are also part of new managerialism in government and efficiency discourses and practices, along with principles of transparency and accountability; and a political economy that includes the proactive disclosure of government contracts, procurement, and open corporate registries. Location also matters. For example, in Canada, the Personal Information Protection and Electronic Documents Act (OPC, 2021) regulates how personal data are shared with the private sector, while in the European Union the General Data Protection Regulation (European Union, 2021) governs data protection. Actors and their subjectivities also matter, as they bring different approaches, priorities, skills, and knowledge; for instance, open data are different for a scientist, an app developer, a chief data officer, a company, a government administrator, and an environmentalist.

By looking at open data as a discursive regime, and examining how its constitutive elements have evolved in different places across

time, it becomes possible to imagine what this social and technological assemblage might look like in a smart context. This framing may lead to greater systems thinking, which, I argue, is required when it comes to smart cities, digital twins, and digital strategies.

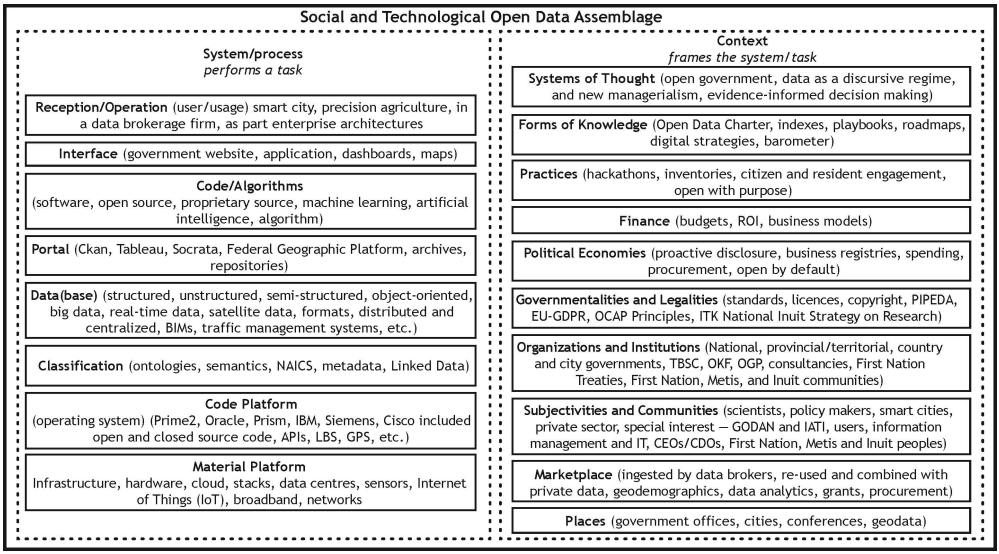

Figure 1.1. Open Data Social-Technological Data Assemblage for Canada. *Source*: Based on Kitchin's Framework 2014.

2. The Legacy of Technical Conceptualizations of Open Data

Critical data scholars, situated broadly in the domains of critical social science and science-and-technology studies, accept that the usual technological conceptualization of data as unbiased, objective, and neutral scientific facts about the world is limited and narrow. Unfortunately, for open data to be qualified as such, they are assessed against these types of criteria, and the legacies of those definitions persist within open data programs. For example, the foundational Open Definition (OKF 2005, 2016/17), a document stemming from the open source and free software movement, is also used to compare and assess open data in the Global Open Data Index (GODI), and was developed by the Open Knowledge Foundation, a global non-profit organization. Applying this definition has resulted in hundreds of thousands of datasets being made open, but without a critical assessment of whether these open data were of good quality or filled any

knowledge gaps, such as equity and inclusion or disability and accessibility. As for data quality, the following elements ought to be considered: lineage or provenance, positional accuracy, attribute accuracy, completeness, logical consistency, semantic accuracy, and temporal information (Guptill & Morrison, 1995). Particularities about data quality vary according to the epistemologies of the domains a dataset stems from and the subjectivities of data authors (Lauriault et al., 2008). For instance, a biologist, astronomer, and spatial data or population health specialist will each have their own specific approach, yet they would most likely agree that knowing and reporting on the quality of the data they use and produce is a requirement of their practice. A lack of knowledge about data quality also precludes the possibility of scientific analysis, and affects the quality of the results; but also, a lack of geospatial or semantic interoperability, and basic standardized framework data, makes linking open data either a very laborious process or nearly impossible. Even though Canada ranked high according to GODI, a lack of data quality or a spatial referent means that important social and political analytical work, such as comparing the outcomes of national health programs or educational achievement, is not possible. Moreover, those kinds of data are rarely found in portals in the first place since these are not the data types listed as being important according to GODI's criteria. During the pandemic, the list of essential datasets a nation state ought to publish also proved to be lacking. In Canada, while data were published, there remained a distinct absence of foundational open data on retirement homes, residences for people with disabilities, and disaggregated equity and health data in general (Lauriault, 2020a). Technical conceptualizations of open data have resulted in many open datasets being available in open data portals, but we cannot create much evidence-informed policy with them because of a lack of quality, interoperability, and data gaps.

The fact that open data indices generally assess data at a national scale compounds the problem. In federations like Canada, where health, education, and cities are jurisdictions of the provinces, data are even less likely to be interoperable or standardized, as the GODI does not assess inter-jurisdictional cooperation or data from a systems perspective. This is unfortunate because there are mechanisms, beyond standardization, that support coordinated efforts. The Geomatics Accord, signed by Canada's provinces and territories and federal government in 2001, is one such example, and this has led to

the promotion and development of trusted key national-framework datasets such as the road network file and a governance structure (CCOG, n.d.). While the Canadian Council on Geomatics is lauded for this collaboration, the focus remained in geomatics and not on framework data related to health regions or socio-economic data, as was seen during the pandemic in Canada. There are also models such as spatial data infrastructures, most notably the Canadian Geospatial Data Infrastructure (NRCan, 2020) and the Arctic Spatial Data Infrastructure (ASDI, 2021), national mapping strategies (OSI, 2017), and scientific-data-producing communities such as the Global Earth Observation System of Systems (GEOSS), among many others, that have well-developed data cultures and policies, practices, legalities, standards, and technologies from which to learn. Furthermore, Indigenous data considerations are also absent from these technical conceptualizations. For First Nation, Métis, and Inuit communities, cosmologies, ontologies, and epistemologies about what constitute Indigenous data, data models, and open data differ from conceptions found in Western notions of empiricism and science. These discourses can be read in works on Indigenous statistics, decolonizing research methodologies, and data sovereignty (Walter & Andersen, 2013; Tuhiwai-Smith, 1999; Kukutai & Taylor, 2016; FNIGC, 2021). For example, the First Nations Information Governance Centre (FNIGC, 2021) states that First Nations

> assert data sovereignty and support the development of information governance and management at the community level through regional and national partnerships. We adhere to free, prior and informed consent, respect nation-to-nation relationships, and recognize the distinct customs of nations.

This includes sovereignty over data in the possession of the Crown but that are considered to belong to First Nations in a post-colonial and reconciliation context. As just described, actors involved with open data in public administrations are different from actors in science, and both differ from Indigenous conceptions. Public administrators assess the qualities of their open data in terms of a set of technical and licence criteria; scientists consider data quality to be a primary concern, and often build interoperable data systems accordingly; while decolonization and data sovereignty, in addition to data quality, are concerns for First Nation, Métis, and Inuit communities.

Each of these actors operate in different contexts and have different subjectivities, resulting in different data practices and conceptualizations, which are often at odds. For example, an open by default approach is incommensurable with the FNIGC's (2021) principles of ownership, control, access, and possession (known as OCAP), since there are data about Indigenous people held by the Crown that should only be made open when there is agreement with the Indigenous communities or peoples they are about.

