Special issue: Governing Data as a Resource

Technology and Regulation

Designing Data Governance for Data Sharing: Lessons from Sidewalk Toronto

Teresa Scassa*

data governance, smart cities, data trust, open data, data ownership

Teresa.Scassa@uottawa.ca

This paper considers data governance for data sharing through the lens of the data governance scheme proposed by Sidewalk Labs as part of its Master Innovation Development Plan (MIDP) for a 'smart city' development on the waterfront of Toronto, Canada. Recognizing the diverse interests in the data that might be collected in the development, the MIDP called for the creation of an Urban Data Trust (UDT) as a data governance body to address both the collection and the sharing of the novel category of 'urban data'. This paper uses the example of urban data and the UDT to illustrate some of the challenges that are central to data governance for data sharing and offers a critique that draws upon the idea of governance of a knowledge commons. The paper identifies some of the issues that led to the failure of the UDT, and extracts key lessons.

1. Introduction

Data is prized as a resource for technological innovation and economic development. Private and public sector companies, researchers, and civil society actors all seek to control and to access data, and governments increasingly seek to facilitate access. Data sharing may be between a few self-selected actors, or on a broad scale, with much in between.[1] Sharing is more complex where data sets include personal information or human behavioral data. These categories of data can significantly impact individuals and communities, raising important questions about privacy, dignity, autonomy, discrimination, expression and association.[2] The increased demand for data sharing creates a concurrent demand for data governance that can address competing claims to rights and interests in the governed data. These are not so much 'ownership' claims, although some may be framed in those terms. Rather, they are claims by groups and individuals to benefit from, or to not be harmed by, data in which they have an interest.

This paper considers data governance for data sharing through the lens of the data governance scheme proposed by Sidewalk Labs as part of its Master Innovation Development Plan (MIDP)[3] for a 'smart city' development on the waterfront of Toronto, Canada. Recognizing the diverse interests in the data that might be collected in the development, the MIDP called for the creation of an Urban Data Trust (UDT) as a data governance body to address both the collection and the sharing of the novel category of 'urban data'. Data governance bodies, whether labelled data trusts or otherwise, have generated considerable interest as a means of facilitating data sharing while accommodating different interests in data. This paper uses the example of urban data and the UDT to illustrate some of the challenges that are central to data governance for data sharing.

This paper begins with a discussion of the concept of data governance for data sharing. Recognizing both the importance of and the unique characteristics of data in a digital society, it draws upon Frischmann, Madison and Strandburg's proposal for a framework for governance that seems well suited to the smart cities context. They define the 'knowledge commons' as "the institutionalized community governance of the sharing, and, in some cases, creation, of information, science, knowledge, data, and other types of intellectual and cultural resources".[4] Their framework recognizes that the goal is not just to store data securely, but rather to share it to build new knowledge and tools in service of a common goal or in accordance with shared values. The first part of this paper explores the 'knowledge commons' as an organizing framework for data governance and links it to the

[1] A diversity of sharing arrangements can fit within the concept of a 'data trust'. See: Jack Hardinges, 'What is a Data Trust?' (*Open Data Institute*, 10 July 2018) https://theodi.org/article/what-is-a-data-trust accessed 27 April 2020.

[2] Although the widespread collection of personal data is most often associated with individual privacy, some scholars raise concerns about other privacy harms to both individuals and communities. See, e.g., Linnet Taylor, Luciano Floridi and Bart van der Sloot, 'Introduction: A New Perspective on Privacy' in Linnet Taylor, Luciano Floridi, and Bart van der Sloot (eds) *Group Privacy* (Springer 2017).

Teresa Scassa is the Canada Research Chair in Information Law and Policy at the University of Ottawa. An earlier version of this paper was presented at the workshop 'Governing Data as a Resource' organized at Tilburg University in November 2019. I am grateful to the thoughtful feedback received from many participants, and particularly to the comments of my paper's discussant, Michael Madison, and the anonymous peer reviewers. Thanks also to Pamela Robinson for her comments on an earlier draft and to my research assistant Tunca Bolca. Although I am a member of Waterfront Toronto's Digital Strategy Advisory Panel, my comments in this paper are my own and do not represent the views of that panel or of Waterfront Toronto.

[3] Sidewalk Labs, 'Master Innovation Development Plan' (*Sidewalk Labs*, June 2019) https://quaysideto.ca/sidewalk-labs-proposal-master-innovation-and-development-plan accessed 27 April 2020 [MIDP].

[4] Brett M. Frischmann, Michael J. Madison, and Katherine J. Strandburg, *Governing Knowledge Commons* (OUP 2014) 3.

particular context of Sidewalk Toronto. Section 2 introduces the Sidewalk Toronto project and considers Sidewalk Labs' initial data governance proposal and how it shifted over time and in response to public reaction. The knowledge commons framework requires a consideration of the background issues that shape the data sharing context, the resources to be shared, and the key governance elements. Each of these issues is dealt with respectively in sections 3, 4 and 5 of this paper. The conclusion identifies the issues that led to the failure of the UDT, and extracts key lessons.

2. Data Governance

The growing importance of data in the information society and economy, and the rise of data-dependent technologies such as Artificial Intelligence (AI) have created a demand for data sharing on an unprecedented scale. While data governance has always been a part of the operational reality of governments and organizations that collect and use data, data governance for data sharing is something quite different. In the first place, it is no longer about the managing of data to meet the needs and legal obligations of a specific organization. Rather, it is about governing data so as to enables its sharing with other entities or organizations to meet polycentric objectives. Data sharing can be broad and indiscriminate (as with government open data regimes), or it can be limited to one or two consenting organizations – or anything in between.

Typically, governance obligations fall on those who 'own' or 'control' data. Conventional forms of data governance – usually within a single organization (whether public or private sector) are premised on some notion of control, whether it is expressed as 'ownership' or as custody over the data. Rights to and interests in data are rooted in law, shaped by policy and practice, and negotiated in private agreements.[5] In some jurisdictions such as Canada and the UK, public sector interests in data are framed as a kind of ownership right.[6] In other jurisdictions there is no specific legal construct, other than a general custodial duty with respect to state information and data. Europe's Directive on Open Data and Public Sector Information,[7] for example, is not framed in terms of public ownership of data or information. Nevertheless, it clearly recognizes obligations of each member state to manage its information and data in the public interest.[8] Public sector right-to-know legislation establishes government as an information steward; it holds it and provides (or denies) public access in the public interest.[9] Open data policies also guide how and in what circumstances public sector data is shared with the public.[10]

A private sector organization may base its rights to control access to and use of its data through a combination of intellectual property law (copyright law[11] and the law of confidential information[12] in particular) as well as physical barriers and the laws that support them (such as trespass, technological protection measures, and criminal law).[13] Access to and use of data is governed by contracts and licences. The organization's data governance practices may also be shaped by data protection laws, as well as evolving standards regarding cybersecurity.

Law also shapes the different interests of individuals and organizations in data. Individuals have interests in their own personal data, notwithstanding any proprietary claims to the same data that might be asserted by public or private sector actors. Public and private sector data protection laws provide a framework for the recognition and exercise of individual rights and interests in personal data. These interests confer a degree of control, including rights of access, erasure (in some circumstances), and portability (in some contexts).[14] As the number and nature of the rights of individuals to their personal data expands, these rights are increasingly labeled 'ownership' rights.[15]

Normally any plan to collect data will include data governance. Where private sector companies collect data, data protection laws establish parameters for managing the data, and these must be integrated into an organization's overall data governance scheme. Data protection laws also establish the data subjects' interest in their data in the

5 Joshua B. Fisher & Louise Fortmann, 'Governing the data commons: Policy, practice, and the advancement of science' (2020) 47 *Information & Management* 237, 237.

6 In Canada, Crown copyright is provided for in s. 12 of the Copyright Act, RSC 1985 c. C-42. In the UK, Crown copyright is found in s. 163 of the Copyright, Designs and Patents Act 1988. See also: Copyright Act 1968, No. 63, 1968 (Australia), Part VII. In the United States, the Copyright Act, 17 USC §107 declares that there is no copyright in works of the federal government. However, this does not prevent state governments from asserting copyright in their works. Different states take different approaches. See: Marketa Trimble, 'U.S. State Copyright Laws: Challenge and Potential' (2017) *Scholarly Works* 1019, 84-85 https://scholars.law.unlv.edu/facpub/1019 accessed 27 April 2020.

7 Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information [2019] OJ L 172/56.

8 It also identifies the different public interests in access to and re-use of public sector information. See, e.g. Ibid., Recitals 16 and 31.

9 Ontario's Information and Privacy Commissioner emphasizes the importance of the role of the public sector as data steward in smart cities under public sector data governance legislative frameworks. (Brian Beamish, 'Open Letter to Stephen Diamond, Chairman of the Board of Directors, Waterfront Toronto' (Information and Privacy Commissioner of Ontario, 24 September 2019) https://www.ipc.on.ca/wp-content/uploads/2019/09/2019-09-24-ltr-stephen-diamond-waterfront_toronto-residewalk-proposal.pdf accessed 27 April 2020.

10 See, e.g., EU Directive on Open Data and Public Sector Information (n 7); Simpler, Faster, Better Services Act, 2019, SO 2016, c 7, Sch 56; Treasury Board Secretariat, 'Directive on Open Government' (Canada, 9 October 2014) https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=28108 accessed 27 April 2020.

11 Although copyright law places facts in the public domain, compilations of data can be protected as 'works'. Article 10(2) of the TRIPS Agreement provides: "Compilations of data or other material, whether in machine readable or other form, which by reason of the selection or arrangement of their contents constitute intellectual creations shall be protected as such. Such protection, which shall not extend to the data or material itself, shall be without prejudice to any copyright subsisting in the data or material itself." (Agreement on Trade-Related Aspects of Intellectual Property Rights, 15 April 1994, [1994] 1869 U.N.T.S. 299, 33 I.L.M. 1197.

12 E.g., art. 39 of the TRIPS Agreement, ibid, establishes criteria for the protection of confidential information. What is rewarded is not just the investment in the collection of commercially important information, but the efforts made to control that information and to maintain its confidentiality. Any 'proprietary' dimensions are rooted in physical and legal control, as opposed to 'authorship'.

13 See, e.g., Teresa Scassa, 'Data Ownership', (2018) *CIGI Papers* No. 187 https://www.cigionline.org/publications/data-ownership accessed 27 April 2020.

14 Rights of erasure and of data portability are features of the EU GDPR in respectively , articles 17 and 20. General Data Protection Regulation, persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (GDPR) [2016] OJ L119/1. The Australian Consumer Data Right also includes a data portability element. See: Treasury Laws Amendment (Consumer Data Right) Bill 2019, Bills Digest No. 68, 2018–19 https://www.aph.gov.au/Parliamentary_Business/Bills_Legislation/Bills_Search_Results/Result?bId=r6370 accessed 27 April 2020.

