CHAPTER 9

Reflections on the Future of Open Data

Pamela Robinson and Lisa Ward Mather

Abstract
This chapter takes the form of an extended postscript, a bridge, rather, between the research conducted and shared by our authors and the future of open data in a world that is currently in a state of rapid flux. The COVID-19 global pandemic, the climate emergency, and our collective efforts to confront systemic racism are among the significant current challenges we face as a society. Each of these challenges, among others, has clear points of connection to data and evidence informing decision-making. These challenges reinforce the pressing nature of the central question of this book: What is the future of open data?

1. Challenges to the Future of Open Data

1.1 Broader Ecosystem Dynamics Impacting the Future of Open Data

The grant that funded the research shared in this book came to an end in early 2019, but this compilation of research was completed in early 2021. During this past year, the ongoing rise of smart cities, machine learning and artificial intelligence, and the COVID-19 pandemic present further challenges to the future of open data that will be explored here.

1.2 The Future of Open Data Needs to Reconcile Tensions with Smart City Efforts

The "smart city" movement continues to gain traction in Canada, despite the term taking on quite different meanings. Key features of a smart city typically include interlinked and networked systems which generate big data that are used to "manage and control urban life in real-time" (Kitchin, 2015). Smart cities generate large amounts of data—from sensors, software, social networking, and surveillance cameras, among other sources—which makes these projects part of the larger debates around open data and open government. As Canadian cities are now beginning to use smart city technologies, this raises questions about what "open" means in the context of a smart city. Many characteristics of a smart city could be designed to be open, such as sensor data that are made publicly available, open-source technology, or progress toward an open government policy. But just because they can be open does not mean they will. There is a huge proprietary advantage to being the "owner" of these large datasets, so the incentive to share the data openly and willingly is low.

Importantly, thus far, municipal governments do not always seem to recognize, or reflect on the significance of, the link between smart cities and open data when developing smart city projects. Round one of Infrastructure Canada's Smart Cities Challenge, which took place from the summer of 2018 to the summer of 2019, provided a window into the ways in which open data could find a place in smart city efforts. The Challenge itself embodied unique aspects of openness; for example, all the applications to the Challenge were required to be posted on municipal government websites at the time of submission (Infrastructure Canada, 2019) and applicants were asked to address the extent to which their projects would include elements of open data (Impact Canada, 2017). But an open challenge does not ensure open data outcomes. It is still early days for evaluating the extent to which "open" is a value the winners actively embrace, but as these projects move from plans into actions, it is an important issue to continue to track.

From the fall of 2017 through the spring of 2020, Waterfront Toronto and Alphabet's Sidewalk Labs began their 29-month interaction exploring a large-scale smart city master plan for 12 acres of land on Toronto's waterfront. Among the many points of contention in this project was the issue of how data—open or not—would be governed (Robinson & Biggar, 2021; Scassa 2020). Data trusts were floated as a

new governance framework that might help mitigate the tensions between open data, private data, privacy, and security (McDonald, 2019), but the early discussion around this intervention raised many more questions than answers.

A particular category of data produced in some smart city projects has called into question the benefit of an open-by-default approach to data. This is human behavioural data, which include, as Bianca Wylie (2018, p. 2) notes, "both aggregate and de-identified data about people." She argues that this type of data, "even when anonymous or aggregate, needs a special approach that may be hard to reconcile with openness" (p. 1). Human behavioural data, collected by sensors, cameras, software, and social media, are particularly problematic because they have value not only for governments, but also for the private sector. Wylie points out that there is a "legislative and policy vacuum regarding consumer protection and technology products, in particular in the context of data products" (p. 5). Because of this, open data policies that are meant to democratize data could become more about "commercialization and outsourcing" (p. 11). Wylie concludes that the process of opening up data should not be slowed "where it is working"; rather, "out of caution, some open data should not be published as such," but they could, under the right conditions, be shared with stakeholders (p. 6). This caution would protect the individual, as well as the wider public.

Public space in the smart city is at the centre of this dilemma (Robinson, 2018, 2019). Many vendors have products that gather data in public spaces, raising expectations that these datasets should be open because they are generated in public/open spaces. But corporations that implement smart city projects on behalf of municipalities can claim ownership of the data they collect in public space. This is problematic for three reasons. One, when public surveillance and data collection are permitted, individuals may no longer have a reasonable expectation of privacy. Two, citizens being tracked in public spaces have often not provided informed consent for their data to be collected, yet these data could be used to manipulate their future behaviour. And, three, if corporate interests own data that are of value to governments, it is possible that the data could be sold back to governments as a service. In this case, governments would be data consumers rather than data stewards.