The Open Data Charter (2015), endorsed by members of the Open Government Partnership (OGP), of which Canada is a member, along with the provinces of Ontario and Quebec as local members, is somewhat better than the original Open Definition (see Chapter 1). The Charter includes six principles that state that data are to be (1) open by default, (2) timely and comprehensive, (3) accessible and usable, (4) comparable and interoperable, (5) for improved governance and citizen engagement, and (6) for inclusive development and innovation. It is a more nuanced approach to GODI and one that factors in institutional arrangements. The fourth principle, that data be comparable and interoperable, is an improvement, but the unfortunate legacy of earlier open data definitions and practices had already gained traction and momentum. The lag time between the adoption of the Open Data Charter and the legacy of institutional practices, combined with open data being administered in a non-systematic way, has resulted in lots of data that are open but which cannot be combined, linked, or compared semantically and spatially. Principles 5 and 6 of the Charter give purpose to the opening of data, and this is where issues of data governance in addition to data for governance ought to be considered. And with the ideals of inclusive development and innovation, values such as fairness, justice, equity, inclusion, and the identification of data gaps and things that are invisible in data ought to also be considered, such as police-shooting data, data on missing and murdered Aboriginal women, or, too, that there are no statistical programs in Canada about people with disabilities. Charter principles include transparency and proactive disclosure initiatives, originating from those involved in access to information, freedom of the press, open contracting, beneficial ownership, and international development. These principles are well suited to the governance of administrative data. The principle of open by default has, however, come under scrutiny of late, as in the context of Indigenous data, and also because it is a labour-intensive process for government

administrators, who often struggle with prioritizing decision-making on what to publish first. The focus is shifting toward publishing with purpose (Open Data Charter, 2018) or, as in the case for a potentially new open government commitment in Canada, Open Data for Results,[1] which aims to mitigate data gaps and invisibilities.[2] Open data programs can also be informed by the experiences of international organizations like the Global Open Data for Agriculture and Nutrition (GODAN), comprised of experts who advocate for and publish data with purpose. GODAN is based in Montréal, and its expert members collaborate to make agricultural and nutritionally relevant data available, accessible, and usable for unrestricted use worldwide, ensuring that the "value chain for agriculture and nutrition is more efficient, innovative, equitable, and accountable" (GODAN 2016).

The Open Data Barometer is another important assessment system. Currently, it evaluates the programs of 30 countries that adopted the Open Data Charter and, for the G20 members, their governments are committed to the G20 Anti-Corruption Open Data Principles (G20, 2015). The Barometer applies the technical criteria of the Open Definition, the technical and policy principles of the Open Data Charter, and the G20 Anti-Corruption Open Data Principles to assess open data readiness in terms of the ability to deliver open data, how that delivery is done, and the political, social, and economic impacts of open data—and it does so quantitatively and qualitatively (Open Data Barometer, 2017). The Barometer is lauded for considering open data more broadly, even though the legacy of the technological criteria for open data persist, as does the lack of attention to data quality. Some organizations, like Open North in Canada, CiviTeo in France, and the City of Ottawa, recognize that technical conceptions of open data are limited, and instead focus on developing data-sharing cultures inside government, as there is also a need to make data

[1] *Open Data for Results* is a new commitment as part of the Government of Canada Open Government, for which the public consultation has just been completed. There has been a general sense of dissatisfaction of the commitments from Government by civil society actors, since what was submitted does not resemble what went out to consultation but some progress has been made nonetheless. https://opengov.konveio.net/open-data-results.

[2] As a member of Canada's Multi-Stakeholder Forum on Open Government, I am one of the proponents for this commitment, which form part of an open government plan for Canada. See https://open.canada.ca/en/multi-stakeholder-forum-open-government.

accessible and interoperable within organizations. This has become increasingly important as cross-organization data analytics teams are forming, in addition to existing geospatial data teams, resulting from the need for an integrated evidence-informed decision-making culture, and also because technical capabilities are increasing. We will see more data analytics teams in cities as more smart systems come online and as a new group of C-level executives (whose title begins with the word "chief") are appointed as a result of digital strategies. These cross-institutional teams are beginning to recognize that legacy administrative data systems in different business units preclude their ability to share data internally, and this is changing how new technologies are procured. For example, the City of Ottawa's Community and Social Services Department collects data in seven different information-management systems, some of which belong to other levels of government, and is now trying to reconcile how it will standardize the collection of equity and inclusion data across these to better inform service delivery, most notably the ongoing public-health response to the COVID-19 pandemic (City of Ottawa, 2021). Technical and data interoperability have been identified as a new priority in how data are produced, as has the need for semantic interoperability if processes are to be automated, such as in the case of artificial intelligence and machine learning (AI/ML). Although not a focus here, another constraint is accessibility, as there are requirements in North America that digital government systems be accessible, which is also the case for open data and their visualization. Currently, this form of accessibility is not assessed by GODI, the Open Data Charter, or the Open Data Barometer, as data about accessibility are not considered to be a key dataset to be published by any of these indicator systems nor is the notion of the accessibility of data for disabled people.

3. Global Data Sharing: A Genealogy

A genealogical approach to understanding the evolution of concepts and practices provides for a deeper analysis of the evolution of the power/knowledge of a discourse (Foucault, 2003). A genealogy historically situates discourse in a specific knowledge-production process. For example, access to data conceptually differs from open data, as it is more about data sharing between a set of specific actors, but it is from this social and technological practice of sharing data that an environment conducive for the emergence of open data exists. The

subjectivities of the data-producing cultures of early actors developing data-sharing practices also differ from those of the open data communities we see today. The former were scientists and data authorities; the latter are the administrators of open data programs and may not necessarily be data authorities nor data owners. The early narrative on access to scientific and spatial data was grounded in systems and infrastructure thinking. Thus, data are part of technical as well as institutional, organizational, collaborative, research, and results-based systems, as data-sharing practices were purpose-driven (e.g., climate modelling). Data here might be in proprietary formats—and may be under a licensing regime, with some restrictions on use by the private sector. This would not be in keeping with open science, where data, methods, techniques, and technologies are open (Foster, n.d.). Nevertheless, data are shared. We need simply think of the multi-billion-dollar Earth observation (EO) community, with its hundreds of public- and private-sector satellite and radar systems circumnavigating the globe, and, within which, data production and sharing is standardized. EO data actors include states, the private sector, and scientific institutions collaborating to share and standardize data toward common goals and for specific purposes, and not simply for the sake of openness. EO principles are about sharing data for sustainable development, resource management, evidence-based decision-making in those areas, and the "benefit of humankind," in somewhat grandiose terms, but also economic viability (GEOSS, n.d.).

Open data, on the other hand, is part of the discursive regime of sharing publicly funded data in the absence of restrictions (OKF, 2005; G8, 2013; Berners-Lee, 2006), epistemically very different from data sharing in science and by Indigenous communities. Open data thinking also coincides with the development of techniques and technologies related to the spatial web and the launch of Google Maps, the advent of OpenStreetMap, and crowdsourcing, and to Web 2.0 platforms such as Facebook and Wikipedia, as well as mobile devices such as smartphones (Lauriault, 2017). The social web and mobile devices enabled people not only to be consumers of data but also to be content producers, creating a new set of data actors who were not necessarily scientists, data authorities, or producers in government but people skilled with coding, open source, APIs, and data science. The antecedents to open data are international natural- and social-science researchers, environmentalists, EO and geomatics communities, governments with spatial data infrastructures,

librarians and archivists, sociologists, and transnational organizations such as the Organisation for Economic Co-operation and Development (OECD) and the United Nations engaged in international and sustainable development, or the European Union engaged with the facilitation of the regional integration of national data assets across borders. Data-sharing and open data actors differ.

Data-sharing has dynamically evolved across time and space for centuries, along with technological capabilities. For example, one of the first compendiums of statistics and maps was the Domesday Book of 1086, and 400 or so years later, data dissemination was accelerated with the invention and adoption of the printing press, in 1455. Governors were also pressured then, as they are today, into making the records of the state available, as exemplified by the 1766 Freedom of the Press Act in Sweden. The Enlightenment and the scientific revolution, with the formation of societies, also formalized and standardized data sharing, as seen in the founding principles of the IMO, articled by Buys Ballot, the organization's first president, in 1873: "It is elementary to have a worldwide network of meteorological observations, free exchange of observations between nations and international agreement on standardized observation methods and units in order to be able to compare these observations" (WMO, n.d.; Buys-Ballot, 1872).

Statistical, social-science, and scientific associations of the late Enlightenment period had similar principles, as was the case for the International Council of Scientific Unions (1931), which merged the International Association of Academies and the International Research Council, which inform the practices of granting councils such as the Canadian Tri-Agency of the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council, and the Social Sciences and Humanities Research Council. The ethos of sharing was later codified into the CUDO-norms of science in 1942 by sociologist of science Robert K. Merton. CUDO stood for communism, universalism, disinterestedness, and organized skepticism, whereby scientists and scientific institutions were encouraged to share the results of their work for the common good, for the purpose of advancing the scientific enterprise, and to ensure that scientific claims were scrutinized before being accepted. As seen in these few examples, it is difficult to pinpoint when and how the practices of data sharing truly began; perhaps, then, as it is now, it was enabled by a social and technological assemblage of

factors at different times, places, and contexts. Not least of these during the Enlightenment was patronage, secularism, literacy, and the means for information to be published, and for it to travel (Anderson, 1986).

