

Special issue on Data Literacy: Notes from the field

Data Literacy projects in Canada: field notes from the Open Data Institute, Toronto node

Andi Argast

Open Data Institute, Toronto, Canada

Corresponding Author.

andi.argast@gmail.com

Lydia Zvyagintseva

University of Toronto, Scarborough, Canada

lydia.zvyagintseva@utoronto.ca

Open data is flourishing in Canada, but there are few formalized data literacy initiatives. Civic technology organizations such as the Toronto Node of the Open Data Institute (ODI Toronto), in partnership with public institutions and advocacy groups, are helping to fill the gap in data literacy through workshops and accessible hackathons. These organizations are collaboratively pursuing the goal of ensuring that open data benefits more than just a minority of technologically privileged Canadians.

Introduction

Open data has firmly taken root in Canada. Evidence of its flourishing is apparent in the open data portals at all levels of government that have made hundreds of thousands of data sets available for public use, and in hackathons focusing on solving challenges as diverse as dementia (DementiaHack, n.d.), healthcare (Hacking Health, n.d.), traffic congestion in Toronto (TrafficJam, n.d.), and international development (Hackathons — Devs Without Borders, n.d.). Similarly, grassroots civic technology initiatives such as Civic Tech Toronto (Civic Tech Toronto, n.d.) and Data for Good chapters, akin to the UK's Data Kind project, are gaining traction in the provinces of Ontario and Alberta (Data for Good - Data Scientists & Devs doing GOOD, n.d.).

Canada's open data community is a hybrid of citizen innovation and government direction at the federal, provincial, and municipal levels. Open government data portals exist in nearly every province and territory, and in many instances municipalities have led the way in

Argast, A., Zvyagintseva, L. (2016). Data Literacy projects in Canada: field notes from the Open Data institute, Toronto node. *The Journal of Community Informatics*, 12(3), 223—231.

Date submitted: 2016-01-26. Date accepted: 2016-06-13.

Copyright (C), 2016 (the authors as stated). Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 2.5. Available at: www.ci-journal.net/index.php/ciej/article/view/1302.

opening data; notable early adopters are the cities of Vancouver, British Columbia and Edmonton, Alberta (McDonald and Wylie, 2014). Through the International Development Research Centre, Canada also sponsors the Open Data Research Network, which looks at open data uses in the developing world (Research project: Exploring the Impacts Emerging of Open Data in Developing Countries, n.d.). Many, if not most, of these open data initiatives aim to be inclusive and equitable, inviting citizens with varying skillsets and interests to participate in the open data, and by extension, open government landscape through use, repurposing, and building on existing data. However, given the scarcity of formal data literacy curricula and initiatives for average citizens, the extent of that inclusivity is arguably minimal.

Canadian information policy

Before examining the importance of data literacy in the Canadian context, it is helpful to understand how historical information policy precedents have shaped Canada's open data landscape; in particular, Canada's 1983 Access to Information Act (Government of Canada, 1983). Along with the access to information legislation, the 1980s gave rise to the marketization of information (Nilsen, 1997), a trend that was mirrored in the US and UK during the same time period (Bates, 2012). In the 1990s, the development of internet access also reflected public/market tensions; the private sector created for-profit digital infrastructure while simultaneously federal government funds linked schools, libraries, and remote communities to the Internet through the Connecting Canadians programmes (Longford and Shade, 2012). The 'under-recognized' tension between the market orientation of information and information as a public good continues to be a defining characteristic of information policy today (Orna, 2008) and has helped shape open data in Canada, creating an environment for hackathons as potential sites of economic development that function in an ostensibly civic framework. Emblematic of this attitude, former President of the Treasury Board, Tony Clement stated, "Data is Canada's new natural resource. The sky's the limit in terms of what we can do with this material (T. B. of C. Government of Canada, n.d.)." This commodification of data released for public use can then be understood as an input for economic growth in the neoliberal post-industrial society. In fact, despite well-intentioned claims that access to publicly funded data on topics such as community housing, transit and recreation statistics will enable citizens to engage more actively with their cities and governments, research suggests that the general public is not positioned to benefit from access to open data (Prado and Marzal, 2013). According to the Open Knowledge Foundation, opening up government data fosters transparency, releasing social and commercial value and participatory governance (What is Open?, n.d.). While there is room for commercial application of open data in this framework, its benefits should not be limited to business alone, and values such as government transparency and participation by citizens in the democratic process should be equally emphasized in government policy and process.