Between the Smart Cities Challenge and Toronto's two-year engagement with Sidewalk Labs on the Quayside project (Robinson &

Coutts, 2019), it is clear that open data is not necessarily fundamentally linked to smart city efforts. The Open Smart City work discussed in Chapter 1 of this volume introduces one approach to better ensure this alignment. The leadership on this approach, coming from Tracey Lauriault, a contributing author to this volume, and Open North (see Lauriault et al. 2019), are now connected to the Smart Cities Challenge through the Community Solutions Network. Will this commitment to open processes and open data continue through other Infrastructure Canada initiatives in the post-COVID era? Will "open" be a fundamental value in future Canadian smart city efforts? Arriving at sensible answers to these questions requires an evaluation of the evolution of how open data figures in new smart city innovation efforts.

1.3 The Future of Open Data Needs a More Nuanced Approach to Whose Data are Gathered and Open

As the open data movement has continued to evolve, a significant tension is emerging. The open data movement's goals of transparent, inclusive, and accountable actions are in contrast to the myriad ways that data can be—and are—readily deployed with the opposite intent or outcomes.

Recent reflections and ongoing research about the potential impacts of artificial intelligence, automated decision-making and machine learning signal cautions for curators and users of open datasets. Poor, equity seeking, and racialized people are subject to more surveillance, and therefore data collection, than people who have more means and political access, and who are white (Eubanks, 2018). From *Weapons of Math Destruction* (O'Neil, 2016) to *Artificial Unintelligence* (Brossard, 2018) to *Black Software* (McIlwain, 2020) to *Data Feminism* (D'Ignazio & Klein, 2020), there is a rapidly expanding scholarship of critical data studies loudly asserting that data-driven efforts, if left unattended or unevaluated, will have a natural tendency to over-serve majority and dominant populations while simultaneously disadvantaging and sometimes harming others. First Nations, Inuit, and Métis Peoples have developed their own governance principles concerning data ownership, control, access, and possession (OCAP) in response to the collection of their data being weaponized against them (FNIGC, 2020).

The late spring of 2020 saw an increase in anti-Black, anti-Asian, and anti-Indigenous racism, leading to large public protests.

The relationships between racialized and equity-seeking communities and the police were central to these protests. While long recognized as problematic, these tensions are connected, in part to the kinds of data gathered about Black, Indigenous, and other racialized people, and how law-enforcement organizations use data (D'Ignazio & Klein, 2020). As the COVID-19 pandemic evolved, public-health disparities in Black, Indigenous, and other racialized communities emerged (Bascaramurty, 2021). The absence of good public-health data in these communities undermined and delayed the delivery of care and support. There is simultaneously too much and not enough data being collected in these communities. The spring of 2020 gave rise to calls for more open, transparent, and accountable data gathering, use, and deployment. So, despite the democratic and inclusive ideals driving the open data movement, the future of open data must attend to these disparities. Open data communities whose membership and leadership do not mirror the diversity of the communities in which they work need to begin building new relationships, with the long-term goal of seeing change over time. Data-driven efforts to address economic, social, spatial, and ecological inequities need to centre the leadership and experiences of the community members experiencing the inequities or these efforts might further entrench persistent settler, colonial, and/or systemically racist systems and practices.

Publishing data often fails to achieve meaningful "awareness" or insight because making sense of data is not easy. As Jean-Noé Landry and Merlin Chatwin describe (2018, p. 4):

> Opening data does not automatically create a data literate public. City officials need to work with potential data users to ensure that they have the right skills to use the data. Officials themselves often require more training to be able to publish and use high quality data. . . . For [many cities], open data has been integrated into their strategy, but it still lacks sufficient human and financial resources to result in meaningful social impact.

One general concern is that, once launched, open data portals seem static and dated, like "abandoned last-minute science fair projects, pie charts sagging because someone didn't use enough glue stick" (Mulholland, 2016). Technology enthusiasts optimistically believe that "if you build it they will come" but research suggests

otherwise (Sieber et al., 2016; Sieber & Johnson, 2015; Johnson et al., 2017). The acts of opening the portal or gathering the data are not enough. For an open data ecosystem to thrive, open data advocates and users across public, private, and civil-society sectors continue to see the transformative potential of open data, and continue to work to achieve its many elusive goals.