Here, I choose to start with the Antarctic Treaty of 1959, which includes the following principles to govern how scientists involved in Antarctic research are to act:

> 1. In order to promote international cooperation in scientific investigation in Antarctica, as provided for in Article II of the present treaty, the Contracting Parties agree that, to the greatest extent feasible and practicable:
>   (a) information regarding plans for scientific programs in Antarctica shall be exchanged to permit maximum economy and efficiency of operations;
>   (b) scientific personnel shall be exchanged in Antarctica between expeditions and stations;
>   (c) scientific observations and results from Antarctica shall be exchanged and made freely available.

I start here because the impact of the treaty is easy to trace; for example, the International Polar Year of 1957 led to the formation of the Scientific Committee on Antarctic Research (SCAR) and, later, the Antarctic Treaty. SCAR scientists were and remain affiliated with global scientific institutions such as the World Data System, the International Science Council, GEOSS, the Committee on Data of the International Science Council (CODATA), the Research Data Alliance, and many others that have advocated for, institutionalized, and operationalized the sharing of data since. Furthermore, SCAR operationalized early data-sharing policies and created one of the first global, standardized, and interoperable scientific data portals (SCAR, 2020).

Making data accessible also became a key international policy at the UN Earth Summit of 1992, which mandated nations to collect and manage their data and information assets, and to build capacity and openly share them. Chapter 40 (UN, 1992) opens with the following statement:

> 40.1. In sustainable development, everyone is a user and provider of information considered in the broad sense. That includes

> data, information, appropriately packaged experience and knowledge. The need for information arises at all levels, from that of senior decision makers at the national and international levels to the grass-roots and individual levels. The following two programme areas need to be implemented to ensure that decisions are based increasingly on sound information:
> a. Bridging the data gap;
> b. Improving information availability.

Chapter 40 also featured a broad base of data actors, and not just scientists. Indigenous Peoples and regional communities were included because they possess important local knowledge, which comes in many forms, and translating that knowledge into digital data is vital to sustainable development. New governing structures have emerged to protect these data; one example is the FNIGC's OCAP principles, discussed above, which do not sit easily with open data by default. There are similar principles by Inuit in Canada, as seen in the Inuit Tapiriit Kanatami's National Inuit Strategy on Research (NISR) (ITK, 2018). The subjectivities of the FNIGC's (n.d.) OCAP and the ITK's NISR are situated in a post-colonial discourse that asserts sovereignty over the knowledge of First Nations and Inuit in Canada. This is important since private data about Indigenous communities are often possessed by the Crown but arguably should be owned by Indigenous Peoples and communities. Public access in this case would require a nation-to-nation form of negotiation and agreement. First Nation, Métis, and Inuit data should be governed differently, even though these data are often not in their communities' possession, such as archival data recorded by explorers, since for Indigenous Peoples these are considered private, are part of their story of colonialism, and might be about sacred sites or potentially sensitive from an ecological or biodiversity perspective. First Nations would argue that any data about them ought not to fall under an open-by-default policy, and that the sharing of these ought to be negotiated.

Access to data is also about capacity building, open science, and the restructuring of government institutions involved in science and statistics broadly, but also about building open data and data-sharing infrastructures such as GEOSS or spatial data infrastructures. In Figure 1.2, I illustrate how open data emerged as a concept and include some important global milestones.

Open data as an international concept is thought to have formally appeared in 2005–2006 with the Open Definition published by the OKF, and with the "Give us back our crown jewels" and Free Our Data campaign by *Guardian* journalists Charles Arthur and Michael Cross (2006 [March 6], 2006). Prior to that, scientific and geospatial communities and transnational organizations were developing organizations, data centres, practices, and protocols to share data for the advancement of science, better management of the environment, more efficient public administration, and generally for the betterment of society. Access to data and open data are also related to the open-source movement and the General Public Licence, the Open Source Initiative and the Creative Commons, open-access publishing, and the sharing of the results of publicly funded science, open science, and interoperability, as in the case of Global Map, the Open Geospatial Consortium, and national spatial data infrastructures (SDIs) built for the purpose of data sharing. Climate change and other environmental issues led to Agenda 21, followed by Rio+10 in 2002, where the EO community advocated SDIs in the *Down to Earth* report (2002) The EU project of integrating systems also developed directives to share public sector information, and the 2007 EU INSPIRE directive for SDI. For example, one of the first foundational datasets to be opened was the public use of the Global Positioning System, in 1983, upon which location-based services are built and are a key feature in mobile systems today, being part of every smartphone, wearable device, autonomous vehicle, the Internet of Things (IoT), smart cities, and digital twins. Open access for academic publications—called for in the Budapest Open Access Initiative (2002)—was also key, as journals were mandated to be not only accessible but also to publish data upon which results were based. This is promoted in Canada by the federal Tri-Agency, noted above, which funds the bulk of the research (GoC, 2016). Others were also involved in the sharing of data, notably scientific, transnational, and civil-society organizations like the Sunlight Foundation, OECD, CODATA, and W3C (the World Wide Web Consortium) for semantic interoperability, later followed by the G8, Open Government Partnership, and the Open Data Institute (ODI). Important agreements such as the Open Data Charter and indices like the Open Data Index and the Open Data Barometer also came online.

Open data as a discourse has normalized practices, and is becoming routinized and operationalized in governments, but it is

disjointed as the focus is primarily on administrative data and, to a lesser extent, survey and science data. Although it is evolving, open data is not a systems-based approach, it is a policy, as it was in the sciences, where there was a purpose to sharing data within a community of practice or an epistemic community. This is in contrast with public-sector administrators who create data for the purpose of managing and operating government programs. New open data institutions are forming, but these are situated within a data-management and an information-technology (IT) context, informed by the management of government records and governed by what are known as C-level officers (chief technology, chief information, chief data, etc.), data-protection officers, and sometimes those who lead digital strategies and who manage data as objects in keeping with new managerialist forms of governing.

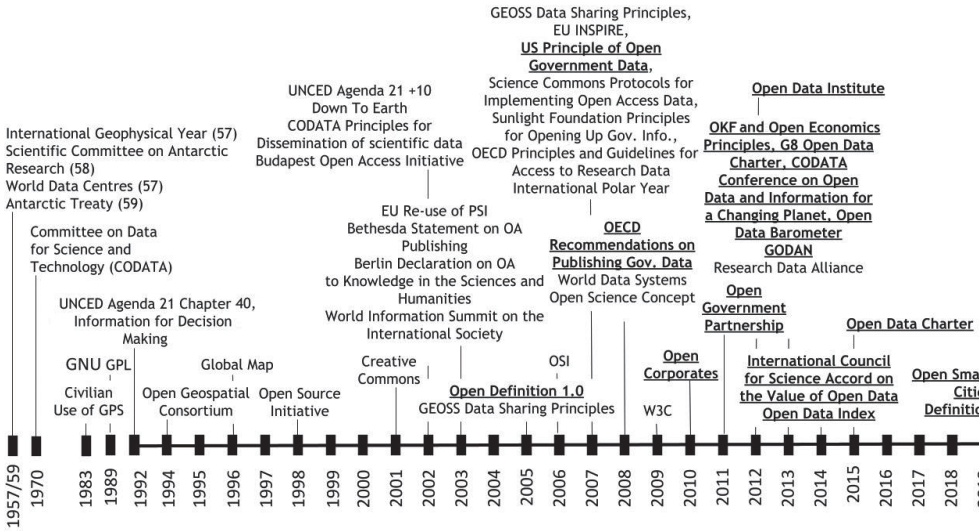

Figure 1.2. Genealogy of Global Milestones influencing Open Access and Open Data in Canada.
Source: Tracey P. Lauriault.

As seen in the genealogy in Figure 1.2, open data did not come from nowhere; it has a history and a provenance. This timeline includes only a selection of institutions, and is therefore partial as it does not include all related intellectual property initiatives, nor does

it include a list of the global actors involved in the protection of Indigenous local and traditional data; it does, however, demonstrate that there are international actors that influence government administrations and civil-society actors. It is interesting to note the epistemological shift—access to data was situated in science, for open data it is new managerialism—and the subjectivities of these different approaches led to different outcomes. It is this shift in epistemologies and actors that, I suggest, has led to open data portals and the publishing of poor-quality data, and the lack of data about complex socio-economic issues such as homelessness, disability, and equity including framework or foundational data. The focus is more on administrative and public-sector data and less on the well-developed and scientific practices of government data such as statistics, mapping, and research data derived from publicly funded science. It is also the reason why data are published as discrete objects as opposed to being parts of sets of authoritative data, records sets, or systems such as open-science monitoring, and why there is a lack of standardization in terms of name space, tags, and spatial referents. These are important if different datasets are to be linked or joined for national scale and analytical purposes. Furthermore, data policies governing administrative data are neither as robust nor as integrative as the practices of the physical and social sciences, geospatial data infrastructures, statistical agencies, and research data, or of open science. This has implications for smart systems and digital twins: Will they be standardized, and will data quality be considered; will they be interoperable; and more importantly, how will these be governed? This will require systems thinking about data, AI/ML, sensors, and related infrastructures as social and technological assemblages.

