Data Governance and Digital Infrastructure:

Analysis and Key Considerations for the City of Toronto

Final Report - June 2020

OpenNorth

OpenNorth

Table of Contents

Table of Contents i

List of Figures iv

List of Tables iv

List of Abbreviations iv

Acknowledgements v

Authors v

Executive Summary vi

Data governance: Concepts and scope vi

Regulatory and legislative ecosystem for smart city data governance vi

Analysis of data governance mechanisms vii

Key considerations and next steps vii

1. Introduction 1

2. Data governance: Concepts and scope 3

2.1. Defining data governance 3

2.2. Domain scope 6

2.3. Organizational scope 7

2.3.1. Single organization data hierarchy 7

2.3.2. Inter-organizational data governance 8

2.3.2.1. Data partnerships 8

2.3.2.2. Data intermediaries 8

2.4. Data scope 11

2.4.1. Smart cities, personal data and privacy 12
2.5. Conclusion 14
3. Regulatory and legislative ecosystem for smart city data governance 16
3.1. Regulatory principles, approaches, and instruments 16
3.1.1. Rights-based approaches 16
3.1.1.1. Fair information principles 17
3.1.1.2. Data ownership 20
3.1.1.3. Indigenous approaches 22
3.1.2. Risk-based approaches 24
3.1.3. Standardization 26
3.2. Legislative context applicable in the City of Toronto 27
3.2.1. Federal legislative context 28
3.2.2. Ontario legislative context 30
3.3. Contrasting Canadian and international legislative frameworks 31
3.3.1. The European Union's General Data Protection Regulation (GDPR) 32
3.3.2. California Consumer Privacy Act (CCPA) 33
3.3.3. Comparing the GDPR and the CCPA 34
3.4. Conclusion 34
4. Analysis of data governance mechanisms 36
4.1. Methodology 36
4.2. Overview of case studies 37
4.3. Observing governance mechanisms in practice 47

4.3.1. Structural mechanisms 47
4.3.1.1. Leadership mechanisms 47
4.3.1.2. Compliance mechanisms 48
4.3.2. Procedural mechanisms 50
4.3.2.1. Planning 50
4.3.2.2. Data acquisition 53
4.3.2.3. Data security and access controls 54
4.3.2.4. Data storage 56
4.3.2.5. Data sharing and publishing 57
4.3.3. Relational mechanisms 58
4.3.3.1. Communication and education 58
4.3.3.2. Coordination 60
4.3.3.3. Stakeholder engagement 61
4.4. Conclusion 64
5. Key considerations and next steps 65
5.1. Key considerations 65
5.2. Future research and next steps 66
References 68

List of Figures

Figure 1: Abraham, Schneider, and vom Brocke's conceptual framework for data governance ..4

List of Tables

Table 1: Examples of smart city technologies used for various city services.....12

Table 2: A taxonomy of privacy breaches and harms.....25

Table 3: Description of all 20 case studies 37

List of Abbreviations

AI	Artificial Intelligence
CCPA	California Consumer Protection Act
FIPPA	The Freedom of Information and Protection of Privacy Act (Canada)
GDPR	The General Data Protection Regulation (European Union)
IoT	Internet of Things
MFIPPA	The Municipal Freedom of Information and Protection of Privacy Act (Ontario)
PHIPA	The Personal Health Information Protection Act (Ontario)
PII	Personally Identifiable Information
PIPEDA	The Personal Information Protection and Electronic Documents Act (Canada)

Acknowledgements

The authors would like to thank Dr. Elizabeth Judge (University of Ottawa) and Jacques Priol (CIVITEO) for contributing expertise, research materials, and writing to early drafts of this report.

Authors

Open North is Canada's leading non-profit organization working for data & technology that empower transparent, accountable and inclusive communities. Open North drives research, capacity-building and network collaboration across and within sectors to advance the responsible and effective use of data & technology.

Lead author: Steven Coutts – Research Analyst, Applied Research Lab, Open North

Co-author: Sarah Gagnon-Turcotte – Director, Applied Research Lab, Open North

Recommended citation:

Coutts, Steven, and Sarah Gagnon-Turcotte. "Data Governance and Digital Infrastructure: Analysis and Key Considerations for the City of Toronto." Open North, 2020.

Executive Summary

As digital infrastructure becomes a part of our cities at an unprecedented scale, the City of Toronto will increasingly have to reckon with the potential risks and impacts involving data that accompany it. The City is currently developing a policy framework and governance model to guide the introduction of connected, smart technologies. It has commissioned this report to anchor its work in the learnings and insights from the data governance field.

This report identifies relevant data governance examples across the world. It investigates a broad range of governance practices and mechanisms in use to protect data, manage risks, and ensure democratic accountability. It is structured around a conceptual framework developed by Abraham, Schneider and vom Brocke during a systematic literature review. It draws considerations from 20 use cases identified by Open North through extensive desk research.

Data governance: Concepts and scope

Data governance in the smart city context is an emerging field. Definition and concepts are still evolving. Therefore, the first section of the report explores and defines key concepts and approaches. We define data governance as follows:

Data governance determines who makes different decisions, how they make them, and how they are held accountable for their role in maintaining or controlling the data of an organization or group.

We introduce three aspects of data governance scope - domain scope, organizational scope, and data scope – and conclude by identifying personal information and the emerging privacy issues associated with it as the center of our research.

Regulatory and legislative ecosystem for smart city data governance

Understanding the range of complementary **legislative and regulatory approaches** to data governance is key to leveraging data for the public interest while improving multi-stakeholder collaboration and mitigating privacy concerns. These approaches include **rights-based approaches,** including fair information principles, data ownership, and Indigenous data governance principles, **risk-based approaches,** and **standardization.**

While the legislative and regulatory context composes only one part of data governance, it is crucial in setting the enabling conditions for different data governance approaches and mechanisms. In response to more stringent privacy legislation introduced in other jurisdictions, Canada is reviewing its laws to determine how they can better respond to the opportunities and

challenges of a digital society. Although the legislative landscape for data governance is set to evolve in the coming years, we discuss various laws governing the collection, sharing, and disclosure of personal information, both in the Canadian context and internationally, with a focus on the protection of personal data and privacy rules.