Open Data and socioeconomic considerations

At various levels of government, open data policies are often implemented using a combination of "education and training, voluntary approaches, economic instruments, and legislation and control" (Huijboom and van den Broek, 2011) and while Canada has made use of voluntary and economic approaches, from our perspective, open data training and education are lacking. The former approaches may reinforce existing socioeconomic

inequities in which more affluent, technologically sophisticated citizens have access to and can effectively use open data tools and technologies. Although demographic information about Canadian open data users remains anecdotal, Jo Bates' findings in the UK indicate a worrying bias toward a technocracy, where "the social background of core OGD [open government data] advocates tends to be elite: highly, often elitely, educated white men from the higher social classes (Bates, 2012)." Although Bates notes that grassroots OGD organizations were inclined to be more diverse, "it would still be accurate to describe it as a community with significant social privileges (Bates, 2012)." In the context of Canada's multicultural society, it is possible that a demographic study of the open data community here would reveal a less homogenous construction than in the UK. Yet, given that the structural inequities of the digital divide persist, a pattern of "empowering the empowered" (Gurstein, 2011) is certainly likely in Canada if care is not taken to include a diverse range of open data users in the processes of data access, use, and sharing.

Approaches to Data literacy

Data literacy and the ability to effectively use data (Gurstein, 2003) are critical measures to ensure an inclusive and equitable engagement with open data in Canada. Data literacy, and the related concepts of statistical literacy and basic numeracy, defines who will actively participate in the shaping of open data projects and who will be the passive recipient of these initiatives (Twidale, 2013). Without a data literate public, the benefits of open data tend to accrue to government itself and businesses (Prado and Marzal, 2013). Competencies such as understanding, finding, manipulating, and interpreting data are a useful starting point in discussing the objectives of data literacy, but it is important to consider the needs of a particular community before prescribing a data literacy curriculum on a broader scale. Arguably, strategically connecting citizens facing specific issues with the data sets that can help alleviate those problems is more valuable than a generalized, universal data literacy plan (McDonald and Wylie, 2014) and in fact, this is the approach that the Toronto Node of the Open Data Institute has been taking in our data literacy efforts over the past two years. Whether it is through more effective use of data (Gurstein, 2003) or data literacy programs, there is no doubt that making open data more accessible and useable to a broader cross-section of the public will ensure that the ideals of openness—accountability and transparency in particular—will remain relevant.

The Toronto node of the Open Data Institute

The Toronto Node of the UK-based Open Data Institute (ODI) launched in 2013, as the Canadian Open Data Institute (CODI), is a volunteer-run initiative with a mission "to support the advancement of Open Data access and use in Canada (Canadian Open Data Institute, n.d.)." The organization initially championed education, improving data literacy, and skills development, with early projects focusing on municipal data standards, a data request framework, and an academic open data research repository (Canadian Open Data Institute, n.d.). In 2014, the Open Data Institute revised its structure, eliminating any national nodes or franchises, and CODI became known as the Open Data Institute's Toronto Node. Toronto is one of only two North American nodes; the other is based in Chicago (ODI Nodes, 2015).

As a volunteer-run organization, ODI Toronto must be strategic in selecting its priority projects. In autumn 2015, we chose to support two key initiatives: a “Data Jam” at the annual Ontario Nonprofit Network Conference, and a public workshop and hackathon hosted by the Toronto Public Library. Data literacy was a key element in each of these projects, and in both instances, we focused on a basic introduction to open data concepts, rather than an advanced curriculum exploring numeracy or analysis. The following sections outline each project in detail.