The open data community also did not foresee the emergence of smart cities, digital twins, or big data analytics since this epistemic community was generally more attuned to e-government, digital strategies, and administrative data, not to the governance of cities, sectors like agriculture, or analytics beyond application development and application programming interfaces (APIs). This is changing, but integration is slow, while it is uncertain if the values of openness will be mapped onto the smart city or digital twin. The concern here is that smart systems may follow the same ahistorical and disconnected trajectory that open data has. For example, if we look to the collection of real-time data from sensors, data are

inseparable from the systems that produce them, and these are often closed and proprietary; or, as Scassa and Diebel (2016) have demonstrated, they may start as open, but when value is discovered, they become closed. Also, if we look to the establishment of IoT and sensor-based approaches from science, as seen in the vast machines (Edwards, 2013) of seismology, EO, and meteorology, these later IoT systems most often practise open science and do not necessarily align with the intentions of corporate platform-based companies or city officials, where the priority is to manage city operations efficiently, innovatively, and economically. But there is no reason why there might not be a mutually beneficial arrangement between cities, the private sector, and civil society actors. It is out of this situated knowledge that the Open Smart Cities Definitions emerged (Lauriault et al., 2019), to develop a way to bridge sectors, and to build on good practices so that these systems are developed and governed with the public good in mind, and this is why looking to the discursive past of open data matters when it comes to managing the future of smart systems, especially since these will often be overseen by the same IT subcommittees that developed open data programs in a city, by the consultancies that advise them, and by platform companies.

4. Open Data in Canada: A Genealogy

As discussed, open data did not suddenly appear internationally as a fully formed concept, nor did it evolve in a consistent manner at the nation-state level. Open data has an international, national, and a local history. In this section, I provide an open data history for Canada as seen in the context of genealogy (Figure 1.3) and the materialities genealogy (Figure 1.4). The items that are in bold and underlined signify open data actors, while those not underlined are access to data actors. Table 1.1 provides a selection of elements related to the context components of a Canadian social and technological assemblage.

4.1 Open Data in Canada

This data-sharing origin story starts with the Canada Institute for Scientific and Technical Information (CISTI), created in 1974 at the National Research Council of Canada. CISTI was created to ensure

that scientists had access to the data and information they need to do their work. These were disseminated with the CAN/OLE online catalogue system and the work of legal scholar and former politician Murray Rankin (1978), who argued that researchers should have access to government information. Librarians were also innovators as they developed data libraries with access, standards, policies, and user guides, as well as technological and user services in the days when data were only accessible on magnetic tape (Ruus, 1982). Furthermore, in 1986, the Canadian Association of Research Libraries created a research data consortium, and, in 1988, other researchers and librarians formed the Canadian Association of Public Data Users. Discourse on access to data also featured in government reports, most notably the mid-1980s Ministerial Task Force on Program Review, commonly known as the Nielsen Task Force. Led by Deputy Prime Minister Erik Nielsen (1984), it conducted an extensive inventory of data assets and concluded that these data should be made publicly available (1984). The Progressive Conservative government of Brian Mulroney, however, did not take this advice and instead created a cost-recovery regime for data, making government data cost prohibitive, especially statistical data. Some might say that this action spearheaded the movement to make data open and accessible in Canada (Humphrey, 1994). This also influenced the creation of data-purchasing consortia, whereby organizations pool their economic, technological, skill, and institutional resources to purchase and share data under a consortium licence (StatCan, 2019). In 1992, sociologist Paul Bernard (1992, p. 21) from the Université de Montréal argued that "knowledge is fundamental to economic development and democratic life in advanced societies; and the information gathered by statistical agencies is an important component of that knowledge. It is essential that such information be made available to researchers and to the public so that it can be used in debates and decision-making."

Sociologists in 1992 were also responsible for *Liberating the Data: A Proposal for a Joint Venture between Statistics Canada and Canadian Universities* (Watkins, 1992), which led to the creation of the federal Data Liberation Initiative (DLI) in 1994, and, for the first time, Statistics Canada disseminated data on the Internet via FTP (file transfer protocol). The DLI made data open to faculty and students but not to the public, as the Statistics Canada licence was restrictive. Outside the academy, statistical data were inaccessible as they were cost

prohibitive. As a result, community-based social-planning councils in the mid-1990s also developed data-purchasing consortia, such as the Geographic and Numeric Information Systems (GANIS) and the Canadian Council on Social Development's (CCSD) Community Data Program. These groups coalesced hundreds of community-based organizations in urban and rural areas across Canada to co-purchase customized cross-tabulated data about socio-economic issues that were aggregated at local geographies, under a special consortium licence from Statistics Canada. In this case, hundreds of NGOs collaborated to gain access to data to study Canada's most marginalized communities.

There were other epistemic groups, such as the geospatial community. The *Atlas of Canada*, first published in 1905, started publishing maps online in the 1980s, and it launched the world's first Internet, open source, and web atlas in 1999. The geospatial community also launched the first open data portals with GeoGratis (1993) and GeoBase (1994), and formed the Canadian Geospatial Data Infrastructure in 1999. This was one of the world's first open-source, open-access, open-architecture, open-specifications, and standards-based data infrastructures. The geospatial community also spearheaded the first discussions to openly licence data under Crown copyright (2008). As seen in Figures 1.3 and 1.4, librarians, archivists, sociologists, and researchers were advocating for the release of social science and research data, while the geospatial community were disseminating their data in open spatial data infrastructures. The former group were working against cost recovery and outside the administration, while the latter were data producers within the administration who were developing systems very much in step with addressing specific issues, such as the environment and resource management. This required multisectoral and multi-jurisdictional collaboration, a workaround to cost recovery, and operationalizing technical, policy, and institutional interoperability.

As was the case internationally, open data in Canada emerged as a concept in 2005, with the How'd They Vote application enabling residents and citizens to see how elected officials voted, and to track what they said in the media with the Civicaccess.ca list and the DataLibre.ca blog (Lauriault & McGuire, 2010). The latter two were created by a group of individuals from community Wi-Fi, access-to-data advocacy, librarians, computer scientists, and many others.

They did not come together because of the Open Definition but instead out of a concern that government data such as statistics and elections data were not available. This was also a time when Web 2.0 tools such as Google Maps and mobile devices were coming online. These enabled mashups, and crowdsourcing projects like OpenStreetMap (2004). A new cohort of data users outside of research communities and government administrations, along with autonomous data producers not affiliated with organizations, emerged from social media, Web 2.0, and the proliferation of mobile devices enabled by location-based services (Kitchin et al., 2017). Open data in Canada was also the culmination of ideals, experience, research, practice, and the work of a number of actors, building on the pre-existing initiatives discussed above. It was also the result of a chance encounter between three people involved in community Wi-Fi, web accessibility, and access to data at the 2005 UNESCO World Summit on the Information Society (WSIS) II civil-society conference in Winnipeg. It was there that Lauriault, Lenczner, and Roy met on a panel to discuss open data, accessibility, and community technology initiatives. They were also invited to draft the "Canadian Civil Society Communiqué" that went to the Tunis World Information Society Summit, which included the following in the preamble: "We firmly maintain that democracy is reliant on an informed citizenry and civil society that has access to the data, information, knowledge and technology necessary to keep governments accountable" (UNESCO, 2005).

It was shortly thereafter that CivicAccess, DataLibre, and the G4+1[3] were formed, along with similar groups in Vancouver. The Open Data Summit and the BC Institute for Open Data were developed; along with actors in Toronto, the GO Open Data Conference, the Open Data Institute chapter, Open North, Ajah.ca, and Powered by Data were created. Individuals from these organizations are now part of the Multistakeholder Advisory Group on Open Government. Also, some of these open data actors were also involved in the creation of the Open Smart City Definition (Lauriault

[3] The G4+1 is an informal group of cities—Vancouver, Edmonton, Toronto, Ottawa—that had fledgling open-data programs, and Montréal is the +1 as it launched its open data program later. The group was founded by Lauriault in 2009–2010, at a GTEC Conference in Ottawa, to enable cities to work on common open-data issues. It continues to meet to discuss and resolve common issues.

et al., 2019), and some later went on to form open data civil-society groups, businesses, and new scholarly domains such as critical data studies.