15 Although inaccurate, the term 'ownership' recognizes the steady expansion of these interests. For an example of such usage, see: British Academy, 'Data ownership, rights and controls: Reaching a common understanding: Discussions at a British Academy, Royal Society and techUK seminar on 3 October 2018', (British Academy 2018), 3-4 https://royalsociety.org/-/media/policy/projects/data-governance/data-ownership-rights-and-controls-October-2018.pdf accessed 27 April 2020.

hands of the organization. While, as discussed below, an organization might assert rights over its data, individual interests in personal data must also be respected.[16] The public sector is similarly also responsible for the governance of the data it collects and is bound by laws, including those relating to access to information and data protection, as well as internal policies and directives. In these cases, however, the collecting organization is either private or public sector in nature, and its data governance is shaped by existing legal frameworks.

The smart cities context implicates multiple parties with interests in data. This can include different private sector actors, one or more levels of government, and a range of other stakeholders that include urban residents individually and collectively. In some cases, the nature and/or volume of the data to be collected, the obvious demand for access to the data, the individual or group interests in the data, or the need for compromise between public and private sector partners, may call out for the creation of a new data governance framework to facilitate data sharing according to articulated values. This is particularly the case where there is a more systematic collection of greater volumes of data, along with plans for more extensive data sharing – particularly personal data or human behavioural data.[17] Data governance for data sharing in this context goes well beyond bilateral data sharing agreements and requires a novel approach. Such approaches have come to be labelled almost colloquially as 'data trusts'.

Beneath the label of 'data trust' is a concept of pooled or shared resources subject to a collective understanding around access or use. This evokes the concept of a 'knowledge commons'. This term invokes both pooled resources and collective governance,[18] reflecting a collective decision-making process.[19] Governance is of this sort incorporates collective action and approaches. Michael Madison invokes Elinor Ostrom's concept of governing the commons, where she states that "a core goal of public policy should be to facilitate the development of institutions that bring out the best in humans."[20] According to Madison, commons as governance involves communally or collectively determined principles that shape and enforce managed access to a shared resource.[21]

Frischman et al created a set of questions organized around key issues for analyzing and understanding a knowledge commons. Their framework recognizes four key elements: 1) the background environment or context in which the commons arises; 2) the attributes of the commons, including what resources are to be pooled, who the relevant community members are, and what goals and objectives it is meant to serve; 3) the governance framework for the commons including governance mechanisms, decision-makers, and relevant institutions and infrastructures. Relevant governance issues also include the applicable norms and laws and the ways in which members interact; and 4) what patterns and outcomes are relevant, including the benefits, costs and risks.[22] These questions shape the discussion below, with a particular emphasis on the background and context, the attributes of the commons, and the governance framework. Concerns about patterns and outcomes are integrated into the discussion of the background environment and context. Not only were these not well articulated in the governance proposal, its failure renders them moot. Nevertheless, benefits, costs and risks are part of the public discussion that shaped the development of the governance framework.

This knowledge commons framework identifies and organizes the issues at the core of data governance design. The relevance of these questions is evident in the case of Sidewalk Toronto. Yet as will be seen in the discussion below, there were, in a sense, two parallel processes for developing a governance framework. One came from Sidewalk Labs itself in the form of the Urban Data Trust (UDT) proposed in the MIDP. The other was a kind of public discussion occurring on many fronts that articulated different visions of a commons based in part on other known models and in part on critiques of the UDT. Not only did the public discussion shape the UDT, it likely also informed Waterfront Toronto's rejection of the proposal. In the governance vacuum created by the demise of the UDT, the knowledge commons framework remains a useful tool to shape a new approach to governance.

3. Background Environment and Context: The Origins of the Sidewalk Toronto Smart Cities Project

The knowledge commons framework identifies the background environment and context in which a commons arises as a primary consideration. In the case of Sidewalk Toronto, this context was particular and unusual and clearly played a significant role in shaping both the governance solution proposed and indeed the entire conversation around governance.

The Sidewalk Toronto smart city development originated in a Request for Proposals (RFP) issued by Waterfront Toronto for the development of a portion of port lands in the City of Toronto, Canada.[23] The Sidewalk Toronto[24] development had unique features that sharply distinguished it from other smart city projects. Most importantly, it was not led by Toronto City Council, nor was it part of a broader smart city initiative.[25] This distinguished it from cities such as Barcelona, which reflect a concerted, overall strategy driven by an elected municipal

[16] In Canada, this includes a right to access one's personal data in the hands of an organization. More extensive rights, such as rights to data portability or the right to erasure are features of the GDPR, (n 14).

[17] Diverse data governance frameworks are emerging to address data sharing in a range of contexts. See, e.g., Teresa Scassa and Merlynda Vilain, 'Governing Smart Data in the Public Interest: Lessons From Ontario's Smart Metering Entity' *CIGI Paper* #221 (CIGI 2019) https://www.cigionline.org/publications/governing-smart-data-public-interest-lessons-ontarios-smart-metering-entity accessed 27 April 2020; Open Data Institute, 'Data Trusts: Lessons from Three Pilots' (*ODI* 2019) https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT30ObnYJGibEhspr2v87jg/edit accessed 27 April 2020.

[18] Frischmann et al, (n 4).

[19] Michael J. Madison, 'Tools for Data Governance' (2020) *Technology and Regulation* 29.

[20] Elinor Ostrom, 'Beyond Markets and States: Polycentric Governance of Complex Economic Systems' (2010) 100:3 *American Economic Review* 641, 665.

[21] Madison (n 19).

[22] Frischman et al, (n 4), 20-12.

[23] The focus of this paper is on data governance. For an examination of the broader project, see: Ellen P. Goodman and Julia Powles, 'Urbanism Under Google: Lessons From Sidewalk Toronto' (2019) *Fordham LR* 457.

[24] I use the name 'Sidewalk Toronto' to refer to the proposed development. 'Quayside' is the name of the parcel of land to be developed under the agreement between Waterfront Toronto and Sidewalk Labs. Waterfront Toronto has often referred to the project as the Quayside development; Sidewalk Labs uses 'Sidewalk Toronto'.

[25] Note that Sidewalk Labs largely avoids the 'smart city' label. See, e.g., Sidewalk Labs, 'Sidewalk Labs is reimagining cities to improve quality of life', n.d. https://www.sidewalklabs.com/ accessed 27 April 2020. In the MIDP, Sidewalk Labs writes: "This effort defines urban innovation as going beyond the mere pursuit of urban efficiencies associated with the 'smart cities' movement, towards a broader set of digital, physical, and policy advances that enable government agencies, academics, civic institutions, and entrepreneurs both local and global to address large urban challenges." (*MIDP*, 'Overview' Volume 0 (n 3), 138.

government.[26] It was also different from developments in other Canadian cities such as Montreal[27] or Edmonton.[28] Unlike a smart city in which city council and city officials take the lead, the primary 'public' actor in the Sidewalk Toronto project was not a public body at all. Rather, Waterfront Toronto is a non-profit corporation created by three levels of government – federal, provincial and municipal – to oversee development of Toronto's port lands, [29] in which, due to the vagaries of geography and the Canadian constitution, all three levels of government have an interest.

According to Waterfront Toronto, its mandate is "to deliver a connected waterfront that belongs to everyone, serving as a leading example of innovation and excellence in urban design, a magnet for investment and job creation, and a source of pride and inspiration for Canadians."[30] Created in 2001 (as the Toronto Waterfront Revitalization Corporation), Waterfront Toronto had already coordinated several development projects along Toronto's waterfront. On March 17, 2017 Waterfront Toronto issued an RFP for the development of Quayside, a 12-acre parcel of port land.[31] In May 2017, Waterfront selected a proposal by Sidewalk Labs, an Alphabet company.[32] The parties entered into a Framework Agreement on October 16, 2017.[33] On July 31, 2018, Sidewalk presented its Plan Development Agreement (PDA) to Waterfront Toronto,[34] in which it set out its preliminary proposal for the Quayside project. The proposal was for a 'smart city' to be developed from the ground up, embedded with sensors as part of a vision of the "potential for technology to improve urban life and to create people-centered communities that are more livable, connected, prosperous and resilient."[35] The PDA also referred to the creation of "a destination for people, companies, start-ups and local organizations to advance solutions to the challenges facing cities . . . and make Toronto the global hub of a rising new industry focused on urban innovation."[36] After receiving feedback on the PDA, Sidewalk began working on its Master Innovation Development Plan (MIDP), which was submitted in June 2019 and made public on June 24, 2019.[37]

While the initial press coverage of the Sidewalk Toronto project showed interest in and openness to its futuristic promise,[38] the project quickly sparked a vocal public reaction. Critics raised multiple concerns, including lack of transparency,[39] experimentation on Toronto's citizenry,[40] lack of long-term viability,[41] and insufficient civic participation.[42] Local start-ups were concerned that they would be shut out of the development, and that proprietary standards might create a kind of vendor lock-in.[43] There was considerable public

[26] See, e.g., Tuba Backici, Esteve Almirall, and Jonathan Wareham, 'A Smart City Initiative: the Case of Barcelona' (2013) 4 *J Knowl Econ* 135; Josep-Ramon Ferrer, 'Barcelona's Smart City vision: an opportunity for transformation' (2017) 16 *Field Actions Science Reports* 70; Mila Gasco, 'What Makes a city smart? Lessons from Barcelona' (2016) *49th Hawaii International Conference on System Sciences* 2983. See also the discussion of the concept of an 'open smart city' in : Tracey Lauriault, Rachel Bloom, and Jean-Noé Landry, 'Open Smart Cities Guide V1.0' (Open North 2018) https://docs.google.com/document/d/1528rqTjzKWwk4s2xKuPf7ZJg-tLlRK8WcMZQbicoGTM/edit accessed 27 April 2020.

[27] See, e.g., Montreal Urban Innovation Lab https://laburbain.montreal.ca/en accessed 27 April 2020.

[28] See City of Edmonton, 'Edmonton: Smart City' n.d. https://www.edmonton.ca/city_government/initiatives_innovation/smart-cities.aspx accessed 27 April 2020.

[29] For a description and a map of the Quayside area, see: Waterfront Toronto, 'Quayside', n.d. https://www.waterfrontoronto.ca/nbe/portal/waterfront/Home/waterfronthome/projects/quayside accessed 27 April 2020.

[30] Waterfront Toronto, 'Note to Reader: Waterfront Toronto's Guide to reading the draft Master Innovation and Development Plan proposal submitted by Sidewalk Labs' (28 June 2019), 3 https://quaysideto.ca/wp-content/uploads/2019/06/Note-to-Reader_June-28-2019_Waterfront-Toronto.pdf accessed 27 April 2020.

[31] Waterfront Toronto, 'Request for Proposals: Innovation and Funding Partner for the Quayside Development Opportunity' (17 March 2017) https://quaysideto.ca/wp-content/uploads/2019/04/Waterfront-Toronto-Request-for-Proposals-March-17-2017.pdf accessed 27 April 2020 [RFP]. The RFP called for a "globally significant demonstration project that advances a new market model for climate-positive urban developments" (at 9).

