Opinion

Data Trusts and the Governance of Smart Environments: Lessons from the Failure of Sidewalk Labs' Urban Data Trust

Lisa M. Austin

University of Toronto, Canada
lisa.austin@utoronto.ca

David Lie

University of Toronto, Canada
lie@cccg.toronto.edu

Abstract

Data trusts are an increasingly popular proposal for managing complex data governance questions, although what they are remains contested. Sidewalk Labs proposed creating an “Urban Data Trust” as part of the Sidewalk Toronto “smart” redevelopment of a portion of Toronto’s waterfront. This part of its proposal was rejected before Sidewalk Labs cancelled the project. This research note briefly places the Urban Data Trust within the general debate regarding data trusts and then discusses one set of reasons for its failure: its incoherence as a model. The Urban Data Trust was a failed model because it lacked clarity regarding the nature of the problem(s) to which it is a solution, how accountability and oversight are secured, and its relation to existing data protection law. These are important lessons for the more general debate regarding data trusts and their role in data governance.

Introduction

In 2018, a year after first announcing their partnership, Waterfront Toronto and Sidewalk Labs set out the preliminary terms for the proposed “smart” redevelopment of a portion of Toronto’s waterfront (“Sidewalk Toronto”) and indicated that a guiding design principle would be to “explore novel forms of data governance,” including the establishment of a “data trust” that “could carry a fiduciary responsibility to serve and balance data subject and public interest” (Toronto Waterfront Revitalization Corporation and Sidewalk Labs LLC 2018). Questions about privacy and data governance were central to the public debate about the project from the start, and the data trust proposal was meant to take the project beyond compliance with existing Canadian privacy laws to offer both stronger privacy protections and a stewardship mechanism through which to ensure data use was for the public benefit.

After more consultation, Sidewalk Labs provided a more concrete model it called the “Urban Data Trust” (Sidewalk Labs 2019). This had a short life. In June 2019, the Waterfront Toronto Board Chair indicated that it needed more information regarding whether Sidewalk Labs’ proposals regarding data governance were compliant with the project’s original digital governance principles as well as with Canadian laws (Diamond 2019a). By November 2019, and after serious criticism from Ontario’s Information and Privacy Commissioner, the Urban Data Trust proposal had been dropped (Beamish 2019; Diamond 2019b). With the demise of the Urban Data Trust proposal, responsibility for data governance shifted to Waterfront Toronto and the various levels of government but was still unresolved when Sidewalk Labs announced the cancellation of the project, a cancellation that it said was due to the global pandemic and accompanying economic uncertainty (Doctoroff 2020).

Austin, Lisa M., and David Lie. 2021. Data Trusts and the Governance of Smart Environments: Lessons from the Failure of Sidewalk Labs' Urban Data Trust. *Surveillance & Society* 19(2): 255-261.

<https://ojs.library.queensu.ca/index.php/surveillance-and-society/index> | ISSN: 1477-7487

© The author(s), 2021 | Licensed to the Surveillance Studies Network under a [Creative Commons Attribution Non-Commercial No Derivatives license](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Sidewalk Labs' Urban Data Trust model failed before the project was cancelled, and yet, despite this failure, the popularity of the idea of data trusts persists at the municipal, provincial, and federal levels in Canada. For example, in December 2019, the City of Toronto began public consultations regarding its Digital Infrastructure Plan, and one of the ideas it indicated that it was considering was a "municipal data trust" (City of Toronto 2019: 13). At the provincial and federal levels, data trusts are being discussed as part of law reform proposals (Ministry of Government and Consumer Services 2019; Government of Canada 2019).

The Urban Data Trust proposal, and its failure, provides important lessons for these other data trust proposals. While the story of the Urban Data Trust is also, in part, bound up with the broader politics of the Sidewalk Toronto project, including civic opposition to the project through #BlockSidewalk, this note focuses on the basic incoherence of the model. As we will outline in more detail, what it lacked was clarity regarding the nature of the problem(s) to which it is a solution, how accountability and oversight are secured, and the relation between such new models of data governance and existing data protection law.