In terms of operationalizing open data in Canada, cities were the early adopters and innovators, starting with the first open data portal coming online in Nanaimo, British Columbia, in 2009 (which included primarily geospatial data), and the creation of the G4+1 group that still meets monthly. This group pooled their resources to work on the first open licence with the Canadian Internet Public Policy Interest Clinic (2016), the standardization of open data metadata, the sharing of best practices, and promoting open data in cities. The federal government first mentions open data in the Standing Committee on Industry, Science and Technology (Lauriault, 2008), followed, in 2010, by the resolution of Canada's access-to-information and privacy commissioners under the leadership of Commissioner Suzanne Legault, in which open data, open government, and freedom of information are linked (see OPC 2010). In 2011, the Treasury Board Secretariat of Canada launched data.gc.ca, a comprehensive knowledge archive network (known as CKAN), data portal, and Canadian government officials attended the first Open Government Partnership meeting, in 2012 in Brazil, accompanied by civil-society actors from the Community Data Program, the Centre for Law and Democracy, and David Eaves, an open data advocate. And the rest, we might say, is history. Since that time, several provinces have launched open data programs (Hunter & Lauriault, 2020), and close to 90 cities and communities now have open data in one form or another.[4] Go Open Data in Ontario was created in 2014. The Open Data Charter has been adopted and is stewarded by Open North, there is the Multi-Stakeholder Forum on Open Government (2018), and Canada hosted the Open Government Partnership Global Summit 2019, in Ottawa, and the Canadian Open Data Society, incorporated in 2020. Dozens of other important civil-society organizations, such as Transparency International, have formed while open data is normalizing, including as seen in open data directives, road maps and proactive disclosure, open contracting, and beneficial ownership, to name a few.

[4] See the following sources for lists of open-data initiatives in Canada, none of which are complete or up to date: https://open.canada.ca/en/maps/open-data-canada, https://en.wikipedia.org/wiki/Open_data_in_Canada, and http://datalibre.ca/links-resources/.

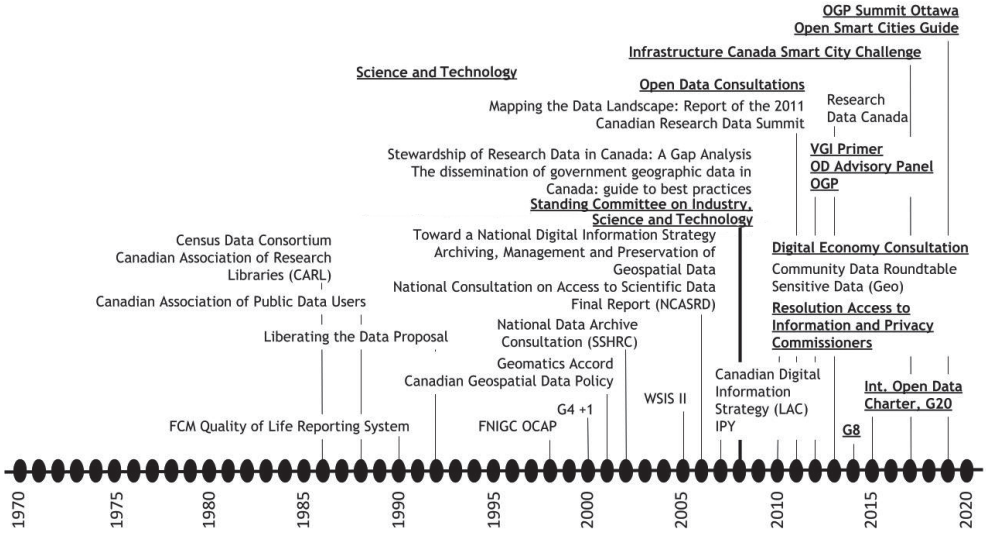

Figure 1.3 Genealogy of Open Access and Open Data Policies and Documents in Canada.

Source: Tracey P. Lauriault.

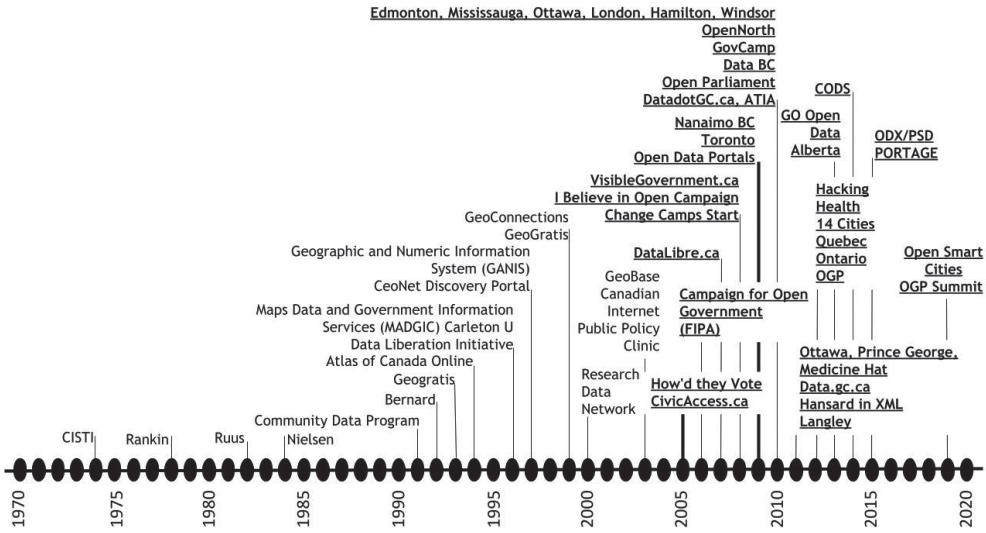

Figure 1.4 Genealogy of Open Access and Open Data Advocates and Initiatives in Canada.

Source: Tracey P. Lauriault.

5. Open Data in Smart Systems

Open data as discourse and practice dynamically evolves; it is not a stand-alone way of doing things, and it is now part of open government programming nationally, and in some provinces and many cities, as technologies and processes evolve, as new policy priorities emerge, and as public engagement and discussion about automated decision-making grow (GoC, 2021). As digital strategies and charters (GoC, 2019) take form, and as smart cities become a priority, open data have become less topical. In Canadian cities, open data programs are being subsumed as part of smart city initiatives; they are becoming a component of the smart city, while smart city data are not necessarily open data and technological systems are not being developed according to the practices of open science or spatial data infrastructures. The pandemic has also clearly demonstrated that open data has not become the norm in certain domains, most notably in public-health reporting (Hunter & Lauriault, 2020).

In 2017, a department of the federal government, Infrastructure Canada, launched a Smart Cities Challenge, where 225 large and small communities, including Aboriginal communities, submitted proposals: 130 were deemed eligible, 20 were short-listed, and four winners were announced on May 14, 2019 (Infrastructure Canada, 2019). This Challenge was interesting as the call made it clear that meaningful smart city outcomes included making data, decision-making, and technological processes open, transparent, and interoperable. The call also mandated that chosen technologies be transferable, and preferably open source and standards-based, for reuse by other communities; that communities have ownership over their data; and that technologies empower and enable communities large and small, as well as traditional and non-traditional partners, to collaborate and strengthen relationships between residents and public organizations, including gender-based analysis (known as GBA+) (Impact Canada, 2018). The procurement of technology and the ownership of data became part of the strategy, and cities had to define challenges to address with technology and also had to consult with their residents to do so. Prior to this, cities in Canada were developing smart city strategies and plans that looked more like a form of networked urbanism, whereby

> big data systems [prefigure and set] the urban agenda and are influencing and controlling how city systems respond and perform . . . cities are becoming ever more instrumented and networked, their systems interlinked and integrated, and the vast troves of data being generated used to manage and control urban life. Computation is now routinely being embedded into the fabric and infrastructure of cities producing a deluge of contextual and actionable data which can be processed and acted upon in real-time. (Kitchin, 2017, p. 43)