There is a good deal of agreement across sectors, including among governments, non-profit organizations, and community groups, about what open data initiatives need to do now. Inside government, staff are routinely having to make the business case for further investments in open data (Robinson & Johnson, 2016). Beyond the entrepreneurial use of these datasets, and despite the open data movement's foundational commitment to democratic principles, there is a need for civic infomediaries (Robinson & Ward Mather, 2017) to advocate for open data release that serves public and/or civic intent as well. Non-profit groups, such as Code for Canada, Code for America, and the Open Data Institute, work to improve government, and the use of government data, from the inside by embedding technologists on fellowships inside governments to help bring new thinking and mobilize new ways of working. From civic hackathons (Costanza-Chock, 2020; Johnson & Robinson, 2014; Robinson & Johnson, 2016) to the leadership of public libraries helping community members begin to understand and use data, the work of civic infomediaries continues to hold space for open data use for the public good.

The COVID-19 pandemic presented a new challenge to the open data movement in that, despite the digital nature of open data, the work of open data civic infomediaries has historically relied on people working together in person to mobilize open data use. Civic hackathons tend to gather people in physical locations. Co-working spaces (e.g., WeWork, the Centre for Social Innovation in Toronto) and civic technology hubs including Civic Hall (New York) and Civic Hall Canada (Toronto) have been built around the belief that by gathering like-minded people together, creative combustion can emerge from the collision of ideas and people in shared spaces. Civic technology groups across North America (e.g., Smart Chicago, Civic Tech Toronto) regularly meet in person or virtually, via weekly hack nights, for example. The Code for Canada and Code for America fellowship models have their fellows working inside government offices, side by side with government staff. Public libraries have made significant investments in data-literacy programming and access to shared

technology through lending programs and innovation hubs. All these examples have connected digital datasets with physical and material locations aimed at animating their use. If the civic future of open data depends upon civic infomediaries continuing to gather people to share ideas and to collaborate, then the future of open data will, like many other important pursuits, need to evolve new techniques for working together.

2. The Future of Open Data is . . . Emergent and Evolving

Academic research is easier to conduct on static or completed subject matter. Research in real-time in collaboration with partners presents a wide range of challenges. The research conducted here, with its focus on the future of open data, has straddled the opportunity to evaluate the ongoing evolution of open data ecosystems while also tracking a series of conditions that are impacting how that ecosystem may continue to evolve. As open data ecosystems have matured, the research shared here sheds new light on the nuance and texture needed in the kinds of data gathered and deployed, in the governance frameworks to regulate and advance open data use, and as concerns the participation and engagement by open data actors.

Thus far, open data ecosystems have shown their capacity to adapt and respond to these changing dynamics. The persistence of the COVID-19 pandemic, among other significant societal challenges, suggests further turbulence and challenge ahead. Across these challenges it is clear, the work of opening data for private and public good is an asymptotic pursuit that will require ongoing attention, investment, and evaluation, and refinement and revision of actions.

References

Bascaramurty, D. (2021, January 26). Racialized Canadians have some of the highest rates of COVID-19 infections in the country. Who can allay their doubts about taking the vaccine? *Globe and Mail*. https://www.theglobeandmail.com/canada/article-racialized-canadians-need-the-covid-19-vaccine-more-urgently-than-most/

Brossard, M. (2019). *Artificial Unintelligence*. MIT Press.

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. MIT Press.

D'Ignazio, C., & Klein, L. (2020). *Data feminism*. MIT Press.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

First Nations Information Governance Centre (FNIGC). (2020). *Introducing a First Nations data governance strategy*. https://fnigc.ca/news/introducing-first-nations-data-governance-strategy.html

Impact Canada. (2017). *Smart Cities Challenge: Applicant guide*. https://impact.canada.ca/en/challenges/smart-cities/applicant-guide

Infrastructure Canada. (2019). *Smart Cities Challenge*. https://www.infrastructure.gc.ca/cities-villes/index-eng.html

Johnson, P. A., Corbett, J. M., Gore, C., Robinson, P., Allen, P., & Sieber, R. (2015). A web of expectations: Evolving relationships in community participatory geoweb projects. *ACME: An International E-Journal for Critical Geographies*, *14*(3), 827–848.