Analysis of data governance mechanisms

Through our case study analysis, we observed how organizations used a variety of **structural, procedural, and relational mechanisms** to unlock the value of their data while minimizing risk. The analysis of **structural mechanisms** shows that trust, representativeness and accountability are at the center of data governance and are supported by ethical and other compliance instruments. Many **procedural mechanisms** we identified had limitations reflecting a similar need to embed them in principled governance. **Relational mechanisms** illustrate how stakeholders' capacity building and engagement are an integral part of data governance.

We find that the framework proposed by Abraham, Schneider, and vom Brocke is useful in providing a high-level taxonomy of data governance mechanisms from a research perspective. However, it should not necessarily be viewed as a step-by-step recipe for data governance.

Key considerations and next steps

Finally, from the observations drawn from the case study analysis, we synthesized several key success factors under the following themes:

- **Define a clear set of guiding values for data governance**
- **Lead with governance, not technology**
- **Build trust and social license through collaboration and transparent communication**
- **Anticipate new risks for individuals created by new data sources**

We then propose a set of activities that may aid in research and organizational alignment and support internal collaboration on data governance.

- **Engage internal stakeholders**
- **Develop a research agenda**
- **Establish and reinforce feedback loops**
- **Stay connected to national and international conversations**

OpenNorth

Data Governance and Digital Infrastructure:

Analysis and Key Considerations for the City of Toronto

1. Introduction

Toronto is rapidly moving towards becoming a smart city through the introduction of connected technologies - collectively termed **digital infrastructure**.[1] Data and technology initiatives have the potential to improve the lives of residents as well as improve City operations. However, as the recently-cancelled Sidewalk Labs proposal highlighted, they also raise questions about how such developments – including the data they collect and use – should be governed.

In February 2019, in recognition of these challenges, Toronto City Council directed City staff to develop a policy framework and governance model associated with digital infrastructure, and a work plan for implementation. City Council provided further direction in June 2019 for City Staff to evaluate policies on ethical digital standards and create a code of technological practices.

The City of Toronto recognizes the need to take stock of its preparedness on data and technology issues – including data protection, risk management, procurement processes, and democratic accountability. It also recognizes the need to set expectations and build trust in how future digital infrastructure projects will be addressed at the municipal level.

Open North's research responds to the above needs through the following activities:

- Identifying a variety of data governance examples in different contexts across the world;
- Conducting a broad scan of data governance mechanisms used in these examples;
- Analyzing selected cases' response to critical data governance issues and challenges; and
- Synthesizing key considerations to inform the City of Toronto as it moves forward with the creation of its Digital Infrastructure Plan.

[1] The City of Toronto defines digital infrastructure as "infrastructure that creates, exchanges or uses data or information as a part of its operation. Digital infrastructure includes physical structures, cabling and network systems, software systems, data standards and protocols. Some examples include sensors (cameras, GPS sensors, microphones, etc.), broadband and telephone networks, Wi-Fi, apps and open data standards." City of Toronto, "City of Toronto Digital Infrastructure Plan - Discussion Guide," December 2019, 3, https://s.cotsurvey.chkmkt.com/lib/48827/files/1541.pdf.

While we have addressed our analysis to the City of Toronto context as much as possible, this report does not provide detailed guidance on the adaptation or implementation of these mechanisms in the City of Toronto's context.

Our research was guided by criteria set out by the City of Toronto (Appendix A), which ensured that we investigated a wide range of activities and processes, from high-level principles and policies to operational protocols and software tools, with examples drawn from different organizations and sectors from around the world.

Yet, as we note in our conclusion, data governance in the smart city context is still an emerging field. Therefore, tracking and measuring the outcomes of specific initiatives will require future research.

The remainder of the report is divided into four sections.

- First, in **Section 2,** we introduce the conceptual understandings of data governance and personal data that guide the analysis to follow.
- Next, in **Section 3,** we outline key federal, provincial, and First Nations legislative context impacting privacy and data protection in the City of Toronto, as well selected international legislation.
- Then, in **Section 4,** we analyze case examples for evidence of governance mechanisms at work in various contexts.
- Finally, in **Section 5,** we summarize key considerations emerging from the analysis and conclude by recommending areas for further research.

2. Data governance: Concepts and scope

In this section, we begin by presenting a definition of data governance that has guided our research. Next, we outline a conceptual framework that describes and provides a structure for the different elements of data governance. Then, we introduce different ways to think about the scope of data governance: at the domain-level, the organizational level, and the data level. After considering the variety of types and sources of 'big data' collected by digital infrastructure in a smart city context, personal data privacy emerges as the center of our research. Finally, we introduce several broad lenses through which to view data governance.

2.1. Defining data governance

To begin with, we need to define **data**. At a basic level, data can be defined simply as "attributes, properties, or characteristics that describe."[2] In an information technology context, data can be defined as "computerized representations of models and attributes of real or simulated entities."[3]

The more data an organization collects, the higher the need to direct its use in an ongoing and systematic way. Different disciplinary understandings of data and data governance mean that it can be difficult to create clear distinctions between principles, policies, and practices that constitute data governance. As a starting point, we can draw upon definitions from information technology (IT) for guidance. According to Weill, whereas IT management is about what decisions are made, IT governance is about who makes them and how they are held accountable.[4]

This view provides a separation between two levels of abstraction: the 'what' (governance) and the 'how' (management). The Data Management Book of Knowledge (DAMA-DMBOK), for instance, describes data governance as "the exercise of authority, control, and shared decision making over the management of data assets."[5] Kooper, Maes, and Lindgren propose that thinking through the lens of corporate IT governance may be limiting because it is more focused on resource management than the value that can be created through information creation, use, and exchange.[6]