Yet, when Sidewalk Labs, a subsidiary of Google's parent company, Alphabet, announced its plans for the Toronto Quayside project at around the same time as Infrastructure Canada launched the Smart Cities Challenge, there was no data-governance strategy, and data were not going to be open. In May 2020, Sidewalk Labs (Carter & Rieti 2020) withdrew its participation from the Quayside project. Prior to these two big initiatives, cities across Canada had begun developing their own smart city programs, and as they were doing so, it became clear that these were innovation- and efficiency-driven, were being organized in the information-management and information-technology (IM/IT) division, and were part of data analytics plans. Citizens in these early smart city programs were not engaged, and open data programs were being subsumed as part of smart city units, while the data derived from smart city technologies were not necessarily going to be open (Lauriault et al., 2019). This is also in part related to the activities of corporate consultancies, technology alliances, and platform companies which have the ear of cities and advocate for closed proprietary systems (Lauriault et al., 2019). Many small, big, and real-time data are being generated by smart cities, and these may include administrative data from intake systems and big data and real-time data generated by sensors and cameras, some of which are private data or behavioural data, which are private and personal in nature. These smart city data bring forward issues related to algorithmic decision-making, and "open data" are no longer on the mainstage, and those data that can be open may be "locked up" into procurement agreements that impede sharing. In the case of smart cities, a new data-enclosure movement might be afoot, and perhaps we need to look at open science as a possible framework to ensure that data and technologies are open, procured, and deployed in the public interest. Open North's Open Smart City

Guide has advocated for this, and has since evolved into a city self-assessment instrument, education modules, several policy briefs, procurement plans, research, and a community solutions network, with several cities adopting these practices.[5] Will there be an open digital twin?

6. Conclusion

Canada has a unique open data social and technological assemblage (Figure 1.1, Table 1.1), and genealogy (Figures 1.3 and 1.4). And technical conceptualizations of open data persist, which limit the scope of open data. This chapter has framed open data as an assemblage to demonstrate how open data is a discursive regime that includes many content and context components that are both social and technical that, across time, space, and actors, constitute Canada's open data story. Also, Canada's open data story is situated in a global context, as seen in Figure 1.2. Open data, in reality, includes many interconnected yet disparate smaller assemblages found in many cities, provinces, and territories, and in the federal government and the international arena, and most often localized in IM/IT divisions. In Canada, open data as a discourse emerged from the efforts of separate and mostly distinct scientific, research, and geospatial communities that included granting councils, sociologists, librarians, archivists, and scientists. For example, the geospatial community created spatial data infrastructures; research-intensive universities developed their own social-science infrastructure, now coming together under the PORTAGE, a data preservation system by CARL and trusted digital repositories of data, the New Digital Research Infrastructure Organization (called the Alliance); the National Research Council focused on science; and community-based organizations created their own data-sharing consortia and portals (Figures 1.3 and 1.4). These access-to-data communities seeded the terrain. Open data in Canada was initiated by cities, first in GIS units and later in IM/IT departments, which are often disconnected from planning or social services, and it was new managerialist in tone, with data portals and indicators rendered in dashboards and maps. This

[5] See, on the Open North website, https://opennorth.ca/publicationdetail?id=3Ptq7I6gVIfzBfl2ZAYoNs https://opennorth.ca/publications, and on the Future City Network website, https://futurecitiescanada.ca/programs/community-solutions-network/.

became evident with pandemic reporting by federal, provincial, and territorial governments, which made their data accessible in dashboards but not necessarily as open data (Hunter & Lauriault, 2020). Open data as operations and as a discourse has since become common at all levels of government. Open data, however, evolved separately from the early efforts of those involved in access to data; they constituted different actors and communities of practice in different sectors. Open data today still lacks much of the systems thinking of the access-to-data communities: there are fewer standards; the quality of the data and metadata is inconsistent; there is a lack of interoperability; there are few integrative framework datasets that stitch the country's assets together, although data linkage projects are emerging; and data discovery associated with metadata and tagging is poor. The corollary is that we have thousands of open datasets, and perhaps open-by-default practices might give way to publishing data with purpose, as was the case with the access-to-data communities, and this might improve data quality. And as discussed, there is also a tension between open data by default and Indigenous data that has yet to be resolved. And as seen during the ongoing pandemic, there are important datasets that just do not exist for some of Canada's most vulnerable communities, and there has been a lack of ethical and intersectional frameworks of equity and inclusion to inform the production of these data and the creation of important framework data (Linton & Lauriault, 2021; Hunter & Lauriault, 2020; Lauriault, 2020b).

Canada is now home to several fledgling smart cities, and although these may have subsumed open data programs, smart city technologies and the data they produce are by no means open. The winners of the Smart Cities Challenge were announced, their data will be open as per the requirement of the call, there will be public engagement about technological decisions, and here we may witness big and real-time open data coming from the short-listed and finalists (INFC, 2017). But what of all the other smart cities and digital twins: Will they be open? Open-smart city aspirations are becoming a reality; the definition counters the enclosure of data and technology, and it is becoming a made-in-Canada model. Of concern, however, is the lack of systems and infrastructural thinking overall in data and technology spaces; be they large social and technological systems such as smart cities, digital twins, or smart grids for utilities, we do not yet see data-governance plans that are systems-based, integrated,

interoperable, standardized, and open. Will open data and open smart cities be included in digital strategies and become a core principle of data-governance plans?

It is also the hope that good and established practices from open science, open data, open source, open platforms, open government, and emerging engagement processes will be mapped onto them, but most importantly that there will be open smart cities, "where residents, civil society, academics, and the private sector collaborate with public officials to mobilize data and technologies when warranted in an ethical, accountable and transparent way to govern the city as a fair, viable and livable commons and balance economic development, social progress and environmental responsibility" (Lauriault et al., 2019).

It is important to learn from the history of open data to better understand its future and to steer a better course for smart systems. Here it is argued that open data will need to be governed as part of large social and technological systems; that its discourse ought to be about democratic deliberation and not just a new managerialist exercise; and that we move beyond narrow technological conceptions and include ethics, public good, and sustainability. Current smart city actors, in addition to public officials, data scientists, and software engineers, may want to look at established and experienced epistemic groups, such as those in the spatial-sciences and remote-sensing communities, and other scientific communities such as natural resources, meteorology, and oceanography or glaciology, where sensor-based open science has been practised for decades. There are many good practices to emulate, most especially when it comes to spatial data infrastructures. Open data in Canada evolved in the absence of the early actors who practised data sharing—scientists, early adopters in the spatial sciences, sociologists—and it is hoped here that open smart cities can steer the course away from more corporatist and closed smart cities and toward open smart systems. The future of open data, as it normalizes, will be subsumed as part of large and smart-technological systems, and we need to ensure that data and the systems that generate them remain open. Open data actors will need to work with open smart city and digital-twin actors to ensure that data remain open, and city actors need to think of smart cities and digital twins as more than operations, and to consider them to be part of urban plans and as part of public space that ought to be deliberated. Finally, for all, good governance includes, among other things, data and technological governance,

and an inherent part of doing technological citizenship (Barney, 2004), since we do live in technological societies (Feenberg, 2011) after all.

Elements	Canada
Governments	Federated constitutional monarchy of provinces and territories with clear divisions of power between the Federal Government, Provincial / Territorial Governments and Cities. Population: 37+ Million
System of Thought	Open Government, Transparency, Accountability, Efficiency, Innovation, E-Government
Form of Knowledge	Open Government Charter Standards, WSF/WMS Metadata, ISO19115, Dublin Core Directives Policies Reports Political Platforms Indicators / Metrics Digital Strategies / Charters Enterprise Architecture Guides, Readiness Guides, Roadmaps, Theory of Change
Governmentalities and Legalities	Open Licences The Personal Information Protection and Electronic Documents Act (PIPEDA) Privacy Act P/T Privacy Legislation Municipal Freedom of Information and Protection of Privacy Freedom of Information Legislation Access to Information Legislation International Aid Transparency Initiative Proactive Disclosure Open Corporates AI Impact Assessment Data Trusts
Organizations and Institutions	*Government Open Data Programs* • Treasury Board Secretariat • Natural Resources Canada • Information and Privacy Commissioners • Provincial Governments • Municipalities and Regional governments • Most federal departments • Cities *Civil Society* • Multistakeholder Advisory Group • OpenNorth, Powered by Data, Data for Good, Open Data Institute • Centre for Law and Democracy • Open Corporates • Transparency International

Table 1.1. A Selection of the Attributes that form the Social and Technological Open Data Assemblage for Canada.
Source: Tracey P. Lauriault.

References

Anderson, B. (1983). Census, map, museum. In *Imagined communities: Reflections on the origin and spread of nationalism* (rev. ed., pp. 163–186). Verso.