The Ontario Nonprofit Network conference Open Data Jam

The Ontario Nonprofit Network (ONN) is an advocacy body for the province of Ontario’s not-for-profit sector, and the annual conference is a hallmark event for its members. For the 2015 conference, ODI Toronto partnered with Data for Good Calgary to deliver a two-part “Data Jam” or workshop to an audience of Ontario-based nonprofit and charitable organizations.

Before describing the sessions, it is helpful to understand the complexity of the relationship between open data and the Canadian nonprofit sector. Data about registered Canadian charities—excluding organizations that are incorporated as nonprofit entities—is publicly available from the Canadian Revenue Agency (CRA) through the federal government’s open data portal (2013 Charities Listings, n. d.). This data has been used by organizations such as the charity advocacy group Imagine Canada to create dashboards and charity look-up tools for nonprofit organizations (Charity Focus – Home, n. d.). Although the CRA’s regulatory requirements have created a wealth of data about charitable organizations, there is little evidence to suggest that most nonprofit organizations are able to leverage this data for business intelligence or other purposes. Studies of charitable and nonprofit organizations in Canada and the US indicate that many lack the ability to effectively use basic digital tools (The 9th Annual Nonprofit Technology Staffing & Investments Report, 2015), let alone possess the requisite data literacy to use resources such as the CRA data set.

Given our audience’s frame of reference, the goal of the ONN Data Jam sessions was to introduce the approximately 30 participants to the basic features of the open data landscape and then embolden them to seek out additional resources for future open data projects. The sessions also discussed the necessity of having basic digital infrastructure in place—e.g. a database or similar repository—in order for organizations to collect and analyse both their own data and other relevant open data sources. The first session of the Data Jam, entitled “Getting to Know Your Data + Civic Tech 101” provided a solid understanding of open data definitions and where to find relevant open data sets in federal, provincial, and municipal portals. The second session, “Managing the Data Lifecycle - Keeping Information Assets Alive” expanded on the theme presented in the first as well as exploring how organizations should be managing their own data. The approach for both workshops was to demystify open data for an audience that was familiar with the term, but had little practical experience with opening, cleaning, and manipulating data sets. To help participants understand how open data was relevant to their own organizations, the workshop explored the types of small data—typically programmatic and budgeting related—that all organizations generate but typically do not use effectively.

For organizations that are challenged by a lack of data literacy, partnerships with the public and private sector are often critical to data-informed projects. During our Data Jam, Geoff Zakaib from the Calgary chapter of Data for Good presented a case study of one such partnership from a successful DataThon held in May 2015 with Calgary's Distress Centre. Over the course of a weekend, volunteers from Data for Good Calgary correlated data from the centre's Crisis Line and 211 with data sets including weather patterns and suicide statistics. Results were analysed and visualised, and the insights generated from the Datathon justified the creation of a new data analyst role at the organization (Argast et al., 2015). This case study of a successful cross-sector partnership offers a promising avenue for nonprofit organizations lacking the internal resources to analyse their own data; however, establishing such a partnership is not without its challenges. Therefore, a core component of the ONN Data Jam was to familiarize the audience with the nascent Canadian 'civic tech' movement—grassroots groups of engaged and data literate citizens—and to position nonprofit workers to participate more fully in civic tech projects as subject matter experts.

