CHAPTER FOURTEEN

DATA ACCESS IN CANADA: CIVICACCESS.CA*

Tracey P. Lauriault and Hugh McGuire

> Data constitutes a critical national resource, one whose value increases as the data become more readily and broadly available.[1]

There is a global movement to liberate government-'owned' data sets, such as census data, environmental data, and data generated by government-funded research projects. This open data movement aims to make these datasets available, at no cost, to citizens, citizen groups, non-governmental-organisations (NGOs) and businesses. The arguments are many: such data spur economic activity, help citizens make better decisions, and help us understand better who we are and where we are going as a country. Further, these data were collected using tax dollars, yet the government holds a monopoly which makes data available only to those able to pay the high access fees, while some data are not made available at all.

The open data movement is lagging in Canada as demonstrated by exorbitant fees for such basics as the data set of postal codes correlated to electoral districts. This data could be used for any number of civic engagement projects, but it costs thousands of dollars due to Statistics Canada's policies of cost recovery.

Currently, access to government data is hampered by four main factors: i) the high cost of available data sets; ii) arbitrary decisions about availability of data sets to the public; iii) restrictive licenses; and iv) inaccessible data formats.

Formed in 2007, Citizens for Open Access to Civic Information and Data[2] is a loose grouping of academics, activists, and citizens concerned with promoting data liberation in Canada. The grouping includes lawyers, copyright experts, librarians, archivists, cartographers, engineers, communications activists, open source programmers, and new media designers. The two main objectives of CivicAccess are:

- encourage all levels of governments (e.g. federal, provincial, municipal) and sectors (health, environmental, education) to make civic data and information available to citizens without restrictions, at no cost, in usable open formats

- encourage the development of citizen projects using civic data and information.

The long-term vision is a country in which citizens, specialists, professionals, academics, community groups and even businesses can work together, developing innovative information

---

[1] www.nap.edu/catalog.php?record_id=id4871.

[2] www.civicaccess.ca

access and visualisation tools, better decision-making models, and more tools responsive to the needs of the citizens. Liberating data will spur grassroots research on important social, economic, political and technical areas, currently hampered by lack of access to and high cost of civic data. Further, we want to link the debate about data to questions of government transparency and accountability, which pivot on access to accurate, reliable, and timely data.

But first, we need access to that data.

WHAT ARE CIVIC DATA?

Civic data are a public good, and more specifically, are 'numerical quantities or other factual attributes generated by scientists, derived during the research process through observations, experiments, calculations and analysis'.[3] They are also 'facts, ideas, or discrete pieces of information, especially when in the form originally collected and unanalysed',[4] and also, from the Report of the National Science Board, 'numbers, images, video or audio streams, software and software versioning information, algorithms, equations, animations, or models/ simulations'. Distinctions are made between raw or level 0 data and derived, refined, synthesised or processed data. Raw data are normally unprocessed; examples include digital signals from a sensor or an instrument (e.g. unprocessed satellite image, thermometer), facts derived from a sample collected for an experiment (e.g. blood sample, ice core), and facts collected by human observation (e.g. mine tailings, census). Computations and data manipulations are related to research objectives and methodologies. Refined or processed data are raw data that have been manipulated, undergone computational modelling, been filtered through an algorithm, sorted into a table or rendered into a map. In these cases, access to the models is as important as access to the output results of those data.

Civic data are the data created and maintained by public organisations and paid for by the public purse as part of the ongoing day-to-day activities of governing. Public data can include crime data at the neighbourhood scale, the number of traffic violations for certain streets, election results, census data, road networks, non-private health data, government expenditure data, school board catchment area boundaries, aggregated test results, environmentally sensitive or contaminated areas, or basic framework map data that include census areas, administrative boundaries, postal code areas and geo-referenced satellite images. Framework data are particularly important as these are the foundational data sets upon which other datasets can be organised. Civic data also includes those created as part of government funded research organisations such as the Social Sciences and Humanities Council of Canada (SSHRC) and the Natural Sciences and Engineering Research Council (NSERC) or any other outsourced publicly funded data and information creation activity.