Data Trusts

The term "data trust" is used in a number of diverse, and sometimes conflicting, ways. One set of meanings converges around providing ways to reduce technical and legal complexities regarding data sharing, including the privacy concerns that arise from data protection law—the public and private-sector laws modeled on the Fair Information Practice Principles that regulate "personal information."¹ For example, a prominent UK report describes data trusts as a solution "to improve trust and ease around sharing data" and talks about them in terms of "proven and trusted frameworks and agreements" (Hall and Pesenti 2017: 2). The language of "trust" in this context does not refer to the model of a legal trust but instead describes trusted methods for data sharing. Other models that focus on the complexity problem make use of data intermediaries of some kind. For example, the Web Science Institute focuses more on the technical aspects, stating that the "basic idea of a data trust is a virtual place where data is made available to share" (O'Hara 2019: 17). In other works, we have proposed a model of "Safe Sharing Sites" to manage the privacy and security issues associated with data sharing (Austin and Lie 2020).

Another set of meanings converges around providing a novel data governance mechanism for data stewardship. For example, the Open Data Institute (ODI) (2019: 2) recently defined a data trust as "a legal structure that provides independent stewardship of data," while others have defined them as an agile structure through which to create new models of governance (Wylie and McDonald 2018). Even in this governance context, the language of "trust" has been a source of confusion and contestation. In law, a "trust" is a very specific kind of legal arrangement where a trustee manages trust property on behalf of the trust beneficiaries. In its review of data trust models, the ODI concluded that a legal trust was actually not a suitable model for a data trust although it continued to use the language of trust. However, the ODI came to its conclusion regarding the unsuitability of a legal trust based on a number of erroneous claims regarding trust law, which have been helpfully refuted by Jun, Penner, and Wong (2019). Others have defended the legal trust model, arguing that it can empower individual data subjects and enhance individual control over personal information (Delacroix and Lawrence 2019; Rinik 2020). One of the reasons why the model of a legal trust has been so attractive to some when thinking about how to ensure responsible decision making is because trustees have a fiduciary obligation to their beneficiaries (Fewer et al. 2019). Indeed, quite apart from ideas of legal trusts or other legal models for data trusts, there is a burgeoning literature on the potential role—and potential problems—of fiduciary obligations in relation to data (Balkin 2016; Khan and Pozen 2019). Increasingly there is also a recognition of the need to have different kinds of governance structures

¹ Internationally, the GDPR (2016) is the most well-known example. Within Canada, the statutes most applicable to the Sidewalk Toronto project are PIPEDA (2000), FIPPA (1990), and MFIPPA (1990). This type of legislation regulates "personal information" or "personal data," which refer to information that is about an "identifiable" individual.

with distinct names, including data cooperatives (Ada Lovelace Institute 2021) and data commons (Scassa 2020).

At the time when Sidewalk Labs was developing its Urban Data Trust model, data trusts were considered to be everything from legal models to organizational practices to standardized methods to technical platforms and data stores that address data sharing; they could seek to facilitate data sharing through reducing complexities or seek to ensure the protection of important interests through enhanced obligations or participatory governance processes. The problem is that these are all very distinct concerns. Enhancing individual control over personal information is one thing, ensuring public access to non-personal information is another. Addressing privacy considerations yields one set of strategies; addressing questions of data harms and data justice in the use of non-identifying information yields a very different set of strategies.

Where does the Urban Data Trust fit within this landscape? As we outline in the next section, it was primarily a data governance model rather than one meant to manage the complexities of data sharing. This was largely because it proposed to manage privacy issues through de-identifying any personal information. However, we argue that this strategy was problematic as it actually reduced accountability and oversight in relation to privacy while increasing complexity. In terms of a governance model, Sidewalk Labs followed the ODI conclusions regarding the unsuitability of the legal trust model and, despite continuing to use the language of “trust,” proposed a non-profit corporate structure for its Urban Data Trust (Sidewalk Labs 2019: 423). However, we argue that its resulting model was in reality two models put together incoherently. One model was an open data vision, which contemplated broad public access to data. The other was a stewardship model that contemplated applying a range of norms to data use decisions. But these two models have different underlying assumptions and require different forms of accountability and oversight. As a result, the Urban Data Trust was a failed model.