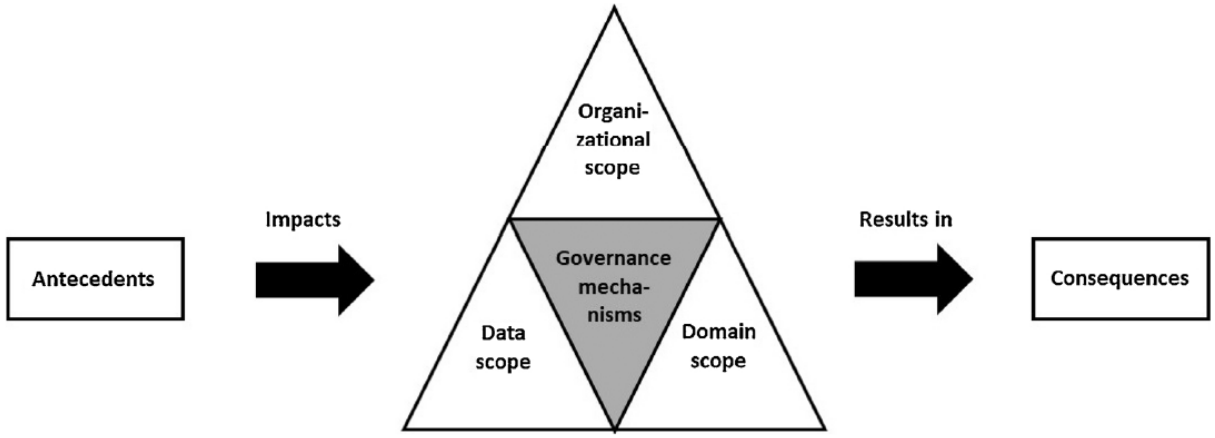
---

[2] Bruhn, "Identifying Useful Approaches to the Governance of Indigenous Data," 2.
[3] Chen et al., "Data, Information, and Knowledge in Visualization," 12.
[4] Weill, "Don't Just Lead, Govern: How Top-Performing Firms Govern IT."
[5] Mosley et al., *DAMA-DMBOK*, 1.
[6] Kooper, Maes, and Lindgreen, "On the Governance of Information," 196.

To think about data governance more expansively, we need a conceptual framework. A conceptual framework provides a vocabulary with which to talk about different elements of data governance. As noted earlier, data governance is often framed in terms of internal corporate business processes. However, there are different roles and responsibilities which need to be considered in a municipal context (for example, the public interest). It is essential to account for both the existing conditions (e.g., legal context, readiness on data issues) and consequences of data governance to understand which factors motivate the adoption of different data governance practices and the effects of those practices.

An example of such a conceptual framework identified through our research process is Abraham, Schneider, and vom Brocke[7] (Figure 1), in which the authors synthesized a common understanding of data governance from an extensive review of research and practitioner publications between 2001-2019. This conceptual framework offers useful insight into how data governance practices emerge, are implemented, and result in outcomes.

Figure 1: Abraham, Schneider, and vom Brocke's conceptual framework for data governance



Antecedents are contingent factors such as the external legal and policy context (see Section 3), as well as internal organizational practices and culture that shape what data governance mechanisms may be employed in a given instance.[8]

At its core, data governance is composed of **governance mechanisms** that come in several varieties: structural mechanisms, procedural mechanisms, and relational mechanisms.[9] These

[7] Abraham, Schneider, and vom Brocke, "Data Governance," 428, fig. 3.
[8] Abraham, Schneider, and vom Brocke, 432.
[9] Abraham, Schneider, and vom Brocke, 427.

mechanisms will be described in greater detail, with examples in Section 4.3, but for now, they can be summarized as follows:

- **Structural mechanisms** determine reporting structures, governance bodies, and accountabilities. They set out what the roles and responsibilities are and allocate decision-making authority.[10]
- **Procedural mechanisms** ensure that data is recorded accurately, held securely, used effectively, and shared appropriately. They comprise the data strategy, policies, standards, processes, procedures, contractual agreements, performance measurement, compliance monitoring, and issue management.[11]
- **Relational mechanisms** facilitate collaboration between stakeholders and include communication, training and coordination and decision-making.[12]

Scope refers to where and over what data governance mechanisms are applied. This has three components: a domain scope, organizational scope, and data scope:

- **Domain scope** refers to the areas in which data governance is intended to help achieve particular goals, such as data quality, data security, data architecture, data lifecycle, metadata, and data storage and infrastructure.[13]
- **Organizational scope** represents the expansiveness of data governance, which can be intra-organizational or inter-organizational. An intra-organizational scope (within an organization or at the project level) focuses on maintaining data quality and integrity. An inter-organizational scope refers to data governance between two or more firms or organizations, which can require more focus on security, sharing, and data standards.[14]
- **Data scope** refers to the type of data to which data governance applies, such as traditional data (including transactional data, reference data, and master data) and big data (including biometric data, sensor and machine-generated data, and web and social media data).[15]

Finally, data governance has **consequences**, both in the short and long-term. In the short term, implementation of data governance can result in improved data processing and operational

[10] Abraham, Schneider, and vom Brocke, 428–29.
[11] Abraham, Schneider, and vom Brocke, 429–30.
[12] Abraham, Schneider, and vom Brocke, 430.
[13] Abraham, Schneider, and vom Brocke, 431–32.
[14] Abraham, Schneider, and vom Brocke, 430–31.
[15] Abraham, Schneider, and vom Brocke, 431.

efficiency. In the long-term, data governance can result in better mitigation or avoidance of risks, such as privacy breaches, and increased trust in the organization. [16]

In this report, we use the following working definition to guide our analysis:

Data governance determines who makes different decisions, how they make them, and how they are held accountable for their role in maintaining or controlling the data of an organization or group.

In the following sections, we unpack the different scopes which apply to data governance: domain scope, organizational scope, and data scope.

2.2. Domain scope

While the goals that an organization wants to achieve by governing its data are antecedent to data governance itself, these goals will determine what the focus areas or **domain scope** of a data governance program will be. Data governance programs commonly address goals in several decision domains including: [17]

- **Data quality**: E.g., ensuring data is able to be used as intended in a given context
- **Data security**: E.g., controlling internal and external access and protecting privacy
- **Data architecture**: E.g., developing and maintaining an enterprise data model
- **Data lifecycle**: E.g., setting up processes and procedures that specify what happens to data from the point of collection, organization, usage, storage, sharing, archiving, up to deletion.
- **Metadata**: E.g., classifying data according to sensitivity level, provenance, and retention period.
- **Data storage and infrastructure**: E.g., managing hardware and software capacity in order to support data quality, security, and lifecycle needs.

There may be other decision domains such as coordination or communication, particularly where data governance occurs between two or more organizations, as discussed in the following section. Regardless, it is essential to consider the areas in which decisions will be made because different objectives require different structural, procedural, and relational mechanisms to implement them.

---

[16] Abraham, Schneider, and vom Brocke, 432–33.
[17] Abraham, Schneider, and vom Brocke, 431.