Through our involvement with groups such as Civic Tech Toronto, we have found that these initiatives try to promote user centred design, a position that theoretically privileges nonprofit workers' experiences in a hypothetical project's data analysis and design process. Yet, we have also observed that it can be difficult for less data-literate individuals to assert themselves and their priorities in these highly technical environments. To help ease this tension, each Data Jam participant was tasked with completing a Civic Tech Product Brief Worksheet, which aimed to assist participants in thinking through all aspect of a data-related problem, and ultimately to better equip nonprofit workers to partner with civic tech groups and other data-focused organizations. The worksheet loosely followed a design-thinking process structured to draw out a more fulsome understanding of both the data-related problem and possible solutions. Once completed, the Creative Commons licensed and freely available worksheet can be used to frame discussions with potential partners. Post-workshop, ODI Toronto plans to connect interested nonprofit organizations with civic tech groups to bring some data-related projects to fruition, and is currently developing a workflow to make this possible. A further goal is to facilitate additional data literacy workshops for a similar audience, with a focus on using open source tools such as Open Refine to clean and analyse data.

Toronto Public Library workshop and hackathon

Libraries have played an important role in connecting Canadians to technology since the days of the government-funded Internet access programs in the 1990s (Longford et al, 2012). Recent initiatives such as the Toronto Public Library's Digital Innovation Hubs, which offer services ranging from 3D printing to Arduino and Raspberry Pi loans to drop-in Maker clubs, have continued to strengthen the relationship between libraries and citizens through the use of technology. Building on this affinity, ODI Toronto partnered with the Toronto Public Library (TPL) and the City of Toronto's open data team for TPL's first open data-focused initiatives in autumn 2015.

The first project with TPL was a free, publicly available evening workshop delivered by ODI Toronto that aimed to provide a foundation in the principles of open data, examine how open data functioned in the context of the City of Toronto, and explore relevant examples of open data usage on a municipal level. Similar to the ONN Data Jam, the goal of the TPL workshop was to promote data literacy by providing participants with a fundamental understanding of

open data, as well as to assess how open data could be used by citizens and by the library system itself. In order to ensure the workshop was accessible, we developed a curriculum that did not require participants to have or to use a computer, certainly a challenge when exploring the topic of open data. Instead, activities relied on non-digital resources such as chart paper and markers; for instance, part of the workshop explored the Data Spectrum, a concept developed by the ODI to help represent the differences between closed, shared, and open data (The Data Spectrum, n.d.). Workshop participants were encouraged to draw a simplified version of the spectrum on chart paper, and then use sticky notes to represent where they thought different types of data might lie on the spectrum. This activity enabled participants to discuss issues of data security and privacy, and to better understand the role of shared and open data in the everyday life.

The demographics of the TPL workshop participants were extremely varied and the workshop discussion reflected their diverse interests. Although many individuals indicated that this was their first exposure to open data concepts, throughout the workshop they continually requested information on free and open source tools for mapping projects and actively discussed how open data could facilitate more informed interactions with the City of Toronto's urban planning department. The participants also expressed interest in having additional workshops to focus on these issues. No formal success metrics for the workshop were in place, although the animated discussion and detailed questions we fielded at the conclusion of the event indicated that participants connected with the material. Clearly, open data is a highly engaging topic because it touches on many areas of civic life, resonating with many people in very intimate ways. As a forum for raising general awareness about open data and its role in city-building and library services, the inclusive and low-tech public workshop worked well. However, presenting information in a lecture style format without the aid of computers restricted how hands-on and technical the material could be, and ODI Toronto hopes to use the library's computers for future workshops.

Our second project, the Toronto Public Library's first-ever hackathon, was held over a weekend following the workshop. For this event, staff at the Toronto Public Library cleaned and released 36 data sets; this data was made available only for the duration of the event, as the TPL was still working to create sustainable and long-term open data policies and processes. Data sets included catalogue and circulation data, library branch information, and data about library programs, classes and services. Participants were introduced to and encouraged to draw on the City of Toronto's municipal open data portal, as well as Wellbeing Toronto (Wellbeing Toronto, n. d.), a mapping application that allows users to combine various socioeconomic indicators in a single visual interface. Subject specialists from the City of Toronto were on hand to assist participants with the municipal data sets. Additionally, help with Open Refine was available for anyone interested in cleaning datasets.