Arctic Spatial Data Infrastructure (ASDI). (2021). *Home page*. Retrieved June 2, 2021, from https://arctic-sdi.org/

Arthur, C., & Cross, M. (2006, March 6). Give us back our crown jewels. *Guardian* [UK]. Retrieved July 9, 2019, from https://www.theguardian.com/technology/2006/mar/09/education.epublic

Arthur, C., & Cross, M. (2006). Free our data campaign. *Guardian* [UK]. Retrieved Aug. 18, 2021, from https://www.theguardian.com/technology/2006/mar/09/education.epublic

Barney, D. (2004). *The network society*. Polity Press.

Barney, D. (2007). *One nation under Google: Citizenship in the technological republic*. Hart House Lecture Committee. Retrieved May 10, 2019, from http://darinbarneyresearch.mcgill.ca/Work/One_Nation_Under_Google.pdf

Bernard, P. (1991). *Discussion paper on the issue of the pricing of Statistics Canada products*.

Bernard, P. (1992). Data and knowledge: Statistics Canada and the research community. *Society*, 21.

Berners-Lee, T. (2006). *5 star linked data*. Retrieved June 9, 2019, from https://www.w3.org/2011/gld/wiki/5_Star_Linked_Data

Budapest Open Access Initiative. (2002). Retrieved June 9, 2019, from https://www.budapestopenaccessinitiative.org/read

Buys-Ballot, C. H. D. (1872). *A sequel to the "Suggestions on a uniform system of meteorological observations*." Royal Dutch Meteorological Institute.

Canadian Association of Public Data Users. (n.d.). *About*. Retrieved May 11, 2019, from https://capdu.wordpress.com/about/

Canadian Association of Research Libraries. (n.d.). *About* CARL. Retrieved June 9, 2019, from http://www.carl-abrc.ca/about-carl/

Canadian Council on Geomatics (CCOG). (n.d.). *Canadian Geomatics Accord*. Retrieved May 11, 2019, from http://www.ccog-cocg.ca/en/accord

Canadian Internet Policy and Public Policy Interest Clinic. (2016). *CIPPIC proposal: An open licensing scheme for traditional knowledge*. Retrieved April 27, 2021, from https://cippic.ca/en/TK_Open_Licensing_Proposal

Carter, A., & Rieti, J. (2020, May 7). *Sidewalk Labs cancels plan to build high-tech neighbourhood in Toronto amid COVID-19*. CBC News. Retrieved May 7, from https://www.cbc.ca/news/canada/toronto/sidewalk-labs-cancels-project-1.5559370

City of Ottawa. (2021). Information shared in person as part of a presentation given as community partners with the Master Class in Critical Data Studies of 2021 at Carleton University, Ottawa.

Cosgrove, D. E. (2001). *Apollo's eye: A cartographic genealogy of the earth in the western imagination*. Johns Hopkins University Press.

Couldry, N., & Mejias, U. A. (2018). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television and New Media*, *20*(4), 336–349. Retrieved May 10, 2019, from https://journals-sagepub-com.proxy.library.carleton.ca/doi/10.1177/1527476418796632

Dalton, C., & Thatcher, J. (2014). What does a critical data studies look like, and why do we care? Seven points for a critical approach to "big data". *Society and Space*. Retrieved May 10, 2019, from https://societyandspace.org/2014/05/12/what-does-a-critical-data-studies-look-like-and-why-do-we-care-craig-dalton-and-jim-thatcher/

DeLanda, M. (2016). *Assemblage theory*. Edinburgh University Press.

Edwards, P. (2013). *A vast machine: Computer models, climate data, and the politics of global warming*. MIT Press.

European Union. (2021). *EU general data protection regulation*. GDPR.eu. Retrieved June 2, 2021, from https://gdpr.eu/

Feenberg, A. (2011). *Agency and citizenship in a technological society*. Lecture presented to the Course on Digital Citizenship, IT University of Copenhagen. Retrieved May 10, 2019, from https://www.sfu.ca/~andrewf/copen5-1.pdf

First Nations Information Governance Centre (FNIGC). (1998). *The First Nations principles of OCAP*. Retrieved May 10, 2019, from https://fnigc.ca/ocapr.html

First Nations Information Governance Centre (FNIGC). (2019). First Nations data sovereignty in Canada. *Statistical Journal of the IAOS*, *35*(1), 47–69. https://doi.org/10.3233/SJI-180478.

First Nations Information Governance Centre (FNIGC). (2021). *Home page*. Retrieved May 24, 2021, from https://fnigc.ca/about-fnigc/

First Nations Information Governance Council (FNIGC). (n.d.). *OCAP principles*. Retrieved June 2, 2021, from https://fnigc.ca/ocap-training/

Foucault, M. (2003). *The essential Foucault: Selections from essential works of Foucault, 1954-1984*. The New Press.

Foster. (n.d.) *What is open science?* Retrieved June 2, 2021, from https://www.fosteropenscience.eu/node/2326

G8 (2013) *G8 Open Data Charter: UK Action Plan 2013*. Retrieved August 15, 2021, from https://www.gov.uk/government/publications/g8-open-data-charter-national-action-plan/g8-open-data-charter-uk-action-plan-2013

G20. (2015). *G20 Anti-Corruption Open Data Principles*. Retrieved May 11, 2019, from http://www.g20.utoronto.ca/2015/G20-Anti-Corruption-Open-Data-Principles.pdf

GeoBase. (2001). *About*. Retrieved July 9, 2019, from https://web.archive.org/web/20151125163309/http://ftp2.cits.rncan.gc.ca/pub/geobase/pdf/About.pdf

GeoConnections. (2008). *The dissemination of government geographic data in Canada: Guide to best practices*. GEOSCAN. Canadian Geospatial Data Infrastructure, Information Product 8, (ed. ver. 2). https://doi.org/10.4095/288853

GeoConnections. (2012). *Canadian geospatial data infrastructure vision, mission and roadmap—the way forward.* GEOSCAN. Canadian Geospatial Data Infrastructure, Information Product 28e. https://doi.org/10.4095/292417

GeoGratis. (n.d.). *Portal*. Retrieved July 9, 2019, from https://www.nrcan.gc.ca/science-and-data/science-and-research/earth-sciences/geography/topographic-information/10785

Global Earth Observation System of Systems (GEOSS). (n.d.). *Principles.* Retrieved June 9, 2019, from https://www.earthobservations.org/geo_community.php

Global Earth Observation Systems of Systems (GEOSS). (n.d.). *About.* Retrieved July 9, 2019, from https://www.earthobservations.org/geoss.php

Global Open Data for Agriculture and Nutrition (GODAN). (2016). *Agricultural open data package working group*. Retrieved May 10, 2019, from https://www.godan.info/working-groups/agriculture-open-data-package-working-group

Global Open Data Index. (2017). *About*. Retrieved July 9, 2019, from https://index.okfn.org/about/

Government of Canada (GoC). (2016). *Tri-agency open access policy on publications*. Retrieved June 2, 2021, from https://www.ic.gc.ca/eic/site/063.nsf/eng/h_F6765465.html

Government of Canada (GoC). (2019). *Canada's digital charter in action: A plan by Canadians, for Canadians.* Retrieved June 2, 2021, from https://www.ic.gc.ca/eic/site/062.nsf/eng/h_00109.html

Government of Canada (GoC). (2021). *Responsible use of artificial intelligence (AI).* Retrieved June 2, 2021, from https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/responsible-use-ai.html

Government of Sweden. (1766). *Freedom of the Press Act*. Retrieved May 24, 2021, from https://sweden.se/society/20-milestones-of-swedish-press-freedom/#

Guptill, S., & Morrison, J. L. (1995). *Elements of spatial data quality.* International Cartographic Association Commission on Spatial Data Quality. Elsevier.

Hughes, T. P. (1987). The evolution of large technological systems. In W. Bijker, T. P. Hughes, & T. Pinch (Eds.), *The social construction of large*

technological systems: New directions in the sociology and history of technology (pp. 45–77). MIT Press.