2.3. Organizational scope

Organizational scope concerns how many separate actors are involved in a given data governance arrangement. [18] On one end of the spectrum, data governance can be applied **within a single organization** (intra-organizational or corporate governance). Data governance can exist **between two or more organizations** directly or via an intermediary body. On the opposite end of the spectrum is data governance **among numerous organizations or actors** – instances of which have been variously called 'data commons,' 'data collaboratives' or 'data cooperatives.'

2.3.1. Single organization data hierarchy

Data governance within a single organization - corporate data governance – can occur at either the project-level or firm-level. [19] It exists to manage and ensure the availability, accessibility, quality, consistency, auditability, and security of data in an organization.[20] It is generally simpler than data governance involving multiple stakeholders, since data is collected and used for internal purposes. Bruhn notes that among the advantages of a single-organization data hierarchy "are its straightforward accountabilities and ability to tailor data creation, management, and storage very closely to the needs of the institution." [21]

Of course, complex organizations also likely need to manage access conditions to data and ensure data is secure from cybersecurity threats. If it is a public sector organization, it may also be subject to stricter regulatory obligations in terms of how they handle specific data. The City of Toronto, for example, may only collect information falling under explicitly named categories in MFIPPA (see Chapter 3 for more detail).

However, it is increasingly rare for an organization to only deal with the data it creates. The contemporary global economy is built on data flows between organizations. Smart cities are often dependent on exchanges of data with outside parties, including private sector digital infrastructure providers. For this reason, data governance must increasingly mediate the relationships between these different actors.

---

[18] Abraham, Schneider, and vom Brocke, 430–31.
[19] Abraham, Schneider, and vom Brocke, 431–32.
[20] Bruhn, "Identifying Useful Approaches to the Governance of Indigenous Data," 4.
[21] Bruhn, 4.

2.3.2. Inter-organizational data governance

Collaboration between organizations involving data raises additional governance issues. Almost inevitably, introducing more actors with a variety of interests creates the need for more complex governance mechanisms, including data exchange standards, service level agreements, and data sharing agreements.[22]

2.3.2.1. Data partnerships

Data partnerships involve two or more parties **co-governing their data**. Bruhn notes that public sector organizations (including large ones such as municipal governments) operate more along the lines of a data partnership since the data they collect does not only concern the organization itself (e.g., clients, employees, transactions) but individuals and organizations external to it (e.g., residents, local businesses). Rather than simply protect their data, public organizations have an incentive to share data across and between departments and agencies.[23] Successful data partnerships may be enabled by trust between parties, a legislative mandate, reciprocal data needs, and matching capacity. To accomplish this, they may set up joint committees to oversee the data governance framework and establish protocols and procedures by which they will share data with one another. [24] Data governance occurs together between parties, with no portion of the relationship outsourced to a third party.

2.3.2.2. Data intermediaries

Since negotiating data ownership, control, and access issues can be difficult and time-consuming, an organization may engage an external actor - a **data intermediary** - to take on these responsibilities. Data intermediaries can help negotiate relationships between data producers and data users in cases where data is sensitive, where one or more parties has less power, resources, or expertise, or where there is a desire to increase trust in the process.

Data intermediary organizations can take several forms, and many concepts have been discussed, such as data brokers, data cooperatives, and data trusts. These are generic models, and, since their real-world implementation may differ substantially from its ideal type, it can be difficult to tell them apart. Looking beyond labels, a key element appears to be the degree to which individual control is retained by the data provider or given over to the intermediary.

---

[22] Abraham, Schneider, and vom Brocke, “Data Governance,” 431.
[23] Bruhn, “Identifying Useful Approaches to the Governance of Indigenous Data,” 5.
[24] Bruhn, 6.

In the smart city context, the model that has received the most attention is the **data trust**.[25] While still a nascent form with few existing examples in practice, recent research has attempted to define the role of data trusts in resolving emerging data challenges.[26] For example, the independent non-profit Open Data Institute (ODI) has explored data trusts through a project funded by the UK Government's Office for Artificial Intelligence and Innovate UK.[27] ODI defines a data trust as: '**a legal structure that provides independent stewardship of data.'** Characteristics described by ODI include: [28]

- As a **steward** of data, a data trust can decide who has access, under what conditions and to whose benefit.
- Data trusts are **independent** from the organizations that hold the data, and prospective data users
- A data trust's **trustees** take on a legally binding responsibility to ensure that the data is shared and used to the benefit of a particular group of people and organizations, as well as other stakeholders affected by its use.
- While data trusts cannot take the form of 'trusts' in a legal sense, they use **legal structures** and forms that take inspiration from them, including the concept of **fiduciary duty:** a legal obligation to act in the best interest of a particular group of people or organizations.
- A data trust takes this concept of enabling an independent institution to hold something – and importantly, to make decisions about its use – and applies it to data.

Before the cancellation of the Quayside project on Toronto's waterfront,[29] a 'civic data trust' was proposed by Sidewalk Labs as a mechanism for safeguarding data collected on-site.[30] Subsequently, a proposal was put forward by the Toronto Region Board of Trade that envisioned a "civic data hub" overseen by the Toronto Public Library.[31] While neither of these proposals have been implemented, the public response to them underscored that independence, legitimacy, and ensuring beneficial use are valued traits of a potential trusted data intermediary,[32] whatever its precise form.

---

[25] Wylie and McDonald, "What Is a Data Trust?"
[26] Delacroix and Lawrence, "Bottom-up Data Trusts."
[27] Open Data Institute, "Data Trusts: Lessons from Three Pilots."
[28] Hardinges, "Defining a 'Data Trust.'"
[29] Vincent and Rider, "Sidewalk Labs Pulls out of Toronto's Quayside Project, Blaming COVID-19."
[30] Sidewalk Labs, "Digital Governance Proposals for DSAP Consultation," 12.
[31] Ruttan et al., "Bibliotech: Beyond Quayside: A City-Building Proposal for the Toronto Public Library to Establish a Civic Data Hub."
[32] Data trust and trusted data intermediary appear to be substantially similar concepts

Data trusts may be useful in cases where multiple parties have different interests in a data set, such as sensitive data (i.e., data subjects may want to limit its use while researchers or companies want to maximize its use). [33] Indeed, it may be more useful to define trusted data intermediaries by the range of issues they could negotiate, which can include:[34]

- Ownership, storage, and usage rights for derivative data generated by the use of primary data;
- Acceptable standards and practices for identifying data subjects, ensuring anonymity, and algorithmic re-identification;
- Access and permissions for analysis;
- Responsibility to inform across the layers of relationships, from data subjects to secondary users;
- Security requirements;
- Tracking and recording the provenance of data and changes in terms over time;
- Alignment and communication of multiple sources or licenses;
- Regulatory and legal conditions of different institutional partners.