Johnson, P. A., & Robinson, P. (2014). Civic hackathons: Innovation, procurement, or civic engagement? *Review of Policy Research*, *31*(4), 349–357. https://doi.org/10.1111/ropr.12074

Kitchin, R. (2015). Making sense of smart cities: Addressing present shortcomings. *Cambridge Journal of Regions, Economy and Society*, *8*(1), 131–136. https://doi.org/10.1093/cjres/rsu027

Landry, J. N., & Chatwin, M. (2018). Making cities open by default: Lessons from open data pioneers. Retrieved August 31, 2021, from https://opennorth.ca/publications/2ce88qelsh6m30thifiniy_en

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

McDonald, S. (2019) *Reclaiming data trusts*. Centre for International Governance Innovation. https://www.cigionline.org/articles/reclaiming-data-trusts

McIlwain, C. D. (2020). *Black software: The Internet and racial justice, from the AfroNet to Black Lives Matter*. Oxford University Press.

Mulholland, J. (2016, March 10). 6 ideas to help government realize open data's transformative power. *Government Technology*. https://www.govtech.com/data/6-Ideas-to-Help-Government-Realize-Open-Datas-Transformative-Power.html

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishers.

Robinson, P. (2019). Public space in a smart city. In N. Ahmed, M. Claudel, Z. Ibrahim, C. Pandolfi, & B. Wylie (Eds.), *Some Thoughts* (p. 43). https://some-thoughts.org/

Robinson, P. (2018). Design interventions for a digital venture. *Spacing*, *46* (Spring), 30–31.

Robinson, P., & Biggar, J. (2021). Seeing the city as a platform: Is Canada's Smart Cities Challenge a step in that direction? In A. Zwick & Z. Spicer

(Eds.), *The platform economy and the smart city: Technology and the transformation of urban policy*. McGill-Queen's University Press.

Robinson, P., & Coutts, S. (2019). The case of Quayside, Toronto, Canada. In L. Anthopoulos (Ed.), *Smart city emergence: Cases from around the world* (pp. 330–350). Elsevier.

Robinson, P., & Johnson, P. A. (2016). Civic hackathons: New terrain for local government citizen interaction? *Urban Planning*, *1*(2), 65. https://doi.org/10.17645/up.v1i2.627

Robinson, P. J., & Ward Mather, L. (2017). Open data community maturity: Libraries as civic intermediaries. *Journal of the Urban and Regional Information Systems Association*, *28*, 31–38.

Scassa, T. (2020). Designing data governance for data sharing: Lessons from Sidewalk Toronto. *Technology & Regulation*, *2020*, 44–56. https://techreg.org/index.php/techreg/article/view/51

Sieber, R. E., & Johnson, P. A. (2015). Civic open data at a crossroads: dominant models and current challenges. *Government Information Quarterly*, *32*(3), 308–315.

Sieber, R. E., Robinson, P., Johnson, P. A., & Corbett, J. M. (2016). Doing public participation on the geospatial web. *Annals of the American Association of Geographers*, *106*(5), 1030–1046. https://doi.org/10.1080/24694452.2016.1191325

Simpler, Faster, Better Services Act, 2019, S.O. 2019, c. 7, Sched. 56.

Wylie, B. (2018). *Open data endgame: Countering the digital consensus* (CIGI Paper No. 186). Centre for International Governance Innovation. https://www.cigionline.org/publications/open-data-endgame-countering-digital-consensus

About the Authors

Pamela Robinson is Professor and Director of the School of Urban and Regional Planning at Ryerson University. Throughout her career as a planner, her research and practice have focused on complex, emergent challenges that Canadian communities face. Her current research focuses on the question of who is planning the Canadian smart city? She writes for Spacing.ca about sustainability, technology, and civic engagement in Canadian cities. Robinson was a member of Waterfront Toronto's Digital Strategy Advisory Panel, was an inaugural member of the Multi-Stakeholder Forum for the Government of Canada's Open Government Partnership, and is an advisor on the Toronto Public Library's Innovation Council.

Lisa Ward Mather is an Urban Planner who graduated from Ryerson University's Master of Planning program. She has done research in

planning policy and is also interested in the power of new technology to improve planning processes for the public sector, the private sector, and the public. She has worked on projects that explore the role of planning around municipal open data, artificial intelligence, civic technology, and smart cities technology.