[32] Sidewalk Labs, 'About Sidewalk' n.d. https://www.sidewalklabs.com/ accessed 27 April 2020. In its 2018 report, the Auditor General of Ontario criticized the very short time period allowed for responses to the RFP. See: Office of the Auditor General of Ontario, Annual Report 2018, Vol. 1, 31 http://www.auditor.on.ca/en/content/annualreports/arreports/en18/2018AR_v1_en_web.pdf accessed 27 April 2020.

[33] Waterfront Toronto, 'Framework Agreement among Toronto, Waterfront Revitalization Corp., Sidewalk Labs LLC and Sidewalk Toronto LP' (16 October 2017) https://www.waterfrontoronto.ca/nbe/wcm/connect/waterfront/035e8ad1-6ba2-46f6-8915-707176baa40f/Framework+Agreement_Executed_SUPERSEDED.pdf?MOD=AJPERES accessed 27 April 2020.

[34] Waterfront Toronto, 'Plan Development Agreement between Toronto Waterfront Revitalization Corporation and Sidewalk Labs LLC', as amended, July 31, 2018 https://waterfrontoronto.ca/nbe/wcm/connect/waterfront/73ac1c93-665b-4fb8-b19b-6bfa23c2a427/PDA+July+31+Fully+Executed+%28002%29.pdf?MOD=AJPERES accessed 27 April 2020.

[35] Plan Development Agreement (n 34) 2-3.

[36] Plan Development Agreement (n 34) 3.

[37] MIDP (n 3). After review and feedback, the MIDP was followed by a Digital Innovation Appendix, released. See: Sidewalk Labs, 'Master Innovation Development Plan: Digital Innovation Appendix' (14 November 2019) https://quaysideto.ca/wp-content/uploads/2019/11/Sidewalk-Labs-Digital-Innovation-Appendix.pdf accessed 27 April 2020.

[38] Kate McGillivray, 'Inside Quayside, the hyper-modern, tech-friendly development coming to Toronto's waterfront' (*CBC News,* 10 May 2017) https://www.cbc.ca/news/canada/toronto/quayside-waterfront-toronto-1.4108717 accessed 27 April 2020; Patrick Lynch, 'Sidewalk Labs Announces Plans to Create Model Smart City on Toronto's Waterfront,' (*Arch Daily,* 17 October 2017) https://www.archdaily.com/881824/sidewalk-labs-announces-plans-to-create-model-smart-city-on-torontos-waterfront accessed 27 April 2020; Andrea Hopkins & Alistair Sharp, 'Toronto to be home to Google parent's biggest smart city project yet' (*Financial Post,* 17 October 2017) https://business.financialpost.com/technology/google-to-be-anchor-tenant-at-toronto-innovation-hub-government-source accessed 27 April 2020; David George-Kosh, 'Alphabet's Sidewalk Labs to Create 'Smart' Neighborhood on Toronto Waterfront' (*Wall St Journal,* 17 October 2017) https://www.wsj.com/articles/alphabets-sidewalk-labs-to-create-smart-neighborhood-on-toronto-waterfront-1508266001 accessed 27 April 2020.

[39] Alanna Rizza, 'Critics call for more transparency for Sidewalk labs neighbourhood in Toronto' (*CTV News*, 8 December 2018) https://www.ctvnews.ca/sci-tech/critics-call-for-more-transparency-for-sidewalk-labs-neighbourhood-in-toronto-1.4210519 accessed 27 April 2020; Brian Barth, 'The fight against Google's smart city' (*The Washington Post*, 8 August 2019) https://www.washingtonpost.com/news/theworldpost/wp/2018/08/08/sidewalk-labs, accessed 27 April 2020; Mariana Valverde, 'The controversy over Google's futuristic plans for Toronto' (*The Conversation*, 30 January 2018) http://theconversation.com/the-controversy-over-googles-futuristic-plans-for-toronto-90611 accessed 27 April 2020 .

[40] See, e.g., Molly Sauter, 'Google's Guinea-Pig City', (*The Atlantic*, 13 February 2018) https://www.theatlantic.com/technology/archive/2018/02/googles-guinea-pig-city/552932/ accessed 27 April 2020; Star Editorial Board, 'Sidewalk Labs community can't be just a techno-experiment' (*Toronto Star*, 10 October 2018) https://www.thestar.com/opinion/editorials/2018/10/10/sidewalk-labs-community-cant-be-just-a-techno-experiment.html accessed 27 April 2020.

[41] This was one of the many concerns on a published list of questions for Sidewalk Labs. See 'Key (Mostly Unanswered) Questions Regarding Sidewalk Toronto Project' n.d. https://cfe.ryerson.ca/key-resources/guidesadvice/key-mostly-unanswered-questions-regarding-sidewalk-toronto-project accessed 27 April 2020. See also: John Lorinc, 'A Mess on the Sidewalk', (The Baffler, March 2019) https://thebaffler.com/salvos/a-mess-on-the-sidewalk-lorinc accessed 27 April 2020.

[42] See, e.g., Jathan Sadowski, 'Google wants to run cities without being elected. Don't let it' (*The Guardian*, 24 October 2017) https://www.theguardian.com/commentisfree/2017/oct/24/google-alphabet-sidewalk-labs-toronto accessed 27 April 2020.

[43] Kurtis McBride, 'Monetizing Smart Cities' (*Building*, 24 August 2018) https://building.ca/feature/monetizing-smart-cities/ accessed 27 April

outcry over issues of privacy, surveillance, and data sovereignty.[44] As the project evolved, some critics questioned the business plan underlying the deal, voicing concerns that it might be a 'real-estate grab' orchestrated by Sidewalk Labs.[45] The opposition culminated in a #BlockSidewalk movement,[46] and the Canadian Civil Liberties Association launched a lawsuit against Waterfront Toronto and the three levels of government in April of 2019, alleging that the project breached residents' constitutional rights.[47] While some have criticized opponents for their resistance to the benefits that the proposal might have for Toronto,[48] the lack of democratic/civic participation in the high-technology development was, for many, a fundamental defect.[49] Ultimately, the project involved private development with significant consequences for more than just land, creating new governance challenges. The project fell outside of traditional public sector participatory governance frameworks and outside of traditional land development paradigms.[50]

Although the original RFP called for plans to develop the Quayside district, the MIDP submitted by Sidewalk Labs distinguished between Quayside and the Innovative Design and Economic Acceleration (IDEA) District, and made proposals for both. The IDEA district included Quayside, but was much larger.[51] While Quayside represented a 4.9 hectare or 12-acre area, the IDEA district extended over 77 hectares or 190 acres. Sidewalk Labs suggested that issues of scale were behind this geographic extension. The data governance scheme proposed in the MIDP was for the larger IDEA district.[52]

Sidewalk Labs and Waterfront Toronto both sought to backfill the perceived democratic deficit with extensive public consultations leading up to the release of the MIDP and continuing afterwards.[53] The MIDP itself sought to allay many of the concerns raised by opponents of the project. Sidewalk Labs stepped back from the 'smart cities' label, recasting the project as one focusing on urban innovation.[54] Its data governance scheme (which changed shape from the PDA to the MIDP) was designed to address multiple concerns relating to the collection and sharing of data within the proposed development.

The initial proposal for Quayside framed it as a high-tech smart city development from the ground up, with a digital layer fully integrated from the design stage.[55] However, the proposal contained no clear plan for data beyond assurances that privacy would be protected through deidentification at source and the adoption of Privacy by Design (PbD) principles.[56] Data governance was an awkward issue for this project. This might have been in part because Waterfront Toronto is not the agent of any particular government and is itself not a party that would assert 'ownership' in generated data. The process by which the MIDP came about was therefore different from normal government procurement. In addition, the proposed development was not clearly either public or private in character. The project had

2020.

44 Bianca Wylie, 'Sidewalk Toronto and the Manufacturing of Consent — Thoughts Heading into Public Meeting 2 of 4' (*Medium*, 18 April 2018) https://medium.com/@biancawylie/sidewalk-toronto-and-the-manufacturing-of-consent-thoughts-heading-into-public-meeting-2-of-4-9acd289e9fa8 accessed 27 April 2020; Laura Bliss, 'How Smart Should a City Be? Toronto Is Finding Out' (*Citylab*, 7 September 2018) https://www.citylab.com/design/2018/09/how-smart-should-a-city-be-toronto-is-finding-out/569116/ accessed 27 April 2020; John Lorinc, 'A Mess on the Sidewalk' (n 43). On the issue of data localization, Sidewalk Labs was initially resistant to the concept. See, e.g., Alyssa Harvey-Dawson, 'An Update on Data Governance for Sidewalk Toronto' (Sidewalk Labs, 15 October 2018) https://www.sidewalklabs.com/blog/an-update-on-data-governance-for-sidewalk-toronto/ accessed 27 April 2020.

45 See, e.g., David Skok, 'Cracks in the Sidewalk' (*Macleans*, 15 February 2019) https://www.macleans.ca/opinion/cracks-in-the-sidewalk/ accessed 27 April 2020; Bianca Wylie, 'Sidewalk Toronto: Here's the Business Model Framework' (*Medium*, 7 June 2018) https://medium.com/@biancawylie/sidewalk-toronto-waterfront-toronto-digital-strategy-advisory-panel-meeting-1-before-6a158971eb65 accessed 27 April 2020.

46 #BlockSidewalk, n.d. https://www.blocksidewalk.ca/ accessed 27 April 2020.

47 Canadian Civil Liberties Association, 'CCLA Commences Proceedings Against Waterfront Toronto' (16 April 2019) https://ccla.org/ccla-commences-proceedings-waterfront-toronto/ accessed 27 April 2020.

48 Stephanie Marotta, 'Business leaders push for Sidewalk Labs smart-city development to be built on Toronto's waterfront' (*Globe and Mail*, 4 July 2019) https://www.theglobeandmail.com/business/article-business-leaders-push-for-sidewalk-labs-smart-city-development-to-be/ accessed 27 April 2020. This article references an open letter published by local business leaders. See Toronto Region Board of Trade, 'Open Letter from Civic Leaders', 4 July 2019 https://www.bot.com/Portals/0/NewsDocuments/742019Civic%20Leaders%20Open%20Letter%20final.pdf accessed 27 April 2020.

49 Bianca Wylie, 'Democracy or Sidewalk Toronto. You Can Have One But You Can't Have Both' (*Medium*, 14 May 2019) https://medium.com/@biancawylie/democracy-or-sidewalk-toronto-you-can-have-one-but-you-cant-have-both-a40e4d1d8daa accessed 27 April 2020; Michael Oliviera, 'Critics decry lack of 'democratic participation' over Sidewalk Labs' proposed neighbourhood' (*Toronto Star*, 2 May 2018) https://www.thestar.com/news/gta/2018/05/02/critics-decry-lack-of-democratic-participation-over-sidewalk-labs-proposed-neighbourhood.html accessed 27 April 2020. See also Goodman & Powles (n 23).

50 An interesting analogy might be made with projects that have significant environmental impacts. These projects necessarily combine economic and development priorities with complex public interest and environmental concerns. In the environmental regulation context, there are complex frameworks for the assessment and approval of such projects. It is also worth noting that the concept of 'social licence' has its roots in the environmental context. See: Kristen van de Biezenbos, 'The Rebirth of Social Licence' (2019) 14 *McGill J. Sust. Dev. L.* 149.