Data trusts should not be viewed as an all-encompassing solution, nor as inherently democratic or fair.[35] If a data trust appears to be appropriate, it requires building necessary relationships, aligning stakeholders around a common beneficial purpose, as well as creating strategies and policies to ensure ongoing oversight.[36] All of these activities require continuous effort to maintain.[37]

It is important to note that it is unclear at this point whether current Canadian law allows for the creation of data trusts modelled on legal trusts. However, some of the potential building blocks of data trusts are being considered in updates to into legislation such as PIPEDA. For example, the concept of data portability would allow individuals to move their data under the care of a trusted intermediary of their choosing (see Section 3.1.1.). Whether these proposed changes would necessitate the creation of a new organizational form or the adaptation of existing ones remains to be seen.

The organizational scope of data governance depends on how many actors are involved. For example, corporate data (internal operations, data focusing on institutional needs, etc.) can be dealt with in a single organization data hierarchy. Targeted data sharing can be governed

---

[33] Wylie and McDonald, “What Is a Data Trust?”
[34] Bernholz, “Workshop Summary: Trusted Data Intermediaries,” 3.
[35] McDonald, “Reclaiming Data Trusts.”
[36] Wylie and McDonald, “What Is a Data Trust?”
[37] Open Data Institute, “Data Trusts: Lessons from Three Pilots,” 47–48.

directly between parties or (especially if there are additional layers of concerns such as privacy, ethical usage, etc.) using a trusted intermediary. Finally, where widespread sharing is desired and maintaining strict control is less of a priority, a data commons approach may be most beneficial. As we will see in the next section (data scope), different types of arrangements will be better suited to certain kinds of data.

2.4. Data scope

Every data governance program must specify which type of data is its focus - its data scope. This is important not only for defining the purview of a data governance program, but also because different kinds of data may need to be governed differently.[38] Abraham, Schneider, and vom Brocke distinguish between **traditional data** and **big data**.[39]

Traditional data includes data that governments have long collected about their citizens (e.g., administrative records, national censuses). These activities are typically well-defined by legislation: under MFIPPA, for example, the City of Toronto may only collect personal information necessary to administer municipal services. Still, these efforts have been limited by the human resources available to collect and catalogue the data, and the space required to store physical records.

By contrast, **big data** consists of "huge volumes of diverse, fine-grained, interlocking data produced on a dynamic basis, much of which are spatially and temporally referenced."[40] Big data is primarily collected by connected devices – including what is referred to as the Internet of Things (IoT) – which capture information through a variety of means including sensors and cameras. [41] While the collection of big data implicates many kinds of devices, including smartphones, wearable devices, online platforms, and more, we will focus our discussion on technologies directly used by municipal governments.

Data collected automatically through various types of sensors may be the most significant component of a municipal smart city data governance framework, as it is implicated in numerous aspects of municipal services and operations. The following table (Table 1) provides some typical examples of city functions and services involving automated data collection such

---

[38] Abraham, Schneider, and vom Brocke, "Data Governance," 431.
[39] Abraham, Schneider, and vom Brocke, 431.
[40] Kitchin, "Big Data."
[41] Sensors and actuators embedded or placed on different structures can measure specific conditions including, but not limited to, levels of light, humidity, water, temperature, gas, chemicals, electrical resistivity, acoustics, air pressure, weight, movement, and speed.

as traffic management, waste management, public transit, law enforcement and other government services. [42]

Table 1: Examples of smart city technologies used for various city services

City service	Example technologies[43]
Traffic management	Roadside sensors (Bluetooth) detect smartphones and other devices carried by passengers in vehicles as a way of measuring traffic presence, density, and flow.
Public transit	Smart fare cards record the time and location where they are tapped
Energy and environmental monitoring	"Smart grid" technologies monitor water and power use RFID chipped waste and recycling bins
Law enforcement	Traffic cameras, red-light cameras use license plate recognition to enforce fines for infractions
Government services	Free public WIFI hotspots may collect information about devices that connect to its network

Smart city technologies and the data they collect can help municipalities operate more efficiently and sustainably, improve service delivery, and make better decisions. However, the large volumes of data collected through the daily interaction between residents and the technologies and systems that help the city function can also bring risks for individuals.

2.4.1. Smart cities, personal data and privacy

So, what is different about smart city technology that the data it collects requires special governance? Up to this point, we have been using data to mean **representative data** – that is, data that expresses or represents aspects of real-world observations, computations, experiments, or record-keeping. However, the transition toward 'smarter cities' is predicated on knowing more and more about residents,[44] which involves two other types of data: implied data

[42] Fewer, "Open Smart Cities FAQ."
[43] These technologies may or may not be in use in the City of Toronto and are provided for illustrative purposes only.
[44] Kitchin, Lauriault, and McArdle, "Knowing and Governing Cities through Urban Indicators, City Benchmarking and Real-Time Dashboards."

and derived data.[45] **Implied data** are inferences or predictions, for example based on a person's online activities, while **derived data** are "produced from other data" such as individual counts aggregated and averaged over time. [46]

Some visions of the 'smart city' involve creating more points of connection between our digital devices and online identities, blurring the boundaries between our public and private lives and between our physical and digital experiences. Proponents of this vision claim that the creation of predictive models and profiles can make the city more responsive, efficient, and tailored to residents' needs. On the surface, the city may look much the same as before, but the underlying technologies and data analytics that support them have advanced to a point where just a few data points may be enough to identify an individual. As more data points are collected and compiled, it becomes possible to build up an accurate profile of an individual – including their location, habits and preferences – as they go about their daily life.[47]

Whether this data processing is carried out by government or private sector actors, it has serious implications for privacy, as the Privacy Commissioner of Canada reflected on in the following quotation:[48]

> *Privacy is nothing less than a prerequisite for freedom: the freedom to live and develop independently as a person, away from the watchful eye of a surveillance state or commercial enterprises, while still participating voluntarily and actively in the regular, day-to-day activities of a modern society such as socializing, reading the news, getting information about health issues or simply buying stuff.*