At the TPL hackathon, participants were tasked with creating project or idea prototypes that improved library services and aligned with the library's new strategic plan. In this manner, data literacy was encouraged at the hackathon through engagement with data sets and problem-solving through design. Specifically, the hackathon aimed to foster projects that answered the question of "How can the library make our communities more resilient, more knowledgeable, more connected and more successful?" (Velasco, 2015). The TPL hackathon aspired to be a highly accessible initiative and subsequently drew a wide demographic of participants: the youngest individual was 12 years old, and the oldest was a 75-year-old

retired programmer (Velasco, 2015). As mentors and facilitators at the hackathon, we found the diversity of the participants to be reflected in an inspiring array of answers, presented in the form of prototypes, to the framing question of the hackathon.

In the style of many other hackathons, prizes were awarded to several groups for their final prototypes, which were presented on the second day of the event. The winning participant's project mapped catalogue search results by branch, for instance allowing someone to quickly see which libraries had books in a particular language. This information could then be cross-referenced with municipal data to show how many speakers of that language lived near the library branch, thus creating a more knowledgeable, connected library service (Chua, 2015). Although several teams developed more traditional mobile applications that allowed users to better interact with the library's services and resources, other teams focused on issues of accessibility and data quality. One team developed a speech interface to help improve accessibility when interacting with the library's catalogue for visually impaired users, while three other teams focused on the library's data collection processes, as well as the quality and usability of the data that was made available for the event (Velasco, 2015). All ideas were well received by the TPL staff at the event who felt that these projects would be beneficial for advancing the development of a comprehensive open data policy for the library. Clearly, structured and systematic engagement with open data fosters creative problem-solving, advocacy, and critical reflection on government, civic, and nonprofit services, processes and policies. The case studies presented here are but two specific approaches to focused open data engagement that nevertheless center on literacy, capacity building, and civic participation.

Conclusion

Canada currently lacks any kind of consistent and official data literacy curriculum or a set of national initiatives. It thus falls to civic organization such as ODI Toronto, in partnership with libraries and other public institutions, advocacy groups, and forward-thinking educators to ensure that open data benefits more than just a minority of technically sophisticated Canadians from a particular socioeconomic background.

A comprehensive look at data literacy in Canada is needed to better understand what, if any, democratizing effects open data is having across the country. In the absence of a literacy strategy at the national level, it is crucial for Canada's disparate civic tech groups to build communities of practice that are inclusive and accessible by design. Cross-sector partnerships are also critical; our projects with established organizations such as the ONN and the TPL fostered accessibility to open data by allowing us to connect with individuals who may be excluded from 'traditional' civic tech initiatives. For our part, we plan to continue our data literacy efforts, developing these partnerships and our skill-building workshops to present more advanced open data concepts and to introduce Canadians to free and open source tools for working with open data at a municipal, and eventually, national level. Our hope is that these data literacy initiatives will spark similar projects from coast to coast, and eventually lead to a more democratized and inclusive open data landscape in Canada.

References

- 2013 Charities Listings. (n.d.). Retrieved from <http://open.canada.ca/data/en/dataset/48fa4b32-c92b-4912-b451-c37e590a716a>