51 See the map of the areas under discussion in: Swerhun, Inc., Waterfront Toronto's Public Consultation on the draft MIDP: Round One Feedback Report (Toronto, 19 September 2019) 5, https://quaysideto.ca/wp-content/uploads/2019/09/Round-One-Public-Consultation-Feedback-Report-September-19-2019.pdf accessed 27 April 2020.

52 Waterfront Toronto, Note to Reader (n 30) 5. The expansion of the area to form part of the proposal was not approved by Waterfront Toronto, and the project has since been scaled back: Stephen Diamond, 'Open Letter from Waterfront Toronto Board Chair'(31 October 2019) https://waterfrontoronto.ca/nbe/wcm/connect/waterfront/waterfront_content_library/waterfront+home/news+room/news+archive/news/2019/october/open+letter+-from+waterfront+toronto+board+chair+-+october+31%2C+2019 accessed 27 April 2020.

53 See, e.g., Waterfront Toronto, 'Quayside: Participate in a Public Consultation' n.d. https://quaysideto.ca/get-involved/public-consultation/ accessed 27 April 2020. Sidewalk Labs' public outreach is described in the MIDP (n 3), Volume 0, 67. A summary of Waterfront Toronto's public consultation, carried out after the release of the MIDP, was published in September 2019. See: Swerhun, Inc (n 52) 5.

54 In the introductory Volume to its MIDP, Sidewalk Labs writes: "This effort turned Sidewalk Labs' initial ideas, as expressed in the RFP response, into a development plan with the potential to serve as a demonstration for an inclusive community that puts urban innovation to work for better quality of life." (See MIDP (n 3), Volume 0, 86).

55 PDA (n 34) 49. See also Dan Doctoroff, 'Reimagining cities from the Internet up' (*Medium*, 30 November 2016) https://medium.com/sidewalk-talk/reimagining-cities-from-the-internet-up-5923d6be63ba accessed 28 April 2020.

56 Brian Jackson, 'Sidewalk Toronto commits to Privacy by Design principles amid citizen concerns' (*IT World Canada*, 4 May 2018) https://www.itworldcanada.com/article/sidewalk-toronto-commits-to-privacy-by-design-principles-amid-citizen-concerns/404887 accessed 28 April 2020; Ann Cavoukian, 'De-identifying data at the source is the only way Sidewalk can work' (*Toronto Life*, 4 September 2019) https://torontolife.com/city/de-identifying-data-at-the-source-is-the-only-way-sidewalk-can-work/ accessed 28 April 2020. Privacy by design principles focus on embedding privacy into the design of technology. See: Ann Cavoukian, 'Privacy by design: The 7 Foundational Principles' (Information and Privacy Commissioner of Ontario, January 2011) https://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf accessed 28 April 2020.

definite public dimensions: it involved publicly owned lands, was originally labelled a 'smart city', and it implicated traditional, municipal services. At the same time, it was also a real estate development and a technology innovation hub.[57] The knowledge commons framework demands consideration of the background and cultural context for the knowledge commons. In the case of Sidewalk Toronto, the relationship between a private sector company and a non-profit corporation around the digital integration of public and private sectors within a real estate development/technology innovation lab created a particular challenge for data governance.

The deadline to finalize an agreement based on the MIDP was extended from September 2019 to March 31, 2020,[58] with a possibility for the parties to terminate the PDA by October 31, 2019 if no agreement could be reached on key issues.[59] The project survived the October 31, 2019 cut-off date after Sidewalk Labs agreed to a number of conditions set by Waterfront Toronto. These included abandoning the UDT and avoiding the novel category of 'urban data' both of which are the focus of this paper.[60] Sidewalk Labs subsequently produced a lengthy Digital Innovation Appendix[61], which provided greater detail about its plans and a more cautious approach to data governance, which recognizes that Waterfront Toronto must play a central role.[62] The project came to an abrupt end on May 7, 2020. In a statement released by Sidewalk Labs' Dan Doctorow, the "unprecedented economic uncertainty [that] has set in around the world and in the Toronto real estate market" was cited as the reason for its termination.[63]

In spite of the demise of the project, the UDT and 'urban data' remain of interest and importance both to understand their origin and concept as a novel form of data governance for data sharing, as well as the reasons for their rejection.

4. The Emergence of Key Governance Issues

The second category of considerations in the knowledge commons framework relate to key attributes of the emerging commons, including the nature of the resources to be governed, the members of the relevant governance community, and the goals and objectives of the commons. In part because of the way in which this project evolved, there was considerable pushback around these issues once the plans for the project became public. As a result, at the same time as a governance framework was being developed, there was a parallel set of conversations that raised particular concerns and preoccupations around many of the core attribute issues. This section considers issues that emerged in public reactions and how the shaped the development of what ultimately became the UDT.

Elements of public pushback can be organized into four broad challenges that Sidewalk Labs subsequently sought to address in the data governance scheme that it proposed[64] and later refined in the MIDP.[65] As a result, these four publicly expressed data-related concerns played an important role in shaping the evolution of the governance scheme in the MIDP. The sheer breadth of the concerns made governance increasingly complex, perhaps overburdening the proposed framework.

The first set of issues related to **data sharing and access**. The initial announcement of the project raised concerns among local technology developers who felt that it might exclude them from opportunities to participate in the development of smart city technology in Toronto, with a large US corporation instead being invited to both shape and occupy the market.[66] Although Sidewalk Labs talked of making data from the project open, the extent of this commitment was unclear.[67] Developers' inclusion issues extended beyond data;[68] nevertheless, there was a desire that smart city data be made available in real-time and under open licences so that developers could use it to generate innovative and competing applications for the city.[69] The data sharing and access concerns were ones that suggested a need for some form of knowledge commons.

Developers also wanted to be able to **participate in the data collection** that would take place within the development zone. In other words, they resisted a vision in which Sidewalk Labs had a monopoly on the applications that would be used to collect smart city data. Sidewalk Labs responded with assurances that it would not monopolize innovation within the district. However, permitting more developers to innovate also meant that there would be new data governance challenges. While Sidewalk Labs could make commitments about data sharing, deidentification, or privacy by design with respect to its

57 See Steve McLean, 'Sidewalk Labs' Sirefman updates Toronto development plans' (*Real Estate News Exchange*, 18 September 2019) https://renx.ca/sidewalk-labs-sirefman-toronto-waterfront-development/ accessed 28 April 2020; James McLeod, 'Did Sidewalk Labs overstep with their masterplan? It certainly raised concerns at Waterfront Toronto' (*Financial Post*, 24 June 2019) https://business.financialpost.com/technology/sidewalk-labs-long-awaited-smart-city-masterplan-raises-concerns-at-waterfront-toronto accessed 28 April 2020. In the MIDP (n 3), Vol 1, 17, Sidewalk Labs describes its 'Innovative Design and Economic Acceleration(IDEA) District that represents an innovative new development model for how the private sector can support the public sector in tackling the toughest growth challenges.'

58 Waterfront Toronto and Sidewalk Labs, Amending Agreement (31 July 31 2019), 1 https://quaysideto.ca/wp-content/uploads/2019/04/Plan-Development-Agreement-July-31-2018-and-Amendment-July-31-2019.pdf accessed 28 April 2020. This deadline was subsequently extended to take into account the COVID-19 crisis.

59 Ibid.

60 Diamond (n 55).

61 Sidewalk Labs, (n 38).

62 Ibid. These timelines have been further extended as a result of the COVID-19 pandemic.

63 Daniel L. Doctorow, "Why we're no longer pursuing the Quayside project — and what's next for Sidewalk Labs" (*Medium*, 7 May 2020) https://medium.com/sidewalk-talk/why-were-no-longer-pursuing-the-quayside-project-and-what-s-next-for-sidewalk-labs-9a61de3fee3a accessed 8 May 2020.

64 Sidewalk Labs, 'Digital Governance Proposals for DSAP Consultation' (October 2018) https://waterfrontoronto.ca/nbe/wcm/connect/waterfront/41979265-8044-442a-9351-e28ef6c76d70/18.10.15_SWT_Draft+Proposals+Regarding+Data+Use+and+Governance.pdf?MOD=AJPERES accessed 28 April 2020.

65 MIDP (n 3) Vol II, Ch 5.

66 See, e.g., Aeman Ansari, 'Toronto doesn't need Google to build a smart city, says open data expert' (*betakit*, 20 November 2017) https://betakit.com/toronto-doesnt-need-google-to-build-a-smart-city-says-open-data-expert/ accessed 28 April 2020; Bill Bean, 'The world is watching as data drives Toronto's Smart City experiment' (*Communitech News*, October 30, 2017) http://news.communitech.ca/the-world-is-watching-as-data-drives-torontos-smart-city-experiment/ accessed 28 April 2020; Jim Balsillie, 'Sidewalk Toronto has only one beneficiary, and it is not Toronto' (*Globe and Mail*, 15 October 2018) https://www.theglobeandmail.com/opinion/article-sidewalk-toronto-is-not-a-smart-city, accessed 28 April 2020.

67 See, e.g., Bianca Wylie, 'Civic Tech: On Google, Sidewalk Labs, and Smart Cities' (*Torontoist*, 24 October 2017) https://torontoist.com/2017/10/civic-tech-google-sidewalk-labs-smart-cities/ accessed 28 April 2020.

68 Some even expressed the concern that discussions around data distracted from issues of ownership/control of the underlying source code. See: Terry Pender, 'Miovision CEO sees great value in Sidewalk Labs data' (*The Record.com*, 3 November 2018) https://www.therecord.com/news-story/9004728-miovision-ceo-sees-great-value-in-sidewalk-labs-data accessed 28 April 2020.

69 Donovan Vincent, 'Who will reap the benefits of Quayside's smart city data?' (*Toronto Star*, 16 December 2018) https://www.thestar.com/news/gta/2018/12/16/who-will-reap-the-benefits-of-quaysides-smart-city-data.html accessed 28 April 2020.

own technologies, it could not do the same for other actors.[70] Instead, it decided to make compliance with the data governance scheme a precondition for participation in the data ecosystem that was being developed.[71] Those seeking to collect data within the IDEA District, or those seeking to use certain types of 'urban data' that were not otherwise available as open data, would have to request permission and comply with requirements established as part of the data governance framework. Not only did this undermine the potential for the design of the kind of consensual data governance framework required for a knowledge commons, the potential scale and cost of managing this more complex data sharing framework, would also have implications for 'openness'. In the MIDP, Sidewalk Labs indicated that there might be fees for approvals of plans to collect or use data submitted to the Urban Data Trust.[72]

A third wave of opposition related to data came from those who were concerned that the ubiquitous collection of data within the smart city posed a **risk to privacy and other values**. Privacy issues had already been anticipated by Sidewalk Labs, which had retained former Ontario Information and Privacy Commissioner Ann Cavoukian as a consultant. Based on principles of Privacy by Design (PbD)[73] Sidewalk Labs had promised that all data it collected would be de-identified at source.[74] However, critics found this unsatisfactory for two main reasons. The first was a growing lack of confidence in deidentification as a means of protecting privacy.[75] In a context in which vast quantities of different types of data are collected and analyzed using big data analytics and AI, reidentification risks are high.[76] A second concern was that even deidentified human behavioural data posed risks of harm both to individuals and to communities. These harms could flow from the use of the data to profile individuals or communities/groups in ways that might impact their access to resources or benefits, or that might incorporate or contribute to bias and oppression.[77] If PbD and deidentification were not complete solutions to the problem, then something more was needed. That something would have to include a mechanism to ensure that the data collected would be used in an appropriate, ethical and responsible manner. This is suggestive of the need for some form of framework for governing the 'knowledge commons'. The UDT proposed in the MIDP was therefore designed to oversee the collection and use of data, under a Responsible Data Use Agreement (RDUA)[78] similar to a privacy impact assessment.