Not only is privacy a pillar of our freedom, privacy has long been a constitutionally protected right under the Canadian Charter that is considered essential for exercising other rights.[49] Individual right to privacy has been interpreted in Canadian case law as including protections for "personal privacy, territorial privacy, and informational privacy".[50] Information privacy refers to

---

[45] Kitchin, *The Data Revolution*, 1.
[46] Kitchin, 1.
[47] Grieman, "Smart City Privacy in Canada," 5–6.
[48] Canada, "A Data Privacy Day Conversation with Canada's Privacy Commissioner."
[49] The right to life, liberty and security of person as well as the right to be secure from unreasonable search and seizure. *Canadian Charter of Rights and Freedoms*, s 7-8, Part 1 of the *Constitution Act*, 1982, being Schedule B to the *Canada Act 1982* (UK), 1982, c 11.
[50] *R v Jarvis*, [2019] 1 SCR 488. https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/17515/index.do

the individual right to control with whom, how much, and for what purpose personal information is disclosed.[51]

Under this regulatory framework, personally-identifiable information (often referred to as PII), that is information that allows the identification of an individual, is accorded more protection than data representing environmental or other non-human phenomena.[52]

As the digital infrastructure expands and collects various types of information, it places individuals at greater risk of identification through the combination of various data sources. For example, information privacy can include:[53]

- locational and movement privacy (to protect against the tracking of spatial behaviour);
- communications privacy (to protect against the surveillance of conversations and correspondence); and
- transactions privacy (to protect against monitoring of queries/searches, purchases, and other exchanges).

A national survey of Canadians conducted in October and November 2018 found that 88 percent of Canadians were "concerned on some level about their privacy in the smart-city context."[54] Protecting against data misuse, data breaches, and exposure to bias and discrimination in the smart city context is therefore essential to create public trust. However, it presents significant challenges. In section 3, we will explore emerging approaches for protecting privacy. We will also see that other rights and principles are coming into play in smart cities initiatives all over Canada.

2.5. Conclusion

In this section, we began by presenting a definition of data governance that has guided our research. Next, we introduced a conceptual framework by Abraham, Schneider and vom Brocke that accounts for how different elements of data governance impact one another. A framework is essential because it can be easy to lose sight of the larger picture when considering each

---

[51] *R v Jarvis*, [2019] 1 SCR 488. See also *R v Dyment*, [1988] 2 SCR 417, citing the control-based definition of Alan F Westin, *Privacy and Freedom* (New York: Atheneum, 1970). https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/375/index.do

[52] Information and Privacy Commissioner of Ontario, "Fact Sheet: What Is Personal Information?"

[53] Kitchin, "Getting Smarter About Smart Cities: Improving Data Privacy and Data Security," 25.

[54] Bannerman and Orasch, "Privacy and Smart Cities: A Canadian Survey," 2.

aspect of data governance in isolation. We will use this framework to provide a structure for the analysis to come.

We introduced three aspects of data governance scope - domain scope, organizational scope, and data scope – and concluded by identifying personal information and the emerging privacy issues associated with it as the center of our research. In the next section, we will introduce several broad lenses through which to approach data governance as well as explore the complex legislative and regulatory context informing data governance in the City of Toronto.

3. Regulatory and legislative ecosystem for smart city data governance

In this section, we provide an overview of current laws as well as recent developments in federal and international legislation. These exist as 'antecedents' within the Abraham, Schneider and vom Brocke model, in that they predict and set the stage for data governance mechanisms to emerge.

While the legislative and regulatory context composes only one part of data governance, it is crucial in setting the enabling conditions for different data governance approaches and mechanisms. We begin with identifying several intersecting principles, approaches, and instruments that frequently emerge in data governance conversations. Next, we highlight some key elements of the provincial, federal, and First Nations legislative and regulatory ecosystem around data protection applicable to the City of Toronto. We note that while Canada has been making progress in updating its privacy legislation, other jurisdictions have advanced towards implementing new approaches in data protection. We conclude this section by briefly exploring features of two of these international approaches and what they tell us about trends in data protection.

3.1. Regulatory principles, approaches, and instruments

Data governance relies on a variety of principles, instruments, and approaches including rights, risk assessment and standards. They are not mutually exclusive; rather, they provide different lenses through which to view data governance. Furthermore, these approaches can be difficult to disentangle from individual instances of data governance with many case studies in our research exhibiting elements of several approaches. This section explores them in greater detail and provides the backdrop against which our case studies have been analyzed.

3.1.1. Rights-based approaches

Rights-based approaches have the goal of establishing acceptable norms for behaviour through broad principles.[55] Rights-based approaches are characterized by a binary logic – an activity either complies or does not.[56] In the context of data protection and privacy legislation, this includes managing risks through fair information principles such as informed individual consent,

[55] *Ex ante*
[56] Gellert, "We Have Always Managed Risks in Data Protection Law," 5.

data minimization,[57] specifying the purpose for data collection, limiting the use of data, protecting data with reasonable security safeguards and providing individuals with data rights such as erasure, rectification, and explanation. Data ownership is another dimension of a rights-based approach to data protection in that property rights entail the ability to control how an asset is used. Finally, Indigenous approaches to data governance, with their emphasis on collective data rights, provide a contrast to models based in the Western legal tradition. These approaches all have strengths as well as limitations.

3.1.1.1. Fair information principles

Rights-based approaches – including commitments, declarations, manifestos, and charters – have proliferated as a means of articulating and demonstrating adherence to a set of values or positions on data governance. While having no legal status or force, statements of principle such as Canada's Digital Charter represent a public commitment to uphold existing digital rights of citizens and introduce new rights, such as data portability. For example, the Declaration of Cities Coalition for Digital Rights, signed by the City of Toronto in October 2019, outlines five key principles supporting human rights including privacy, freedom of expression, and democracy:[58]

> ***Universal and equal access to the internet, and digital literacy:*** *Everyone should have access to affordable and accessible internet and digital services on equal terms, as well as the digital skills to make use of this access and overcome the digital divide.*
>
> ***Privacy, data protection and security:*** *Everyone should have privacy and control over their personal information through data protection in both physical and virtual places, to ensure digital confidentiality, security, dignity and anonymity, and sovereignty over their data, including the right to know what happens to their data, who uses it and for what purposes.*
>
> ***Transparency, accountability, and non-discrimination of data, content and algorithms:*** *Everyone should have access to understandable and accurate information about the technological, algorithmic and artificial intelligence systems that impact their lives, and the ability to question and change unfair, biased or discriminatory systems.*

[57] Data minimization refers to practices that limit the personal data collected and processed from individuals to include only information that is relevant or necessary to accomplish specific purposes.
[58] Cities Coalition for Digital Rights, "Declaration of Cities Coalition for Digital Rights."