Lessons from Sidewalk Labs

What was the Urban Data Trust? According to Sidewalk Labs’ Master Development and Innovation Plan, the proposed Urban Data Trust would “oversee matters of the digital governance of urban data for the [development], including the approval and management of data collection devices placed in the public realm, as well as addressing the challenges and opportunities arising from data use, particularly those involving algorithmic decision-making” (Sidewalk Labs 2019: 420). “Urban data” is a special term coined by Sidewalk Labs to include non-personal data, aggregate data, de-identified data, and personal information “collected and used in physical or community spaces where meaningful consent prior to collection and use is hard, if not impossible, to obtain.” The Urban Data Trust would not be a legal trust, but a non-profit entity with a board of five members representing different skills and constituencies. The board would hire a Chief Data Officer who would oversee the development of a charter, Responsible Data Use (RDU) Guidelines, and oversight and review processes (Sidewalk Labs 2019: 421). Any party proposing to collect or use “urban data” within the development would need permission from the Urban Data Trust and enter into a contract that would “include parameters that govern the collection, use, and destruction of urban data” (Sidewalk Labs 2019: 421). The MDIP speculates that a longer-term version of the Urban Data Trust would be to transform it into a public-sector or quasi-public sector agency through enabling legislation (Sidewalk Labs 2019: 422).

If this describes the basic structure of the Urban Data Trust, the first question we need to ask is: what problem was it trying to solve? Put in terms of our broad categorization of data trust proposals from the previous section, this proposal was responding to data governance issues rather than reducing technical and legal complexities. Sometimes the proposal spoke of using the trust to take privacy protection beyond existing laws, and sometimes the proposal spoke of using the trust to address the many additional considerations of data monopolies, group privacy, group harms, and group benefits that arise when considering ubiquitous data collection in the public (and quasi-public) realm.

The core problem with the Urban Data Trust was that it pursued two different data governance strategies simultaneously and the result was an incoherent model. On one hand, the Urban Data Trust embraced an “open data” model where the foundational idea was that urban data would be “publicly accessible by default” through open-data portals (Sidewalk Labs 2019: 238). This strategy rested upon the assumption that if urban data included any “personal information”—information that is identifiable and generally regulated by data protection laws—then this would be “de-identified by default” (Sidewalk Labs 2019: 456). Assume for a moment that this is possible. If so, then the open data strategy is the flip side of data protection law: if the data are not personal information (which individuals should control) then nobody should control it (including corporations). The point of the Urban Data Trust would be to instill trust in the public that Sidewalk Labs was not going to have preferred access to the data. On the other hand, some of the questions that began to emerge in the public discussions associated with the Sidewalk Toronto project called this open data strategy into question, raising important issues of the nature of power and public benefit, group privacy, data harms, data justice, and algorithmic accountability—even in relation to personal information that was de-identified or information that was never identifying. If data protection law is premised on individual control of information, and open data is premised on nobody controlling information, then the other alternative is collective governance of information. The Urban Data Trust incorporated collective governance concerns into its design by proposing to manage a more capacious category of data (“urban data”) than personal information and to subject this category of urban data to additional data norms, such as the RDU Guidelines.