A fourth issue around **data localization** arose from the considerable opposition to the idea that data collected in the smart city environment might end up stored on servers located outside Canada. On one level this was a privacy issue – Canadians have long been wary about the impact of the U.S. PATRIOT Act[79] on data about Canadians stored in the United States.[80] On another level, it is a data sovereignty issue.[81] Because the data was collected within and about a Canadian city, many saw it has having a public quality and that it should there-

[70] See Gabrielle Cannon, 'City of Surveillance: Privacy Expert Quits Toronto's Smart City Project' (*The Guardian*, 23 October 2018) https://www.theguardian.com/world/2018/oct/23/toronto-smart-city-surveillance-ann-cavoukian-resigns-privacy accessed 28 April 2020; John Buntin, 'Technopolis: Google's Sister Company Wants to Build the City of the Future on Toronto's Waterfront. Should a private tech giant be designing smart cities?' (*Governing*, July 2019) https://www.governing.com/topics/urban/gov-google-toronto.html accessed 28 April 2020.

[71] The MIDP (n 3), is clear that meeting the requirements of the Responsible Data Use Framework (RDUF) is independent of meeting all legal obligations. In other words, developers would not only have to meet the requirements of applicable laws, they would also have to meet what might be additional requirements imposed by the UDT. Adding another layer of compliance– and one for which fees might be charged -- would increase the burden for participation of SMEs.

[72] MIDP (n 3) Vol II, Ch 5, 422 and 434-435.

[73] Cavoukian, 'Privacy by Design' (n 58).

[74] Cavoukian, 'De-identifying data' (n 58).

[75] Concerns over reidentification risk have existed for some time (see, e.g., Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 *UCLA LR* 1701. These are exacerbated with the advance of technology. A recent article found that the reidentification risk was so high even for anonymized medical data that anonymization techniques in use today were unlikely to meet the rigorous norms of the GDPR. See: Luc Rocher, Julien M. Hendrickx & Yves-Alexandre de Montjoye, 'Estimating the success of re-identifications in incomplete datasets using generative models' [2019] 10 *Nature Communications* Article #3069 https://www.nature.com/articles/s41467-019-10933-3 accessed 28 April 2020. By contrast, Yakowitz argues that re-identification risks are exaggerated. See: Jane Yakowitz, 'Tragedy of the Data Commons' (2011) 25 *Harv. J L & Tech* 1.

[76] See Rocher et al, ibid.

[77] Concerns over the adverse impacts of data profiling on individuals and groups are longstanding. See, e.g., David Lyon, *Surveillance as Social Sorting* (Routledge 2002); Oscar H. Gandy, Jr., *Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage* (Routledge 2009); Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (MacMillan 2018).

[78] MIDP (n 3) Vol II, Ch 5, 424-440.

[79] Uniting (and) Strengthening America (by) Providing Appropriate Tools Required (to) Intercept (and) Obstruct Terrorism Act of 2001, Pub L 107-56 [The U.S.A. PATRIOT Act].

[80] For example, the governments of British Columba and Nova Scotia each passed laws that prohibited the storage of certain public data outside of Canada. See: Personal Information International Disclosure Protection Act, SNS 2006, c 3; Freedom of Information and Protection of Privacy Act, RSBC 1996, c 165, s. 30.1. See also: David Loukidelis, 'Privacy and the USA PATRIOT Act: Implications for British Columbia Public Sector Outsourcing' (Office of the Information and Privacy Commissioner of British Columbia, October 2004) https://www.oipc.bc.ca/special-reports/1271 accessed 28 April 2020.

[81] Banks characterizes it as a situation where "vast troves of data in a public-private partnership would be exfiltrated from Canada." He asks, "Once the data is outside of Canada, could Canadian governmental bodies ever reclaim control of that data should future voters decide that this is appropriate for security or other reasons?" (Timothy Banks, 'Will Sidewalk Labs' civic data trust hush critics of Waterfront Toronto?', (*IT and Data Governance*, 23 October 2018) https://timothy-banks.com/2018/10/23/will-sidewalk-labs-civic-data-trust-hush-critics-of-waterfront-toronto/ accessed 28 April 2020. Sean McDonald notes: "Framing data localization around the Canadian Government's enforcement of privacy law narrows the potential benefits of localization, and ignores the threats emanating from internationalizing the processing and storage of public data." (Affidavit of Sean McDonald in Canadian Civil Liberties Assn and Lester Brown v. Toronto Waterfront Revitalization Corporation, et al, Court File No. 211/19, 16 https://ccla.org/cclanewsite/wp-content/uploads/2019/06/Affidavit-of-Sean-McDonald-2019-05-28.pdf accessed 28 April 2020. The term "data sovereignty" is sometimes confused with other concepts such as data residency or data localization. Data localization typically involves legal requirements to store data within a specified jurisdiction. Data residency involves ensuring that enough of a company's data processing activities are 'located' in a legal sense within a country's borders in order to take advantage of certain beneficial laws or policies. Data sovereignty, in its narrowest sense refers to data being subject to the laws of a particular jurisdiction. However, data sovereignty can have a broader meaning, as it does in the context of the Indigenous Data Sovereignty movement. In that context, data sovereignty involves not only claims to self-governance with respect to the storage and management of data about the self-governing community. See, e.g., Tahu Kukutai and John Taylor, eds., *Indigenous Data Sovereignty: Toward an Agenda*, (ANU Press 2016) https://www.oapen.org/download?type=document&docid=624262#page=25 accessed 28 April 2020. Note that the term "data sovereignty" is now also used in relation to person control over personal data. See, e.g., the statement that "sovereign data subjects are those who are in a position to articulate and enforce claims to power about their data." Patrik Hummel et al, 'Sovereignty and Data Sharing' (2018) *ITU Journal: ICT Discoveries*, Special Issue No. 2, 2 https://www.itu.int/en/journal/002/Documents/ITU2018-11.pdf accessed 28 April 2020.

fore be located in Canada.[82] Although Sidewalk Labs initially resisted data localization arguments,[83] by the time the MIDP was published, this commitment had softened somewhat – storage in Canada would take place if adequate facilities existed.[84] The discussion over data localization suggests that the proposed UDT was meant to house the data it governed, rather than simply managing access to the data stored on the servers of the actors that generate it – although this was not entirely clear.[85]

Because this was not a public or city-led project, public concerns could only be raised after the announcement of the project. This led to the development of a governance framework in rather unique circumstances that ultimately proved problematic. Not only did the timing and context prevent the collaborative development of the data governance framework by all stakeholders, the project went ahead before the question of who was to have custody or control over what data was resolved. It is fair to say that many considered that smart city data would, by default, be municipal data under the custody and control of the City of Toronto – at least so far as the data was collected in relation to the infrastructure, streets, and other public spaces of the development.[86] This view was evidently not shared by Sidewalk Labs, although Sidewalk Labs remained cagey on the issue.[87] The issue is important. Contributors of data to a knowledge commons are stakeholders entitled to participate in the shaping of the governance framework. By moving ahead without addressing who was contributing what data to the commons, there could be no consensual governance model.

Sidewalk Labs ultimately proposed an independent data governance body to oversee its data-sharing framework. In doing so, it also attempted to hive off a category of data suitable for governance in this way. Rather than identifying particular data sets, whether controlled by public or private sector actors, that should be pooled and governed collectively in the public interest, it chose to create a whole new category of data – "urban data". Any data falling within the definition of 'urban data' was subject to governance by the Urban Data Trust.

5. Attributes: Urban Data

The second category of questions in the knowledge governance framework asks what resources are to be pooled, who the relevant stakeholders are, and what the goals and objectives are. The category of 'urban data' in the MIDP was, in many ways designed to answer these sorts of questions. As will be seen below, it defined a category of data for governance (urban data), characterized it as a kind of communally shared resource, and identified a fairly general concept of public interest. But, as will be seen below, this category of data was inherently problematic, creating fundamental problems for the data governance scheme. In this sense, the category also interacts with the next set of questions over knowledge commons governance, as the novel category of 'urban data' made it difficult to identify how existing legal frameworks would apply. 'Urban data' was defined as either unowned or communally owned. The data was conceived of as existing independently of its collectors, who would have to seek permission and follow rules regarding its collection. The category of urban data therefore defined the commons in terms of data in a geographic context, rather than data sets collectively pooled by stakeholders to serve common ends.

One reason why the UDT might have been built around 'urban data' could be to avoid the legal barriers to the contribution of public sector data to a communal governance regime. Under the laws of Ontario at the time of the MIDP,[88] the management of data collected by a public sector entity could not simply be delegated to a third party with its own governance rules. The public body was legally required to manage that data according to public sector laws and policies.[89] There was therefore a jarring and unresolved relationship between public ownership as represented by the public sector, and the notion of 'public' or 'communal' ownership of urban data in the UDT. These challenges were not insurmountable, but they might have required some legislative change.

The MIDP defined "urban data" as "information gathered in the city's physical environment, including the public realm, publicly accessible spaces, and even some private buildings".[90] The category 'urban data' was largely based on geography and concepts of public versus private space. Urban data could be personal or non-personal data, and could

[82] Lauriault et al (n 80) 24, state that "Data residency is a critically important consideration for Open Smart Cities because many firms that provide cloud computing for smart cities (Google, Microsoft, etc.) store their data in servers outside of Canada."

[83] See, e.g., Alyssa Harvey-Dawson, 'An Update on Data Governance for Sidewalk Toronto' (Sidewalk Labs, 15 October 2018) https://www.sidewalklabs.com/blog/an-update-on-data-governance-for-sidewalk-toronto/ accessed 28 April 2020.

[84] Specifically, Sidewalk Labs committed "to using its best efforts at data localization, as long as there are Canadian-based providers who offer appropriate levels of security, redundancy, and reliability.", MIDP (n 3), Vol II, Ch 5, 460. In an Open Letter dated October 31, 2019, the Chair of Waterfront Toronto confirmed that the parties had agreed that all personal information would be stored in Canada (Diamond (n 55)) For a critique of Sidewalk Labs' initial approach to data localization, see Affidavit of Sean McDonald (n 82).