> ***Participatory democracy, diversity and inclusion:*** *Everyone should be represented on the internet, and collectively engage with the city through open, participatory and transparent opportunities to shape the technologies designed for them, including managing our digital infrastructures and data as a common good.*
>
> ***Open and ethical digital service standards:*** *Everyone should be able to use the technologies of their choice, and expect the same level of interoperability, inclusion and opportunity in their digital services. Cities should define their own technological infrastructures, services and agenda, through open and ethical digital service standards and data to ensure that they live up to this promise.*

Principles dealing specifically with privacy rights were some of the first to come into existence. The first internationally agreed set of privacy principles, published in 1980 by the Organization for Economic Cooperation and Development (OECD),[59] influenced subsequent privacy regulation around the world. The OECD guidelines were updated in 2013 to address the impact of new technologies.[60] Subsequent legislation such as PIPEDA, the GDPR, and the Council of Europe's Convention 108 has codified similar sets of fair information principles.[61] For reference, PIPEDA's fair information principles are as follows:[62]

> ***Principle 1 - Accountability:*** *An organization is responsible for personal information under its control. It must appoint someone to be accountable for its compliance with these fair information principles.*
>
> ***Principle 2 - Identifying Purposes:*** *The purposes for which the personal information is being collected must be identified by the organization before or at the time of collection.*
>
> ***Principle 3 - Consent:*** *The knowledge and consent of the individual are required for the collection, use, or disclosure of personal information, except where inappropriate.*

---

[59] OECD, "Guidelines Governing the Protection of Privacy and Transborder Flows (1980)."
[60] OECD, "Guidelines Governing the Protection of Privacy and Transborder Flows (2013)."
[61] Privacy International, "The Keys to Data Protection: A Guide for Policy Engagement on Data Protection."
[62] Office of the Privacy Commissioner of Canada, "PIPEDA Fair Information Principles."

> ***Principle 4 - Limiting Collection:*** *The collection of personal information must be limited to that which is needed for the purposes identified by the organization. Information must be collected by fair and lawful means.*
>
> ***Principle 5 - Limiting Use, Disclosure, and Retention:*** *Unless the individual consents otherwise or it is required by law, personal information can only be used or disclosed for the purposes for which it was collected. Personal information must only be kept as long as required to serve those purposes.*
>
> ***Principle 6 - Accuracy:*** *Personal information must be as accurate, complete, and up-to-date as possible in order to properly satisfy the purposes for which it is to be used.*
>
> ***Principle 7 - Safeguards:*** *Personal information must be protected by appropriate security relative to the sensitivity of the information.*
>
> ***Principle 8 - Openness:*** *An organization must make detailed information about its policies and practices relating to the management of personal information publicly and readily available.*
>
> ***Principle 9 - Individual Access:*** *Upon request, an individual must be informed of the existence, use, and disclosure of their personal information and be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.*
>
> ***Principle 10 - Challenging Compliance:*** *An individual shall be able to challenge an organization's compliance with the above principles. Their challenge should be addressed to the person accountable for the organization's compliance with PIPEDA, usually their Chief Privacy Officer.*

While self-regulation through principles is one way of demonstrating voluntary compliance with established values and norms around data governance, it may be a relatively weak approach unless encompassed within a broader regulatory framework and aligned across different areas of government.[63]

It has been argued that rights-based approaches to privacy have limitations. Informed consent, for example, is challenging to implement in practice because many people just click 'yes' to

[63] Bennett and Raab, "Revisiting the Governance of Privacy," 5.

privacy policies without reading them.[64] Furthermore, even if consent has been given for one particular use of a person's data, it is almost impossible to trace data back to an individual after the fact in order to gain their consent to use their data for a secondary purpose.[65] Furthermore, limiting use of data for a specified purpose is challenged by machine learning applications, which are "predicated on the collection of as much data as possible for purposes to be determined as the processing unfolds."[66] Therefore, rights-based approaches – while providing principled guidance – may need to be supplemented with a more practical approach.

3.1.1.2. Data ownership

The question of data ownership rights – who owns what and what that ownership means – often arises in the context of the new digital economy in which data has become a valuable commodity.[67] Data ownership rights "provide a powerful basis for control"[68] which is important because data often involves many competing interests. Just as a company may wish to own its data to commercialize it, a government may assert its ownership right to its data to earn revenue or, conversely, make it available as open data. Under emerging legal frameworks such as the EU's GDPR, individuals can assert rights of consent, erasure, and portability, giving them "quasi-ownership" rights over their personal data.[69]

Data is unlike other kinds of goods in several respects. First, it is non-rivalrous, meaning the original creator can give an exact copy of data to another party without losing any part of the original.[70] Data can also be about multiple people – consider, for example, that your genetic information is also about your family – which makes it difficult for an individual to claim exclusive ownership. One useful frame is to think of a "bundle of data rights" that gives individuals protection in terms of how data about them is treated, including the right to have their data anonymized and used only in a reasonable manner.[71] In Canada, courts have held that individuals may have the right to access and correct their personal information, but these rights stop short of ownership.[72]

---

[64] Scassa, "Enforcement Powers Key to PIPEDA Reform."
[65] Gellert, "We Have Always Managed Risks in Data Protection Law," 3.
[66] Gellert, 3.
[67] The British Academy, techUK, and The Royal Society, "Data Ownership, Rights and Controls."
[68] Scassa, "Data Ownership," 2.
[69] Scassa, 2.
[70] The British Academy, techUK, and The Royal Society, "Data Ownership, Rights and Controls," 5.
[71] The British Academy, techUK, and The Royal Society, 6.
[72] Scassa, "Data Ownership," 13.