TYPES OF OPEN DATA?

Some aspects of the open data movement include the following:

- Open Access (OA,[5] which aims to end restrictive licenses on university research and data as seen in initiatives such as Open Access News)

---

[3] nsf.gov/pubs/2005/nsb0540/.

[4] www.archivists.org/glossary/.

[5] wikipedia.org/wiki/Open_access.

- data visualisation projects which combine design and data in creative ways to make information more accessible, such as Gapminder[6]

- grassroots citizen projects using government data sets to improve cities and towns, such as FixMyStreet.[7]

CIVIC DATA ACCESS IN CANADA

Access to civic data in Canada depends on how much money you have, to which organisations you are a member, and for what purpose you want to use the data.

If you are a university professor or tuition paying university student in Canada, access to data is quite good. This is largely the result of work done by the Data Liberation Initiative (DLI) which is a data purchasing consortium.[8] DLI consortium members pay an annual subscription fee that allows their faculty and students unlimited access to numerous Statistics Canada public use microdata files, databases and geographic files. If you are a student or teacher in Ontario, you may access data from the new Ontario Data Documentation, Extraction Service and Infrastructure Initiative (ODESI)[9] which will target Statistics Canada datasets, datafiles from Gallup Canada and other polling companies, public-domain files such as the Canadian National Election Surveys, and selected files from the Inter-University Consortium for Political and Social Research (ICPSR). Both the DLI and ODESI provide access to a small subset of Canadian citizenry. Their license is very specific about who authorised users are, exclusivity, and how data products cannot be used such as 'in the pursuit of any contractual or income-generating venture either privately, or under the auspices of the educational institution'.[10]

If you work for a government, access to data varies depending on which department and level of government you are in, the rationalisation you have for acquiring that data, and the budget your department or section has. For instance, Environment Canada shares its data quite openly, as does Natural Resources Canada via the GeoConnections[11] GeoGratis.ca or the Geobase.ca[12] programs. In fact, Geobase.ca has one of the most progressive data licensing[13] programs so far seen in the Government of Canada. At the Canadian provincial or city scale, things start to get confusing as licenses differ, as do cost recovery and access policies. Land Information Ontario (LIO) has many data sets in their downloadable catalogue;[14] however, these data are only available through a Government of Ontario Intranet or between and among members of the Ontario Geospatial Data Exchange (OGDE).[15] Municipalities suffer from very restrictive or non-existent data sharing policies that are not uniform across departments.

As an example, the City of Ottawa has different categories of clients for its GIS data:

---

[6] gapminder.org/.

[7] www.fixmystreet.com/.

[8] statcan.ca/english/Dli/dli.htm.

[9] odesi.uoguelph.ca/wiki/index.php/Main_Page.

[10] www.statcan.ca/english/Dli/caselaw/assess.htm.

[11] www.geoconnections.org/.

[12] www.geobase.ca/.

[13] www.geobase.ca/geobase/en/licence.jsp.

[14] tinyurl.com/yufhn5.

[15] www.lio.gov.on.ca/en/Exchange.htm.

- category A, internal municipal clients: no charge for data and rarely require a license agreement

- category B, external municipal clients: are charged a fee to reflect the staff resources consumed in the preparation of the data and sometimes require a license agreement

- category C, external groups needing data for specific projects: are usually charged the same fee as category B clients and must also enter into a signed data license agreement naming a specific project or use

- category D, external groups wishing to commercially market the data: category D clients are expected to pay a fair market rate for any data they want to commercialise.

'for all requests it is expected that the client can demonstrate a legitimate use of the data. This provision ensures that staff resources are not unduly expended on frivolous requests. Additionally, the license must refer to a specific project or use as this helps the City track how the data is being used and by whom'.[16]

There is no 'citizen' category. How you can use, re-use, and represent data is quite restrictive. It would seem logical to have data discoverable and accessible via a data portal. This would result in the City not having to work so hard to micro manage the use of our public data.

Things get really confusing when different levels and departments of government repeatedly sell each other the same data sets with public money. Governments do not have intra-governmental data portals that centralise data acquisitions and share data assets amongst public servants. Duplication of effort and multiple layers of bureaucracy and accounting could be done away with by simply making all the data free to not only citizens but also their governments!