[85] For example, when discussing access to data collected under the supervision of the Urban Data Trust, the MIDP states: "Facilitating access could be accomplished in a variety of ways, from having the Urban Data Trust actually hold the data to having it set rules that require collectors to publish de-identified, aggregate, or non-personal data in real time." MIDP (n 3), Vol II, Ch 5, 434.

[86] See, e.g., the letter of Ontario's Information and Privacy Commissioner that criticizes how the MIDP negates the role of the public sector in governing Sidewalk Toronto data (Beamish (n 9)).

[87] The MIDP makes reference to the need to comply with "all applicable laws" (See, e.g., MIDP (n 3) Vol II, Ch 5, 421.) and identifies both public sector (FIPPA and MFIPPA) and private sector (PIPEDA), data protection laws without specifying which would apply in what contexts (MIDP (n 3) Vol II, Ch 5, 421.).

[88] Ontario has since amended its Freedom of Information and Protection of Privacy Act, RSO 1990, c F.31, to allow for the creation of entities outside government that can engage in the governance of data from multiple sources.

[89] For municipal governments in Ontario, this includes the Municipal Freedom of Information and Protection of Privacy Act, RSO 1990, c M.56. As the Information and Privacy Commissioner for Ontario notes, it would have been possible for the UDT to take a position on privacy different from that of the provincial regulator. See Beamish (n 9), 6. It is perhaps no surprise that the federal government is also contemplating legislative change to facilitate collective data governance in a manner consistent with data protection obligations. See ISED, 'Strengthening Privacy for the Digital Age: Proposals to modernize the Personal Information Protection and Electronic Documents Act' 21 May 2019, https://www.ic.gc.ca/eic/site/062.nsf/eng/h_00107.html accessed on 28 April 2020.

[90] MIDP (n 3), Vol II, Ch 5, 377. 'Publicly accessible spaces' is a complex category. It appears to mean that owners or lessors of publicly accessible private properties– such as retail spaces (MIDP (n 3), Vol II, Ch 5, 426) – would, to the extent that they collect data in these spaces, be collecting 'urban data' and would therefore be subject to the RDUA and the UDT. In the MIDP, Sidewalk Labs uses the example of a parking garage lessor who would need to go through the RDUA process in order to install security cameras in its garage (MIDP (n 3), Vol II, Ch 5, 439-440.) The 'public realm' includes public spaces such as streets or parks. It apparently also includes atmospheric or environmental data. MIDP (n 3), Vol II, Ch. 5, 379, 417

include aggregate or de-identified data.[91] The definition neatly avoided the traditional dichotomy of public and private sector data; the identity of the party collecting the data was irrelevant to its characterization as 'urban data'. The creation of a new category evaded both ownership and control issues, as well as the collaborative approach to governance that different 'ownership' interests would entail. Yet the recognition and reconciliation of diverse interests is both an important process and an outcome of commons governance.

As defined in the MIDP, Sidewalk Labs' 'urban data' has the following characteristics:

1. It is defined based upon where it is collected (i.e. location is a key element in the definition of urban data);[92]
2. The "where" is linked to some concept of shared or communal space;
3. Shared or communal space can cut across the boundaries of publicly and privately-owned spaces;[93]
4. Urban data may include personal information and/or human behavioural data, as well as other types of non-personal data;[94]
5. Urban data is not defined by who is collecting it (i.e. it can be collected by public or private sector actors and possibly even by individuals).[95]

Geography or location was therefore a core component of the definition.

'Urban data' was defined both in terms of what it was and what it was not. For example, 'urban data' is distinct from what Sidewalk Labs labeled as "transaction data". Transaction data was data relating to any specific transactions carried out by individuals with the providers of particular services (such as ride-sharing, utilities, etc.).[96] The distinction between urban data and transaction data was explained by Sidewalk Labs' Alyssa Harvey-Dawson:

> For clarity, we call the original information collected in a physical place in the city "urban data." Urban data is different from data created when individuals agree to provide information through a website, mobile phone, or paper document. It presents unique challenges, including that it could reasonably be considered a public asset, and that it raises potential concerns around surveillance and privacy.[97]

In the MIDP, transaction data was also explained as not fitting within the category of urban data because it would be "unworkable given the lack of a relationship between this kind of data and a specific geography."[98] Yet 'transaction data' would not always be easy to separate from information collected in a physical space. For example, contracting for municipal water services will generate transaction data such as the amounts billed to a particular customer. However, it is unclear whether data about the volume, frequency and timing of water consumption (on which billing is based) is solely transaction data or also 'urban data', since it is linked to a particular geographic location (the point of consumption). Perhaps the answer is that some data would be transaction data when linked to a particular individual, but could become urban data in aggregate or anonymized form. As another example, the MIDP distinguished between data from sensors such as cameras on ride sharing vehicles (urban data for which permission to collect in the IDEA district is required) and consumer trip and payment data, which would be transaction data.[99] Yet arguably, data about the movement of a person from point A to point B within the IDEA District (which is data relevant to the transaction) has links to physical space and could be construed as urban data, particularly if it were useful data for understanding traffic patterns or transit demands. These questions about where transaction data ended and urban data began illustrates how challenging the definition of a novel category of data can be.

A major reason why transaction data was separated from urban data was because it is seen as specific to a contractual relationship between an individual and a service provider, and would be governed by terms of service and a separate privacy policy. In other words, this data was not collectively owned because it was seen as proprietary to the party that collected it from an individual under the terms of a contract. Sean McDonald criticized this distinction between transaction data and urban data, stating that the result is that "the more sensitive the data the more proprietary it would be."[100] Yet this seems precisely part of the rationale. For Sidewalk Labs, urban data was suitable for collective governance because it was 'owned' by no one. The relationship between the individual and the provider both makes the data proprietary and enhances its sensitivity. By contrast, urban data involves no specific relationships. Sidewalk Labs' insistence on geography as a core characteristic of urban data nevertheless created a tension with transaction data because the two categories – urban data and transaction data – depended on different characteristics that were not mutually exclusive.[101] Urban data relied upon collection in shared geographical spaces, while transaction data was defined in terms of specific relations between an individual and an organization. The fact that specific relationships can arise with respect to data that – in aggregate – can provide information about shared public space creates conceptual problems. These are only augmented by the ambiguity around the notion of public versus private spaces. It raises the question of why aggregate transaction data with the appropriate geographical dimensions is not also communally owned urban data.

The definition of urban data is also interesting because it relied upon concepts of 'public' and 'private' tied to geography and in particular to concepts of public and private spaces defined not necessarily in terms of land ownership but in terms of access and usage. This

[91] MIDP (n 3) Vol II, Ch 5, 417.
[92] "The term 'urban data' nods to the fact that it is collected in a physical space in the city and may be associated with practical challenges in obtaining meaningful consent." MIDP (n 3) Vol II, Ch 5, 416.
[93] In the MIDP (n 3) this seems to include privately owned or controlled spaces with public dimensions, such as retail stores, the lobbies of apartment buildings, or public spaces within publicly owned buildings.
[94] For example, in the MIDP (n 3) Vol II, Ch 5, 416, it states that urban data "includes both personal information and information that is not connected to a particular individual." It goes on to say that "Urban data would be broader than the definition of personal information and includes personal, non-personal, aggregate, or de-identified data. . .".
[95] It is not clear whether "There is no obvious means for individuals to consent to its collection" should be a sixth factor in this list, or whether this statement is simply a conclusion that can be drawn from the listed features of urban data. In other words, it is not clear if it is an 'and', or if the problem of consent is considered inherent to data within this category.
[96] MIDP (n 3), Vol II, Ch 5, 416: "urban data would be distinct from more traditional forms of data, termed here "transaction data", in which individuals affirmatively – albeit with varying levels of understanding – provide information about themselves through websites, mobile phones, or paper documents."
[97] Harvey-Dawson (n 45).
[98] MIDP (n 3) Vol II, Ch 5, 427.
[99] MIDP (n 3) Vol II, Ch 5, 427.
[100] Sean McDonald, 'Toronto, Civic Data, and Trust', (*Medium*, 17 October 2018) https://medium.com/@McDapper/toronto-civic-data-and-trust-ee-7ab928fb68 accessed 28 April 2020.
[101] This is also noted in the Ontario Information and Privacy Commissioner's open letter to the Chair of Waterfront Toronto; Beamish (n 9).

jarred with established understandings of data ownership that turn on in who collects or controls the data.[102] Data could be 'urban data' regardless of whether it was collected by public or private sector actors

The linking of urban data to location was probably at least partly driven by concerns over the collection of human behavioural data. At one point in the MIDP, Sidewalk Labs noted that the location elements "may be associated with practical challenges in obtaining meaningful consent."[103] In other words, the data governance scheme was designed to address the privacy problem of the requirement of consent for collection of personal information in a context in which consent would be impractical to request or obtain – such as urban public spaces. Yet since technology might evolve to enable consent in a broader range of contexts, this added another layer of uncertainty about what would constitute 'urban data'.

The consent requirement for the collection of personal data is different in Canada depending on whether the collector is a public or private sector actor. Consent is not required for personal data collection by public sector actors, although notice is.[104] This recognizes the imbalance of power between governments and citizens as making true consent impossible.[105] Instead, data collection by government is legitimized by democratic processes that enable the government's action and the public policy considerations that motivate the collection. Where the collector is a private sector actor, consent is required.[106] The UDT was intended to provide a substitute process to legitimize collection without consent in public spaces by private sector actors.[107] It did so by establishing an independent governance framework that would set the rules for both collection and for subsequent uses of this data. Yet this shifted the role of the UDT from data steward to a kind of data protection authority or even a mini-municipal government. For example, if Toronto's municipal government decided to collect a certain type of data from light standards or from other municipal infrastructure throughout the city, it would have to apply to the UDT for permission to deploy these sensors in the IDEA district – and such permission could be refused, or conditions could be attached.[108] This created the possibility that the UDT could deny permission to the population's democratically elected municipal government to implement a city-wide policy decision.[109]

The problems were not just with new governance for public sector data. The MIDP offered an example of a parking garage operator in the development area who decides to install security cameras. Although use by patrons of the garage is consent-based and transactional, Sidewalk Labs considered the camera data to be 'urban data' subject to the governance regime. Thus, the garage operator, who would already be subject to private sector data protection legislation, would have to go through the RDUA process. It seems problematic to suggest that security camera footage should be contemplated as shareable through the UDT, even in deidentified form. There is no compelling case for public or communal 'ownership' of such data. Any governance process beyond data protection law seems unnecessary. Other problematic "publicly accessible spaces" might include the lobbies of apartment buildings or condominiums, retail stores, shopping malls, or restaurants. In all of these cases, there are already data protection laws that would govern collection of personal information, and in many instances, collection would be for fairly specific purposes such as security. In most cases, governance through privacy legislation would suffice to place strict limits on what could be collected, how it might be used, how notice would have to be provided, and how long the data could be retained. It is unclear what added value would be provided by a further layer of data governance. Adding such data to a data governance regime for data sharing would only raise additional privacy and ethical concerns.