The result was unworkable. Some of the deep tensions in this double-model show up in the fact that the Urban Data Trust pursued an open data model for the “disclosure” of urban data (including its further downstream uses) but pursued a collective governance model for the “collection” and “use” of urban data. For example, the Urban Data Trust was supposed to approve any project in the district that would involve urban data collection or use (Sidewalk Labs 2019: 420). This made the trust more like some versions of land stewardship, like British Columbia’s Islands Trust, which invests significant planning authority into a trust council and local trust committees to make decisions regarding the acquisition and use of lands for conservation purposes (*Islands Trust Act* 1996). By analogy, the Urban Data Trust was envisioned as a trust to manage urban spaces from a data perspective instead of a conservation perspective. The problem with the land analogy is that it is odd to speak of stewarding land while giving it away, whereas data are non-rivalrous and so stewardship of data does include decisions regarding disclosures. Indeed, many of the reasons for wanting the Urban Data Trust to review data collection and use—that there are important norms regarding data harms and data justice that go beyond either individual privacy or ensuring public access—are also reasons for wanting the Urban Data Trust to review disclosure.

Leaving aside the important question of whether land stewardship is a good model for data stewardship, it is important to also note that BC’s Islands Trust is created by a complex statute and not private law tools, which allows for the delegation of public authority (including planning authority), resolves questions of overlapping authority, and provides mechanisms of accountability and oversight. By contrast, the Urban Data Trust would have been created through the tools of private ordering and, because of this, failed to resolve issues of conflicts with municipal planning powers or the democratic governance mechanisms of the City of Toronto. Although the Urban Data Trust was promoted as providing a heightened level of protection for data, it did not incorporate additional protective norms like fiduciary obligations. The entity was to be established, at least initially, through an ordinary contract, and some of its key decision-making regarding collection and use was to involve a risk/benefit analysis and a determination of the “net balance of impact” (Sidewalk Labs 2019: 432).

The Urban Data Trust’s approach to technical and legal complexities regarding data flows was also deeply problematic and bound up with its open data strategy. The project’s basic strategy for managing privacy issues was to require any personal information collected to be “de-identified”—or rendered non identifying—thereby taking it outside the ambit of data protection law and its accompanying legal complexities. For example, Sidewalk Labs proposed that the RDU Guidelines require that personal information collected in the development to be “de-identified by default” (Sidewalk Labs 2019: 456). This

speaks to a technical vision for reducing complexities—that through innovative techniques, like “de-identification,” we can resolve privacy issues. However, determining whether data are “de-identified” is highly contested (Ohm 2010), involves legal interpretation and not just technical methods, and is recognized to require ongoing oversight (WP29 2014). Canada’s regulators have generally considered it part of their role to decide whether an organization is collecting personal information or not, as this determines whether the data processing is regulated by the statutes they are charged with overseeing. For example, in the recent investigation into Cadillac Fairview’s use of Anonymous Video Analytics (AVA) in its shopping malls, the Office of the Privacy Commissioner of Canada determined that the AVA technology used in Cadillac Fairview’s wayfinding kiosks was collecting and using personal information despite claims that this information was de-identified at the source (Office of the Privacy Commissioner of Canada 2020).

In addition to its faith in technical solutions to privacy questions, the Sidewalk Labs proposal had faith in using the tools of private ordering to provide *more* privacy protection than existing data protection laws. There is some truth to this in that if the City of Toronto, rather than the Urban Data Trust, decided to collect personal information, then it would be governed by Ontario’s public sector data protection laws, which would not require de-identification by default. However, the Urban Data Trust did not augment data protection law so much as bypass it. Because “urban data” included (de-identified) personal information but was broader than this, the relationship between the Urban Data Trust and data protection law was unclear (Beamish 2019). Even more problematic is the fact that, because the Urban Data Trust was to be a non-profit corporation, it would not be subject to either Canada’s federal private sector data protection law or provincial public sector data protection law. In other words, the Sidewalk Labs’ proposal would actually take the Urban Data Trust entirely outside of the public regulatory framework for privacy protection. At the same time, any organization involved in collecting information would have to independently assume responsibility for compliance with this regulatory framework while also having to manage requirements imposed by the Urban Data Trust. The proposal therefore *increased* legal complexity while at the same time reducing accountability and oversight apart from vague statements that the Urban Data Trust could eventually be transformed to a public entity through enabling legislation.