If you are from an NGO, data access is cost prohibitive. Many small NGOs pool their resources and develop data purchasing consortia such as the Canadian Council on Social Development Community Social Data Strategy.[17] However, like the DLI, these entities remain closed and exclusive shops. Statistics Canada allows a variety of companies to resell civic data and has also licensed a number of civic data value added distributors.[18]

As a citizen, you have access to incomplete data sets from the Depository Service Program[19] available to you in public libraries. These are suitable for high school projects but not for public participation in a democracy. What we really need is a concerted lobby in Canada that will free public data.

[Note: Since this chapter was written, the City of Ottawa has become a Canadian Open Data City (www.ottawa.ca/online_services/opendata/index_en.html).]

WHY FREE CIVIC DATA?

In a wider, less technical sense, 'data' are what we use to make decisions, so they are a public good. We use data sets to make decisions about how we as individuals should act, and how we as a society ought to do things. All the rules that govern our societies, from agricultural

---

[16] www.ottawa.ca/calendar/ottawa/citycouncil/occ/2010/05-12/csedc/08-DOC1-IT Policy - Data Dissemination.pdf

[17] www.ccsd.ca/subsites/socialdata/home.html.

[18] www.statcan.ca/english/reference/data.htm.

[19] dsp-psd.pwgsc.gc.ca/Reference/guides-e.html.

practices to cooking, to our law systems and social interactions, are the result of our interpretation of the interaction between different data sets over time.

Our ability to collect, analyse and interpret these data, and to make decisions based on them, is what gives humans our particular ability to solve societal problems such as food shortages, disease infestations, and resource depletion.

Democracy has a number of fundamental ideals, including free speech, free press, transparency of government, separation of powers, rule of law, public education, and free markets. All these principles are based on openness of information, or openness of data. In a sense, the basis of democracy is to open up the decision-making process to everyone.

By opening data to more people, you get more interpretations, more proposals of different solutions, better decisions about the best solutions, and in the long run, more successfully-solved problems. We have reached a time when the cost to share datasets is no longer cost prohibitive.

The processing power available on a desktop computer can do an enormous amount with even large datasets. Skilled designers have the ability to interpret, redesign, repackage, and display data in new and important ways, and the social web allows others to contribute to that process.

Transparency and accountability are essential elements of a functional participative democracy, and access to data and information is imperative. Transparency increases as quality data are widely and freely disseminated. Government and the private sector often miss important types of analyses, particularly local, cross boundary or jurisdictional research. For instance, it is cost-prohibitive and technically difficult for a community group to discover and access neighbourhood-scale data from different levels of government to conduct any kind of local community market or demographic analysis. An entrepreneur developing a business plan for a company to operate in four cities in two provinces would quickly discover restricted access to the basic data and information required to understand their market niche, clients, and competitors.

The basic digital data and information upon which we depend are rarely accessible, rarely interoperable, rarely in open formats, and are often prohibitively expensive. Moreover, regressive licensing regimes impede the sharing of data, or worse, there are no licensing regimes at all, which leaves citizens at the whim of the decisions of public servants. This is particularly true at the municipal and school board levels where a lack of clear guidelines often means no access to data for fear of releasing the wrong thing. For Canadian citizens this means that much innovation and knowledge is being thwarted. Worse, we often are forced to pay exorbitant prices for data to study important issues such as poverty,[20] homelessness,[21] or to assess the cost to the health care system of poor air quality.[22]

CIVIC DATA PROJECTS

Wikipedia was launched in 2001, and in seven years has displaced Britannica, the gold-standard English language reference encyclopaedia since 1768. Wikipedia has more articles, is more up-to-date, and, while the accuracy of the information in Wikipedia is a constant work-in-progress, Nature's December 2005 study of scientific articles in the two encyclopaedia found the

---

[20] datalibre.ca/2007/11/30/paying-for-data-to-study-poverty/.

[21] datalibre.ca/2008/01/18/housing-and-homelessness-data-in-canadian-cities/.