- Argast, A., Wylie, B., & Zakaib, G. (2015). Nonprofit Data Jam! Getting to Know Your Data + Civic Tech 101.
- Bates, J. (2012). "This is what modern deregulation looks like" : co-optation and contestation in the shaping of the UK's Open Government Data Initiative. *The Journal of Community Informatics*, 8(2). Retrieved from <http://ci-journal.net/index.php/ciej/article/view/845>
- Canadian Open Data Institute. (n.d.). Home - Canadian Open Data Institute. Retrieved from <http://opendatainstitute.ca/>
- Charity Focus - Home. (n.d.). Retrieved from <http://www.charityfocus.ca/EN/Pages/Home.aspx>
- Chua, S. (2015). Exploring neighbourhood libraries and other notes from the Toronto Public Library Hackathon. Retrieved from <http://sachachua.com/blog/2015/11/exploring-neighbourhood-libraries-and-other-notes-from-the-toronto-public-library-hackathon/>
- Civic Tech Toronto. (n.d.). Retrieved from <http://civictech.ca/>
- Data for Good - Data Scientists & Devs doing GOOD. (n.d.). Retrieved from <http://www.meetup.com/DataforGood/>
- DementiaHack. (n.d.). Retrieved from <http://hackernest.com/dementiahack/>
- Government of Canada. Access to Information Act (1983).
- Government of Canada, T. B. of C. (2011, March 17). Minister Day Launches Open Data Portal [media release]. Retrieved from <http://www.tbs-sct.gc.ca/media/nr-cp/2011/0317a-eng.asp>
- Gurstein, M. (2003). Effective use: A community informatics strategy beyond the Digital Divide. *First Monday*, 8(12). Retrieved from <http://firstmonday.org/ojs/index.php/fm/article/view/1107>
- Gurstein, M. B. (2011). Open data: Empowering the empowered or effective data use for everyone? *First Monday*, 16(2). Retrieved from <http://firstmonday.org/ojs/index.php/fm/article/view/3316>
- Hackathons — Devs Without Borders. (n.d.). Retrieved from <http://www.devswithoutborders.org/breakpoverty/>
- Hacking Health. (n.d.). Retrieved from <http://hackinghealth.ca/>
- Huijboom, N., & van den Broek, T. (2011). Open data: an international comparison of strategies. *European Journal of ePractice*, 12. Retrieved from <http://www.epractice.eu/en/document/5290090>
- Longford, G., Moll, M., & Shade, L. R. (2012). There and Back to the Future Again: Community Networks and Telecom Policy Reform in Canada, 1995–2010. In A. Clement, M. Gurstein, G. Longford, M. Moll, & L. R. Shade (Eds.), *Connecting Canadians: Investigations in Community Informatics* (PDF). Edmonton, AB: AU Press.
- McDonald, K., & Wylie, B. (2014, March). *Opening up Open Data: Making Data Available to Everyone*. Girl Geeks Toronto, Toronto, ON.
- Nilsen, K. (1997). Social science research in Canada and federal government information policy : the case of Statistics Canada. |c1997. Retrieved from <http://hdl.handle.net/1807/10950>
- NTEN. (2015). The 9th Annual Nonprofit Technology Staffing & Investments Report. Retrieved from <https://www.nten.org/article/the-9th-annual-nonprofit-technology-staffing-investments-report/>
- ODI Nodes. (n.d.). Retrieved from <https://theodi.org/nodes>
- Orna, E. (2008). Information policies: yesterday, today, tomorrow. *Journal of Information Science*, 34(4), 547–565.
- Prado, J. C., & Marzal, M. Á. (2013). Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. *Libri*, 63(2), 123–134. <http://doi.org/10.1515/libri-2013-0010>

- Research project: Exploring the Emerging Impacts of Open Data in Developing Countries. (n.d.). Retrieved from <http://www.opendataresearch.org/project/2013/oddc>
- The Data Spectrum. (n.d.). Retrieved from <https://theodi.org/data-spectrum>
- TrafficJam. (n.d.). Retrieved from <http://www.trafficjam.to/>
- Twidale, M. B., Blake, C., & Gant, J. P. (2013). Towards a data literate citizenry. In iConference 2013 Proceedings. iSchools. Retrieved from <http://hdl.handle.net/2142/38385>
- Velasco, A. (2015, November 28). Great Ideas and Teamwork Displayed at Inaugural TPL Hackathon. Retrieved from http://torontopubliclibrary.typepad.com/digital_design_studio/2015/11/great-ideas-and-teamwork-displayed-at-inaugural-tpl-hackathon.html
- Wellbeing Toronto. (n.d.). Retrieved from <http://map.toronto.ca/wellbeing/>
- What is Open? (n.d.). Retrieved from <https://okfn.org/opendata/>