In sum, the Urban Data Trust was a failed model because it lacked clarity regarding the nature of the problem(s) to which it is a solution, how accountability and oversight are secured, and its relation to existing data protection law.

Conclusion: The Way Forward

We can understand the failed Urban Data Trust as a response to the problems of the current regulatory environment and the fact that Canada’s existing data protection laws provide an inadequate structure for governing data in smart environments. The Urban Data Trust was not wrong in thinking that smart environments require new data governance models even if its own model was incoherent.

The mismatch between data protection law and “smart” environments is fundamental. Data protection law is premised on ideas of individual control over personal information and centered on relationships between individuals and (private or public) organizations. But collecting data in smart environments is not easily modelled on the intentional sharing of personal information with an organization providing you with a product or services. Instead, it involves opaque collection of information that may or may not be about people and may or may not be identifiable. This collection is mediated through ubiquitous sensor technology as a by-product of the use of public space, and it is meant to enable multiple future (and often unspecified) innovative uses. Data becomes part of our public *infrastructure* rather than a feature of a particular transaction or relationship, unmoored from specific individuals and from specific projects and purposes, and created through complex private-public partnerships.

Data trusts—and the many other models of data governance that are emerging—can potentially play an important role in managing data in the complex data ecosystem of the twenty-first century, including smart environments. However, we need much more clarity regarding the nature of the problems to which specific

models are solutions, the nature of the interface between these models and data protection law, and the nature of their accountability and oversight mechanisms. It is far from clear that we can do this by simply relying upon the existing tools of data protection law and then augmenting this through private ordering (contracts, corporate structures, legal trusts) and technical means (de-identification) rather than asking the more difficult law reform questions of how to *create* the necessary legal infrastructure for governing data-as-infrastructure.