[22] www.oma.org/phealth/icap.htm.

accuracies to be roughly equivalent.[23] Wikipedia is the most useful encyclopaedia in the world, if, by useful we mean, 'the encyclopaedia that most people use'.

We are beginning to see more examples of civic projects. One example gets right to the nitty-gritty of municipal politics: potholes. Launched in February 2007, the UK project FixMyStreet.com 'is a site to help people report, view, or discuss local problems they've found to their local council by simply locating them on a map'.[24] The project targets such problems as potholes, broken streetlights, and graffiti. It has revolutionised municipal maintenance planning by putting the data collection into the hands of citizens and opening up the planning and decision-making process to many concerned citizens. Problem reports are there for all to see, providing municipal councils more incentive to fix the problems. Another amateur project that turns a light on the political process itself is howdtheyvote.ca, which tracks how Canadian members of parliament vote on individual bills -- information that should be fundamental to our understanding of our representatives in Parliament.

Crimereports.com[25] is a US site built to help citizens get more information about the locations and frequencies of crime incidents in their cities.

These examples of progressive initiatives suggest that we are in the early days of the movement towards opening up government data. Open data allows citizens to build tools that can address issues important to them. More tools of civic engagement through data are starting to appear on the web, and there is much to be done.

WHAT IS CIVICACCESS.CA DOING?

Civicaccess.ca[26] is about liberating public data from public institutions and finding new ways to make data accessible and useful. Individual members are doing incredible things. However, as a collective we have not tackled any big projects. We provide a mailing list[27] with over 200 members across the country that exchange information on issues, innovations, projects and ideas.

The authors of this paper also co-author DataLibre.ca, a CivicAccess.ca inspired blog, to fill a void on this topic. Its readership has been increasing and we are seeing traffic coming from key players in the open access movement, the open data and open source communities, along with members from library and archives associations. Ultimately, CivicAccess.ca is firing up the conversation on access to public data in Canada and we hope to discover and support the creation of innovative open public data projects. So come and join us!

CONCLUSION

Innovation comes from many drivers and sources, but there are two essential prerequisites: a problem in need of solving, and information and data. With a few other ingredients such as intelligence, creativity, and resources, innovation will occur. But the fundamental ingredients in

---

[23] tinyurl.com/yotjyh.

[24] fixmystreet.com/.

[25] crimereports.com/.

[26] civicaccess.ca/.

[27] lists.pwd.ca/mailman/listinfo/civicaccess-discuss.

innovation are always human desires to improve something, and figuring out, based on information, how to improve it.

Solving problems is one fundamental role of governments. By opening up civic data, and allowing citizens and citizen-groups to participate in problem solving, we believe that we will start to see more innovative and better solutions to the problems facing society.

Doing any form of research requiring cross jurisdictional civic data sources that cross domains, sectors and topics is very difficult in Canada. We have discussed the underlying reasons, examined some of the many bottlenecks and roadblocks, and highlighted examples of some progressive initiatives.

The technological solutions to provide free access to Canada's civic data are readily available and relatively inexpensive.[28] What is more difficult is finding the political will to make our civic data public.

RECOMMENDED RESOURCES

Klinkenberg, Brian, 'The True Cost of Spatial Data in Canada', *Canadian Geographer / Le Géographe canadien*, 47: 1, pp. 37–49, published online 20 May 2003, doi: 10.1111/1541-0064.02e11.

McMahon, Ronald C., Saskatchewan Bureau of Statistics, *Cost Recovery and Statistics Canada*. Available at: library2.usask.ca/gic/v2n4/mcmahon/mcmahon.html.

UK Guardian Free Our Data Campaign. Available at: www.freeourdata.org.uk/.

PodCast about CrimeReports.com. Available at: itc.conversationsnetwork.org/shows/detail3459.html.

Stephenson, W. David, *Let my data go! The Case for Transparent Government*. Available at: stephensonstrategies.com/speeches/let-my-data-go-the-case-for-transparent-government/.

---

[28] library2.usask.ca/gic/v2n4/mcmahon/mcmahon.html.