By making geography (particularly 'public space') the primary characteristic of 'urban data', the definition also became dangerously over-inclusive. Sidewalk Labs provided at least two examples of urban data collection in which the problems of over-inclusivity are evident. The first involved the use of an app to collect non-personal data about park usage by a civil society group.[110] The data collection was an automated version of what might otherwise be recorded by volunteers equipped with pens and paper. Sidewalk Labs offered this as an example of data collection that would have to go through the RDUA process and that would require approval by the UDT prior to collection because the data is 'urban data' collected in public space. Another example from the MIDP is the collection of air quality data

[102] In copyright law, for example, authorship of compilations of data is determined based upon who is responsible for the selection or arrangement of the data within the compilation. See, e.g., Geophysical Service Inc v Encana Corp, 2016 ABQB 230, 38 Alta LR (6th) 48, aff'd 2017 ABCA 125, leave to appeal denied 2017 CanLII 80435 (SCC). Under the EC, European Database Directive 96/9/EC of the European Parliament and of the Council of the European Union of 11 March 1996 on the legal protection of databases, [1996] O.J, L 77/20, article 4, ownership is determined based upon who created the database.

[103] MIDP (n 3) Vol II, Ch 5, 416. This is contrasted with "transaction data" for which, according to Sidewalk Labs, consent can be directly obtained from the individual. (See: MIDP (n 3), 426).

[104] See, e.g., Beamish (n 9); Department of Justice, Canada, 'Privacy Principles and Modernized Rules for a Digital Age' (Canada, 21 August 2019), 12-13 https://www.justice.gc.ca/eng/csj-sjc/pa-lprp/dp-dd/modern_1.html accessed 28 April 2020.

[105] For example, a consultation document from Canada's Department of Justice states: "some individuals might fear adverse consequences and feel compelled to consent to the collection of personal information". Since true consent is not possible, collection is instead based on the link to a legal activity by government. See: Department of Justice (n106) 12.

[106] See, e.g., the critique by David Young, 'Sidewalk Labs – Public or Private Data' (*David Young Law*, 2019) http://davidyounglaw.ca/compliance-bulletins/sidewalk-labs-public-or-private-data/ accessed 28 April 2020.

[107] See, e.g., Natasha Tusikov, "'Urban Data' and 'Civic Data Trusts' in the Smart City", (Centre for Free Expression, 6 August 2019) https://cfe.ryerson.ca/blog/2019/08/%E2%80%9Curban-data%E2%80%9D-%E2%80%9Ccivic-data-trusts%E2%80%9D-smart-city accessed 29 April 2020; Keri Grieman, 'Pedestrian Curiosity: A Brief Examination of Consent and Privacy in Swath Section Smart City Spaces' (2019) 7(5) *Spatial Knowledge and Information Canada* 1 http://ceur-ws.org/Vol-2323/SKI-Canada-2019-7-5-1.pdf accessed 28 April 2020. Ontario's Office of the Information and Privacy Commissioner raises concerns that unclear and overlapping roles for private regulators and the UDT create a confusing compliance context. See: Beamish (n 9) 6.

[108] The Ontario Information and Privacy Commissioner (Beamish (n 9), 8) commented on how problematic this would be. In his view, to "expect the City to apply to a non-profit Trust, go through the evaluation process, and commit to contractual undertakings would be inappropriate given the experience, mandate and statutory authority of the City."

[109] In his letter to the Chair of Waterfront Toronto, Commissioner Beamish notes that it is "problematic that, as proposed, the City and other public sector organizations would be expected to apply to the Trust in order to collect or use any Urban Data in the geographical area of the project." (Beamish (n 9) 8). He observes that where the city is required by law to collect data: "To then expect the City to apply to a non-profit Trust, go through the evaluation process, and commit to contractual undertakings would be inappropriate given the experience, mandate and statutory authority of the City."

[110] MIDP (n 3) Vol II, Ch 2, 185. This project was publicized by Sidewalk Labs prior to the development of the MIDP, and was part of the discussion around their RDUA. See: Farrah Merali, 'Sidewalk Labs partners with Toronto groups to collect data for public life study' (*CBC News*, 16 December 2018) https://www.cbc.ca/news/canada/toronto/sidewalk-labs-thorncliffe-park-womens-committee-1.4946336 accessed 28 April 2020.

by an environmentalist.[111] These examples reveal the tension between data in the public domain – free for all to gather and use – and "urban data" in which the UDT would assert some form of control over who collects the data and why. Under this approach, public domain data becomes collective data, subject to control over both collection and use.[112]

This confusion between public domain and collectively 'owned' data was evident in the MIDP. Writing about data governance for Sidewalk Toronto, the company's Alyssa Harvey-Dawson stated "No one has a right to own information collected from Quayside's physical environment — including Sidewalk Labs."[113] At the same time, Sidewalk Labs characterized urban data as a "community or collective asset",[114] suggesting a kind of communal ownership distinct from public sector data.[115] Harvey-Dawson acknowledges the governance gap created by this novel concept when she states: "If no one owns urban data, the question remains: Who manages it in the public interest?"[116] Sidewalk Labs' answer, was of course, the Urban Data Trust, which is discussed in more detail in the following part.

6. Governance: The Urban Data Trust

The third set of questions in the knowledge commons framework addresses governance. This includes a consideration of governance mechanisms and decision-makers, infrastructures and institutions, as well as informal norms and legal structures. In the case of Sidewalk Toronto, the UDT was presented as the governance body for the pool of 'urban data'.

In the PDA, Sidewalk Labs proposed that it would explore the creation of a "data trust" to govern data collected in the Quayside development. This mention of the data trust was short on detail; it was referred to as a "novel form of data governance".[117] This concept evolved into a "civic data trust",[118] which is described by McDonald and Porcaro as "an organizational and legal model that protects the public's interest" in data.[119] Both proposals generated debate and uncertainty about what they meant in terms of governance, with some raising concerns that they could not be 'trusts' in a legal sense.[120] In addition, some critics challenged the appropriateness of using the 'civic data trust' label for the scheme proposed by Sidewalk Labs, which was ultimately a top-down arrangement.[121] In any event, perhaps in response to both sets of criticism, the MIDP, dropped 'civic data trust' and proposed instead an Urban Data Trust, with the qualification that it was not using the word "trust" in its trust law sense.[122] Sidewalk Labs also indicated in the MIDP a reluctance to adopt any solution that depended upon new legal infrastructure (i.e. legislative amendment or new legislation).[123] This reluctance might have been due to a concern about delays and uncertainty that could arise from any solution that would be subject to the vagaries of a political process. Nevertheless, Sidewalk Labs left open the possibility that the new governance body might at some point evolve into a public body, although how this might occur was unclear.[124]

The MIDP contemplated that the final development agreement with Waterfront Toronto would provide for the establishment of the UDT.[125] Once created, it would be a non-profit organization independent of both Sidewalk Labs and Waterfront Toronto. It would have the mandate "to address the digital governance challenges related to urban data while also promoting data driven innovations that benefit

[111] MIDP (n 3), Vol II, Ch 2, 183.

[112] It is difficult to see how a communal data ownership argument could be used to prevent anyone from collecting non-personal data in the public realm. Even in the case of personal data, data protection laws do not prevent the collection of personal data by individuals for purely private reasons, nor do they apply to the collection, use or disclosure of personal information when it is for journalistic, artistic, or literary purposes. (See: Personal Information Protection and Electronic Documents Act, SC 2000, c 5, s. 4(2)(b) and (c); Personal Information Protection Act, SA 2003, c P-6.5, s. 4(3), and Personal Information Protection Act, SBC 2003, c 63, s. 3.) This is due to freedom of expression concerns (Alberta (Information and Privacy Commissioner) v. United Food and Commercial Workers, Local 401, [2013] 3 SCR 733, 2013 SCC 62.).

[113] Harvey-Dawson (n 45). Data in the public domain is not owned. By contrast, a data commons is a pool of data that, although shared, is nonetheless controlled. Observing that data held in a commons is often for specific purposes, Yakowitz describes a data commons as consisting of "public-use research datasets" (Yakowitz (n 76) 6).

[114] MIDP (n 3) Vol II, Ch 5, 418.

[115] Sidewalk Labs gives the example of traffic data, stating: "Since that data originates on public streets paid for by the taxpayers and since the use of that data could have an impact on how those streets operate in the future, that data should become a public resource." MIDP (n 3) Vol II, Ch 5, 418.

[116] Harvey-Dawson (n 45). Goodman and Powles (n 23), 18, argue that "creating a term unrecognized in law, would effectively negate any default privacy setting: everything done within the bounds of the Sidewalk Toronto project would be potentially up for grabs."

[117] Plan Development Agreement (n 34), Schedule 1, 47. The concept of a data trust is quite fluid and open-ended. See, e.g., Hardinges (n 1); Element AI/Nesta, 'Data Trusts: A New Tool for Governance' (*ElementAI*, 2019) https://hello.elementai.com/rs/024-OAQ-547/images/Data_Trusts_EN_201914.pdf accessed on 28 April 2020; 'A Primer on Civic Digital Trusts' (*MaRS*, December 2018) https://marsdd.gitbook.io/data-trust/ accessed on 28 April 2020; Sylvie Delacroix and Neil D. Lawrence, 'Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance' (2019) *International Data Privacy Law* https://doi.org/10.1093/idpl/ipz014 accessed on 28 April 2020; Sean McDonald, 'Reclaiming Data Trusts' (*CIGI*, 5 March 2019) https://www.cigionline.org/articles/reclaiming-data-trusts accessed on 28 April 2020.

[118] See: Harvey-Dawson (n 45). Sidewalk Labs defined the 'Civic Data Trust' as "an independent third party that ensures that value from data goes to the people, communities, government, industry and society from which it was collected and that data privacy and security are protected." (Sidewalk Labs, Digital Governance Proposals (n 65), 12.)

[119] Sean McDonald and Keith Porcaro, 'The Civic Trust' (*Medium*, 4 August 2015) https://medium.com/@McDapper/the-civic-trust-e674f9aeab43 accessed on 28 April 2020. McDonald acknowledges that as the concept is still in evolution, there may be different understandings of what constitutes a civic data trust. Affidavit of Sean McDonald (n 82) 5.

[120] According to the trust model, ownership in data is transferred to the trust which then manages it according to the specified terms. Some argued that data was incapable of this kind of transfer and ownership. A civic trust must also act in the interest of the broader population, and some argued that a data trust would not easily fit within the concept of 'charitable trusts' developed in Canadian law for public benefit trusts. See: Goodman and Powles (n 23), at 19; Mariana Valverde, 'What is a data trust and why are we even talking about it? Sidewalk Labs' magic tricks' (*Centre for Free Expression*, 14 January 2019) https://cfe.ryerson.ca/blog/2019/01/what-data-trust-and-why-are-we-even-talking-about-it-sidewalk-labs%E2%80%99-magic-tricks accessed on 28 April 2020.

[121] McDonald and Porcaro (n 123), state that a civic data trust and its private sector data contributors have a fiduciary duty "to develop participatory governance processes that keep each other in check." This lack of process in the proposed UDT was seen as a key failing.

[122] MIDP (n 3) Vol II, Ch 5, 423.