Humphrey, C. (1994). The case for a Canadian national social sciences data archive. *Government Information in Canada*. 1, No. 2.7. Retrieved June 26 from https://library2.usask.ca/gic/v1n2/humphrey/humphrey.html

Hunter, A., & Lauriault, T. P. (2020). *Tracing COVID-19 data: Open science and open data standards in Canada*. Retrieved May 24, 2020 from http://datalibre.ca/2020/10/13/tracing-covid-19-data-open-science-and-open-data-standards-in-canada/

Iliadis, A., & Russo, F. (Eds.), (2016). Critical data studies. Special Issue in *Big Data & Society*. Retrieved May 10, 2019, from https://journals-sagepub-com.proxy.library.carleton.ca/page/bds/collections/critical-data-studies

Impact Canada. (2018). *Impact Challenge finalist guide*. Retrieved February 14, 2020, from https://impact.canada.ca/en/challenges/smart-cities/finalist-guide

Infrastructure Canada. (2019). *Smart City Challenge*. Retrieved July 9, 2019, from https://www.infrastructure.gc.ca/cities-villes/index-eng.html

Inuit Tapiriit Kanatami (ITK). (2018). *ITK national strategy on research*. Retrieved June 2, 2021, from https://www.itk.ca/national-strategy-on-research/

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

Kitchin, R. (2017). Data-driven urbanism. In R. Kitchin, T. P. Lauriault, and G. McArdle (Eds.), *Data and the city* (pp. 44–56). Taylor & Francis.

Kitchin, R., & Lauriault, T. P. (2014). *Towards critical data studies: Charting and unpacking data assemblages and their work*. Retrieved May 10, 2019, from https://ssrn.com/abstract=2474112

Kitchin, R., Lauriault, T. P. & Wilson M. (Eds.). (2017), *Understanding Spatial Media*, Sage.

Kukutai, T., & Taylor, J. (Eds.), (2016). *Indigenous data sovereignty*. ANU Press.

Lauriault, T. P. (2008). *Open data*. INDU Standing Committee Submission, Government of Canada, House of Commons Standing Committee on Industry, Science and Technology Committees Directorate, Study on Canadian Science and Technology.

Lauriault, T. P. (2017). Open spatial media. In R. Kitchin, T. P. Lauriault, & M. Wilson (Eds.), *Understanding Spatial Media* (pp. 95–110). Sage.

Lauriault, T. P. (2020a). *Where is the national health region map of COVID-19 Cases?* Retrieved May 24, 2021, from http://datalibre.ca/2020/04/07/where-is-the-national-map-of-covid-19-data-by-public-health-units-in-canada/

Lauriault, T. P. (2020b). *COVID-19 data: Data and technological citizenship during the COVID-19 pandemic project description*. Retrieved June 2, 2021, from http://datalibre.ca/tracing-covid-19-data/.

Lauriault, T. P., Bloom, R., & Landry, J.-N. (2019). *The open smart cities guide*. Retrieved May 10, 2019, from https://opennorth.ca/publications/3ptq7i6gvifzbfl2zayons_en

Lauriault, T. P., & McGuire, H. (2010). Data access in Canada: civicaccess.ca. In B. Fitzgerald (Ed.), *Access to public sector information: Law, technology and policy: Vol. 1* (pp. 278–284). Sydney University Press.

Linton, M., Chokly, K., & Lauriault, T. P. (2021). *Invisible people and institutions: No data about custodial institutions for disabled people in Canada*? Retrieved May 24, 2021, from http://datalibre.ca/2021/02/18/invisible-people-and-institutions-no-data-about-custodial-institutions-for-disabled-people-in-canada/

National Research Council. (2002). *Down to earth: Geographic information for sustainable development in Africa*. The National Academies Press. https://doi.org/10.17226/10455

Natural Resources Canada (NRCan). (2020). *Canada geospatial data infrastructure (CGDI)*. Retrieved June 2, 2021, from https://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/10783.

Nielsen, E. (1984). *New management initiatives: Initial results from the ministerial task force on program review*. Government of Canada, Department of Finance.

Officer of the Privacy Commissioner of Canada (OPC). (2010). *Resolution of Canada's Access to Information and Privacy Commissioners*. Retrieved July 9, 2019, from https://www.priv.gc.ca/en/about-the-opc/what-we-do/provincial-and-territorial-collaboration/joint-resolutions-with-provinces-and-territories/res_100901/

Office of the Privacy Commissioner of Canada (OPC). (2021). The *Personal Information Protection and Electronic Documents Act* (PIPEDA). Retrieved June 2, 2021, from https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/

Open Data Barometer. (2017) *Open data barometer–Leaders edition ODB methodology - v1.0 | 15 September 2017*. Retrieved May 11, 2019, from http://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Methodology.pdf

Open Data Barometer. (2019). Retrieved May 11, 2019, from https://opendatabarometer.org

Open Data Charter. (2015). *Principles*. Retrieved May 10, 2019, from https://opendatacharter.net/

Open Data Charter. (2018). *Publish with purpose*. Retrieved May 10, 2019, from https://drive.google.com/file/d/1hYmoTZTDgFe9E8CtxAW6qfbjfj5W93DL/view

Open Knowledge. (2015). *The open definition V 2.1*. Retrieved May 10, 2019, from https://opendefinition.org/od/2.1/en/

Open Knowledge Foundation (OKF). (2005). *History of the open definition*. Retrieved May 11, 2019, from https://opendefinition.org/history/

Open Knowledge Foundation (OKF). (2016/2017). *Global open index methodology*. Retrieved May 10, 2019, from https://index.okfn.org/methodology/

Ordnance Survey Ireland (OSI). (2017). *National mapping strategy*. Retrieved June 2, 2021, from https://www.osi ie/services/national-mapping-agreement/

Rankin, M. (1978). *Access to information vital to researchers*. C.A.U.T. Bulletin.

Ruus, L. (1982). The University of British Columbia data library: An overview. *Library Trends*, *30*(3), 397–406.

Scassa, T., & Diebel, A. (2016). Open or closed? Open licensing of real-time public sector transit data. *JeDEM–eJournal of eDemocracy and Open Government*, *8*(2). https://doi.org/10.29379/jedem.v8i2.414.

Scientific Committee on Antarctic Research (SCAR). (2020). *Welcome to the Scientific Committee on Antarctic Research*. Retrieved June 2, 2021, from https://www.scar.org/data-products/data/

Secretariat for the Antarctic Treaty. (1959). *The Antarctic Treaty*. Retrieved May 12, 2019, from https://www.ats.aq/e/ats.htm

Statistics Canada (StatCan). (2019). *Data liberation initiative consortium licence*. Retrieved June 2, 2021, from https://www.statcan.gc.ca/eng/dli/caselaw/license

Thatcher, J., O'Sullivan, D., & Mahmoudi, D. (2016). Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environment and Planning D: Society and Space*, *34*(6), 990–1006. Retrieved May 10, 2019, from https://journals-sagepub-com.proxy.library.carleton.ca/doi/10.1177/0263775816633195

Tuhiwai-Smith, L. (1999). *Decolonizing methodologies: Research and Indigenous Peoples*. University of Otago Press.

UNESCO. (2005). *UNESCO, Paving the road to Tunis–WSIS II: The views of Canada's civil society on the Geneva plan of action and the prospects for phase II*. Retrieved July 9, 2019, from: https://www.mcgill.ca/channels/news/canadian-civil-society-communiqu%C3%A9-15735.

United Nations (UN). (1992). *Agenda 21, chapter 40, information for decision-making*. Retrieved May 12, 2019, from https://www.itu.int/net/wsis/docs2/pc3/contributions/C013.pdfW3C

VanBuskirk, M. (2008). *The history of the NRC Canada Institute for Scientific and Technical Information 1924-2009*. National Research Council.

Walter, M., & Andersen, C. (2013). *Indigenous statistics: A quantitative research methodology*. Routledge.

Watkins, W. (1992). Liberating the data: A proposal for a joint venture between Statistics Canada and Canadian universities (unpublished).

Watkins, W., & Boyko, E. (1996). Data liberation and academic freedom. *Government Information in Canada*, *3*(2) 2. Retrieved 26 June from http://www.usask.ca/library/gic/v3n2/watkins2/watkins2.html.

World Meteorological Organization (WMO). (n.d.). *History of the IMO*. Retrieved May 12, 2019, from https://public.wmo.int/en/about-us/who-we-are/history-IMO

About the Author

Tracey P. Lauriault is Associate Professor of Critical Media and Big Data, Communication and Media Studies, School of Journalism and Communication, cross-appointed to the master's program in Digital Humanities, and on the advisory board for the Institute for Data Science at Carleton University. She serves on the Multi-Stakeholder Forum for the Canadian Open Government Civil Society Network, and on the board for the Canadian Open Data Summit 2021, among other open data civil-society and government activities. She is Research Associate with the Maynooth University Social Science Institute in Ireland, the Geomatics and Cartographic Research Centre at Carleton University, the Institut national de la recherche scientifique in Montréal, and the Centre for Law, Technology and Society at the University of Ottawa. Lauriault is a citizen of the Algonquins of Pikwàkanagàn First Nation, the Republic of Ireland, and Canada.