References

- Article 29 Working Party (WP29). 2014. Opinion 05/2014 on Anonymisation Techniques. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf [accessed May 14, 2020].
- Austin, Lisa, and David Lie. 2020. Safe Sharing Sites. *New York University Law Review* 94 (4): 581–562.
- Balkin, Jack. 2016. Information Fiduciaries and the First Amendment. *UC Davis Law Review* 49 (4): 1183–1234.
- Beamish, Brian. 2019. Re: Sidewalk Labs’ Proposal. Information and Privacy Commissioner of Ontario, September 24. <https://www.ipc.on.ca/wp-content/uploads/2019/09/2019-09-24-ltr-stephen-diamond-waterfront-toronto-residewalk-proposal.pdf> [accessed October 7, 2020].
- City of Toronto. 2019. Digital Technologies and Your City Public Consultations. <https://www.toronto.ca/community-people/get-involved/public-consultations/#item/746> [accessed October 7, 2020].
- Delacroix, Sylvie, and Neil Lawrence. 2019. Disturbing the “One Size fits All” approach to Data Governance: Bottom-up Data Trusts. *International Data Privacy Law* 9 (4): 236–252.
- Diamond, Stephen. 2019a. Open Letter from Waterfront Toronto Board Chair, Stephen Diamond Regarding Quayside. WATERFRONToronto, June 24. <https://quaysidetoronto.ca/wp-content/uploads/2019/06/Open-Letter-from-WT-Board-Chair-on-Quayside-June-24-FINAL.pdf> [accessed October 7, 2020].
- . 2019b. Open Letter from Waterfront Toronto Board Chair. WATERFRONToronto, October 31. https://www.waterfronttoronto.ca/nbe/wcm/connect/waterfront/1e89d736-3bc1-4513-9bd3-1788dbc77352/Open-Letter-on-Threshold-Issue-Resolution-from-WT-Board-Chair-October-31-2019.pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=1e89d736-3bc1-4513-9bd3-1788dbc77352 [accessed October 7, 2020].
- Doctoroff, Daniel L. 2020. Why We’re No Longer Pursuing the Quayside Project — and What’s Next for Sidewalk Labs. Sidewalk Talk, May 7. <https://medium.com/sidewalk-talk/why-were-no-longer-pursuing-the-quayside-project-and-what-s-next-for-sidewalk-labs-9a61de3fee3a> [accessed October 7, 2020].
- Fewer, David, Sarah Crothers, Brenda McPhail, and Stephanie Perrin. 2019. The Price of Trust? An Analysis of Emerging Data Stewardship Models. https://cippic.ca/sites/default/files/file/Data_Governance_Submission_Draft_31_March_2020.pdf [accessed October 7, 2020].
- Government of Canada. 2019. Strengthening Privacy for the Digital Age. https://www.ic.gc.ca/eic/site/062.nsf/eng/h_00107.html [accessed October 7, 2020].
- Hall, Wendy, and Jerome Pesenti. 2017. Growing the Artificial Intelligence Industry in the UK. Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy, October 15. <https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk> [accessed October 7, 2020].
- Jun, Jeremiah Lau Jia, J.E. Penner, and Benjamin Wong. 2020. The Basics of Private and Public Data Trusts. *Singapore Journal of Legal Studies* 2020: 90–114.
- Khan, Lina M., and David E. Pozen. 2019. A Skeptical View of Information Fiduciaries. *Harvard Law Review* 133 (2): 497–541.
- Ministry of Government and Consumer Services. 2019. Ontario Private Sector Privacy Reform. Discussion paper. <https://www.ontariocanada.com/registry/showAttachment.do?postingId=33967&attachmentId=45716> [accessed October 7, 2020].
- O’Hara, Kieron. 2019. Data Trusts: Ethics, Architecture and Governance for Trustworthy Data Stewardship. Web Science Institute White Papers. University of South Hampton, February 23. <http://dx.doi.org/10.5258/SOTON/WSI-WP001>.
- Office of the Privacy Commissioner of Canada. 2020. PIPEDA Report of Findings #2020-004: Joint Investigation of the Cadillac Fairview Corporation Limited by the Privacy Commissioner of Canada, the Information and Privacy Commissioner of Alberta, and the Information and Privacy Commissioner for British Columbia. Office of the Privacy Commissioner of Canada, October 28. <https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2020/pipeda-2020-004/> [accessed May 17, 2021].
- Ohm, Paul. 2010. Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review* 57 (6): 1701–1177.
- Open Data Institute. 2019. Data Trusts. <https://theodi.org/wp-content/uploads/2019/04/ODI-Data-Trusts-A4-Report-web-version.pdf> [accessed October 7, 2020].
- Rinik, Christine. 2020. Data Trusts: More Data than Trust? The Perspective of the Data Subject in the Face of a Growing Problem. *International Review of Law, Computers & Technology* 34 (3): 342–363.
- Scassa, Teresa. 2020. Designing Data Governance for Data Sharing: Lessons from Sidewalk Toronto. *Technology and Regulation* 2020: 44–56.

- Sidewalk Labs. 2019. Master Innovation and Development Plan (MIDP). <https://www.sidewalktoronto.ca/midp/> [accessed October 7, 2020].
- Toronto Waterfront Revitalization Corporation, and Sidewalk Labs LLC. 2018. Plan Development Agreement. https://storage.googleapis.com/sidewalk-toronto-ca/wp-content/uploads/2019/06/13210552/Plan-Development-Agreement_July312018_Fully-Executed.pdf [accessed October 7, 2020].
- Wylie, Bianca, and Sean McDonald. 2018. What Is a Data Trust? Centre for International Governance Innovation <https://www.cigionline.org/articles/what-data-trust> [accessed October 23, 2020].

Statutes

- Island Trusts Act*. RSBC 1996, c.239.
- Personal Information Protection and Electronic Documents Act* (PIPEDA), SC 2000, c.5.
- General Data Protection Regulation* (GDPR), EU 2016/679.
- Freedom of Information and Protection of Privacy Act* (FIPPA), RSO 1990, c. F.31.
- Municipal Freedom of Information and Protection of Privacy Act* (MFIPPA), RSO 1990, c. M.56.