[123] According to Sidewalk Labs, "housing the Urban Data Trust in a public-sector entity would require new or amended legislation, and the passage of legislation can take time and would need to account for emerging technologies." MIDP (n 3) Vol II, Ch 5, 422. Ontario's Information and Privacy Commissioner, by contrast, urges law reform to provide the necessary legal infrastructure: Beamish (n 9), 8. An Australian report on smart cities notes that "It is those cities that actually enact legislation around their data ecosystem and the panoply of smart cities initiatives that are best placed to shape and control their urban digital futures." ('Governance and the Smart City' (*Energy of Things*, December 2016), 10 https://www.fishermansbend.vic.gov.au/__data/assets/pdf_file/0015/33243/Governance-and-the-Smart-City_EoT_December-2016.pdf accessed on 28 April 2020.

[124] MIDP (n 3) Vol II, Ch 5, 422.

[125] MIDP (n 3) Vol II, Ch 5, 420.

individuals and society."[126]

The UDT as proposed in the MIDP would have consisted of five members (at least initially). The nature and composition of the UDT was dictated by the concept of 'urban data' as being neither public nor private sector data, and subject to some form of 'public' or communal ownership. Thus, the five proposed members were meant to represent different interested parties in this data. One would be chosen for his or her expertise in data governance and legal issues. The other four would represent different 'interest groups': academic, public, private and community.[127] This suggested a commonality in interests within each of these categories – something that could not be safely said about any of them.[128] Beyond this, although the data was seen as being collective or communal data and while it was clearly expected that much of this data might be human behavioural data, the "community" received only one seat on a board of five.[129]

The UDT was meant to govern urban data by controlling who was entitled to collect and use this data, and by setting the terms and conditions. This was to be carried out through the RDUA – a combination of application form and ethics approval request to be filed prior to commencing the collection of data in the designated area. Parties seeking to use urban data collected by someone else would file RDUAs explaining the nature and purpose of their proposed use. The RDUA would require the incorporation of privacy-by-design principles, and would specify that data must be used for a "beneficial purpose" which "must incorporate Canadian values of diversity, inclusion, and privacy as a fundamental human right."[130] The purposes for collection and use would have to be clear and transparent. Data would be deidentified by default, stored securely, and collection would be minimized. Data must not be sold or used for advertising without explicit consent of the data subjects. Those who wish to use data for the development of AI must also conform to responsible AI use principles.[131] The actual RDUA process would be similar to a privacy impact assessment.[132] While Sidewalk Labs acknowledged that the UDT could establish its own guidelines, it proposed the RDUA for at least the initial start up period.[133] Any sensors would have to be mapped and registered with the UDT in a public registry to enhance transparency.[134] Although collected data would be publicly accessible by default, the UDT would have the ability to impose access conditions where this was warranted to protect the public interest. The UDT would also oversee data sharing agreements, access terms and fees. A "data collection and use administration fee" would be part of each data collection/use agreement and would be payable to the UDT to offset its operating costs.[135] The UDT would have the authority (presumably under the terms of the agreements with individual data collectors or users) to audit an organization's practices, to remove sensors in cases of non-compliance, and to seek legal remedies for breaches of conditions.[136] However, as it would not be a public body, nor would it be created by statute, it was unlikely to have any special enforcement powers.[137]

An alternative to the UDT might have been to turn to the public sector for a governance framework. For example, the OIPC suggested that:

> Rather than relying on Sidewalk Labs to develop an appropriate solution, this is an opportunity for the provincial government to take the lead and modernize the laws to address the legislative shortcomings. Amendments could include mandatory requirements for data minimization, additional protections for individual and group privacy, ethical safeguards, and greater enforcement tools for my office, including additional investigation, order making and audit powers.[138]

Public sector governance was specifically rejected by Sidewalk Labs. The Toronto Board of Trade, in a separate proposal, suggested that the Toronto Public Library should operate as a trusted data steward.[139]

In an article on the Sidewalk Labs proposal, Alyssa Harvey-Dawson suggested that the UDT would fill a void because: "Existing laws on urban data do not address ownership."[140] It was thus a concept of governance premised on the idea that the captured data were a communal asset. Yet data exist because someone has captured them, and this act of capture reflects specific choices made by the data collector. In addition, some data, such as personal data, reflect layers of interests. The idea of urban data as a kind of 'terra nullius' masked the existing interests in the data, and it was these interests that needed to be reflected in the design and implementation of a governance framework.

Ultimately, in proposing the UDT, Sidewalk Labs chose a governance model developed unilaterally, and not as part of a collective process involving data stakeholders. It was driven by a sense of urgency that allowed neither collaboration nor even legislative change that might have provided some institutional legal infrastructure. It is perhaps not surprising, therefore, that after its review of the MIDP, Waterfront Toronto rejected both the concept of 'urban data' and the UDT, and

[126] MIDP (n 3) Vol II, Ch 5, 420. Note that Commissioner Beamish (n 9), 7, expresses concerns over limited oversight of the UDT and the fact that it would not be subject to data protection and transparency laws.
[127] MIDP (n 3) Vol II, Ch 5, 420
[128] For example, having a single community representative mistakenly presumes a homogeneous community. It is also not clear whether academia is represented in their research capacity or as a substitute for civil society, which is unrepresented. As for the public sector, three levels of government have an interest in the port lands that are the subject of development, and their interests are not necessarily common. The "business industry representative" presupposes common interests across large, medium and small enterprises.
[129] In Data Trusts (Element AI) (n 123) 21, the authors observe that the UDT "failed to address the types of power imbalances at the core of the issues being discussed, and further exemplified the disenfranchisement of citizens in the decision-making process as to how their personal data is to be used, as the terms of the trust were chosen by Sidewalk Labs in the first place." Commissioner Beamish (n 9) 6, argued that it would be more appropriate to focus on the expertise required by the work of the UDT rather than on representation by sector.
[130] MIDP (n 3) Vol II, Ch 5, 424.
[131] MIDP (n 3) Vol II, Ch 5, 425. Ontario's Information and Privacy Commissioner, in an open letter to Waterfront Toronto, criticized the extent to which the UDT duplicated existing governance regimes for what the Commissioner clearly considers public sector data: Beamish (n 9).
[132] MIDP (n 3) Vol II, Ch 5, 428-429.
[133] MIDP (n 3) Vol II, Ch 5, 424.
[134] MIDP (n 3) Vol II, Ch 5, 433.
[135] MIDP (n 3) Vol II, Ch 5, 422.
[136] MIDP (n 3) Vol II, Ch 5, 435
[137] Enforcement is challenging. The Office of the Information and Privacy Commissioner notes that the UDT would have limited powers of oversight and redress (Beamish (n 9) 7). Commenting on an earlier iteration of the trust, McDonald, supra note 99, expressed concerns that if the UDT were to have the enforcement powers it needed, "we would have to substantially devolve and privatize limited forms of regulatory investigation and punishment authority."
[138] Beamish (n 87) 8.
[139] Donovan Vincent, 'Toronto Public Library should control data collected at Quayside, Board of Trade says' (*Toronto Star*, 9 January 2019), https://www.thestar.com/news/gta/2019/01/09/toronto-public-library-should-control-data-collected-at-quayside-board-of-trade-says.html accessed 3 May 2020.
[140] Harvey-Dawson (n 45).

the parties agreed that the project would move to the next phase without these elements.[141]

While the UDT and the concept of 'urban data' were problematic, their abandonment did not resolve the project's data governance issues. It returned the development to the status quo ante, leaving the private and public sector actors each to manage their data according to existing frameworks. The termination of the project in May 2020 made the immediate issue of data governance moot, although Waterfront Toronto remains committed to developing the Quayside area and any new partner or project may well have to design some form of data governance framework. Recent legislative amendments may have since created more room to innovate in the creation of a knowledge commons in which both public and private sector data can be shared. It remains to be seen whether there will be a willingness among new partners to invest in the design of an appropriate knowledge commons framework.

7. Conclusions

The preceding discussion of the data governance model proposed by Sidewalk Labs for the Sidewalk Toronto development offers an example of a failed governance scheme from which useful lessons may be drawn.

One problem with the UDT as a governance model was that it developed, in part, in response to a diverse range of public criticisms and concerns that were raised following the announcement of the Sidewalk Toronto project. A first problem was the reactive nature of the design of the data governance regime. The knowledge commons and its governance are ideally part of project design from the outset. The concerns were brought forward by many different urban stakeholders, from developers to residents. They included the ability to participate in innovation within the district, concerns over undue surveillance, ethics and human rights, and data localization arguments that combined privacy and sovereignty considerations. An attempt to build governance in response to these diverse concerns led to a data governance framework that tried to do too much and for many different reasons. While Frischman et al observe "Commons governance confronts various obstacles to sustainable sharing and co-operation",[142] not all of the obstacles sought to be overcome by the UDT were about the pooling or sharing of information assets. Rather, some related to the very nature of the development itself. In many ways, the UDT was designed to do too much and to satisfy too many disparate concerns.

A second flaw in the proposal was the decision to base the framework on the novel category of 'urban data'. This category was meant to capture a kind of data in which there might be a multiplicity of stakeholder interests. Yet by basing the definition on a combination of physical geography and uncertain notions of public and private space, the category was both unwieldy and uncertain. Rather than create governance for a pool of data shared by collaborating partners, the MIDP defined a category of data in which no one could claim ownership and subjected it to governance by the UDT. Quite apart from the problems with identifying data as independent of its collectors, this approach distanced the data to be governed from those who would have a clear stake in its governance.

A third flaw was that the governance model proposed was a top-down model originating from a single stakeholder in a complex environment with multiple participants and diverse interests in the data. The lack of an organic process with broad stakeholder engagement was a serious defect. Such a process should have identified who the stakeholders were and then involved them in considering what the data sharing model should look like, what data it should govern, according to what principles, and for whose benefit. It is clear that Sidewalk Labs saw some urgency in the task of designing data governance, found existing legal frameworks lacking, and felt legal change could not happen with sufficient speed or flexibility. Yet all of these factors undermined the legitimacy of what was proposed. "Urban data" was a profoundly problematic category of data, and the "trust" was not a trust in any real sense of the word. The result was a data governance scheme doomed to failure.

The failure of the UDT illustrates the importance of addressing data governance issues at the project design stage; these issues are often intricately intertwined with questions about what data to collect and for what purposes, which in turn are both project design and data governance issues. Further, data sharing necessarily implicates multiple interests, which may be both public and private. The diverse stakeholders need to be able to participate in the conceptualization and design of the data governance model and need representation in its implementation. In this respect, the concept of the 'knowledge commons' is useful and instructive. A knowledge commons does not depend upon the existence of a new type of data. Rather, it is premised upon different data 'owners' choosing to pool or share their data to achieve common goals within carefully set parameters. A knowledge commons shifts the focus from ownership/control over data to governance for sharing, but it does not deny or undermine the rights and interests of those who contribute to the commons. Rather, these form the basis for the interests of the contributors to participate in the governance of the commons.

[141] Diamond (n 55).
[142] Frischman et al (n 4) 23.