While there are several sources of ownership rights under Canadian law – including the *Copyright Act*, *Patent Act*, and *Trade-marks Act* – it is currently unclear whether data are subject to ownership rights. It is helpful to refer to the concepts of representative, implied, and derived data introduced in Section 2.4.1.[73] Current copyright law recognizes a distinction between raw 'facts' or 'ideas' (representative data) and processed 'information'; the latter is subject to copyright protection while the former is not. Compilations of facts, including a data set to which new data is continually being added, may not be covered by copyright protection.

Furthermore, while data may not be subject to copyright per se, software applications (which generate and process data) are protected.[74] The situation becomes more complicated when implied or derived data are involved. For example, courts in both the US and Canada have "found sufficient authorship in data generated either by non-AI algorithms or by complex processes such as those used in the collection of underwater seismic data."[75] Even when data has not directly involved a creative act by a human, there is not necessarily a clear-cut answer to the question of data ownership; it depends on the circumstances.

Smart city technologies often require specialized, proprietary software to function. A city may own a physical piece of smart city technology (e.g., a connected street lamp), but it may license the software from a vendor. Even if the collected data itself is not subject to intellectual property protection, it may be processed and stored in a proprietary format and protected by technological protection measures. This could create 'vendor lock-in,'[76] where the city would be required to pay license fees to a technology vendor to access collected data (which may itself be in the public domain) on an ongoing basis.

Open-source licensed software and data standards have been promoted as a potential solution to lock-in. **Barcelona**, for example, has included the principle of "technological sovereignty" as a critical pillar of its Digital City Plan.[77] Technological sovereignty is based around three principles: The transition and use of free or open-source software, interoperability of services and systems, and the use of open standards.[78] These principles are intended to guide the creation of an open public data infrastructure, placing the public interest at the heart of future smart city

---

[73] Scassa, "Data Ownership."
[74] Fewer, "Open Smart Cities FAQ."
[75] Scassa, "Data Ownership," 10.
[76] Fewer, "Open Smart Cities FAQ."
[77] Barcelona City Council. Office for Technology and Digital Innovation, "Barcelona City Council Technological Sovereignty Guide"; Bria and Bain, "Manifesto in Favour of Technological Sovereignty and Digital Rights for Cities : Ethical Digital Standards."
[78] Barcelona City Council. Office for Technology and Digital Innovation, "Barcelona City Council Technological Sovereignty Guide," 4.

development. Barcelona's model needs further evaluation, but an open approach may be one way to sidestep the complexities of creating new data rights.

In a smart city where technology vendors control much of the data that is created, there is often a call for cities to take back control by 'owning' their data. But data ownership may be a limiting framework. As Teresa Scassa notes, "within this rapidly evolving data environment, and with flexible and adaptable legal tools and principles already in place, a cautious "wait-and-see" approach is preferable to the creation of a new *sui generis* right."[79] Defining new data rights may solve some issues while creating new ones, including limiting access and reuse in ways that are detrimental to the public interest.[80] In the next section, we discuss a set of First Nations principles based in the idea of collective data ownership that, while not being law, respond to a specific set of issues in order to give back control of data to First Nations communities.[81]

3.1.1.3. Indigenous approaches

Historically, researchers and government officials have entered First Nations communities and collected data (including biological samples). This data has been used to stigmatize and cause harm to those communities.[82] This has been allowed to occur because, under Canadian law, privacy legislation applies only to governments. Since many First Nations communities are not recognized as 'governments,' separate legislation applies. For example, population health data associated with First Nations communities, is not subject to provincial health privacy laws, but rather is allowed to be collected and disclosed under the *Access to Information Act*.[83]

The Indigenous data sovereignty movement is rooted in a desire to reclaim control over traditional knowledge and data about Indigenous communities. The OCAP® principles are one Indigenous data governance approach that guides the use of data about First Nations (though some principles may be shared with other Indigenous groups such as Inuit and Métis). In contrast to European settler approaches which privilege individual rights, Indigenous data governance principles center the concept of **collective (or community) privacy rights.**

---

[79] Scassa, "Data Ownership," 17.
[80] Scassa, 17.
[81] First Nations Information Governance Centre, "The First Nations Principles of OCAP®."
[82] Bruhn, "Identifying Useful Approaches to the Governance of Indigenous Data," 9.
[83] Stinson, "Healthy Data: Policy Solutions for Big Data and AI Innovation in Health."

It is important to note that OCAP principles are not recognized in Canadian law and can therefore only be implemented through agreements.[84] These are articulated through the following four principles:[85]

> ***Ownership*** *refers to the relationship of First Nations to their cultural knowledge, data, and information. This principle states that a community or group owns information collectively in the same way that an individual owns his or her personal information.*
>
> ***Control*** *affirms that First Nations, their communities, and representative bodies are within their rights in seeking to control over all aspects of research and information management processes that impact them. First Nations control of research can include all stages of a particular research project-from start to finish. The principle extends to the control of resources and review processes, the planning process, management of the information and so on.*
>
> ***Access*** *refers to the fact that First Nations must have access to information and data about themselves and their communities regardless of where it is held. The principle of access also refers to the right of First Nations communities and organizations to manage and make decisions regarding access to their collective information. This may be achieved, in practice, through standardized, formal protocols.*
>
> ***Possession*** *While ownership identifies the relationship between a people and their information in principle, possession or stewardship is more concrete: it refers to the physical control of data. Possession is the mechanism by which ownership can be asserted and protected.*

While OCAP principles do not apply to non-First Nations data, they illustrate an alternate approach to data governance from which lessons may be drawn – especially in viewing privacy as a collective right as opposed to simply an individual one. In Section 4, we will see one implementation of OCAP principles.

---

[84] First Nations Information Governance Centre, “Understanding the Basics of OCAP®,” 4.
[85] First Nations Information Governance Centre, “The First Nations Principles of OCAP®.”

3.1.2. Risk-based approaches

Risk-based approaches to data protection make use of risk analysis tools to assess and manage risks associated with proposed data processing activities.[86] Risk-based approaches are sometimes positioned as opposite to rights-based approaches, substituting universally-applicable legal principles (i.e., rights) with calculated, contextual risk analysis. Despite these differences, however, Gellert considers them to be "twin practices" as they are both concerned with "which [data] processing can take place and under what conditions."[87]

Risk-based approaches build on existing individual privacy rights in human rights treaties, constitutional law, statutory law, and data protection regimes. [88] A risk-based approach "aims to bridge the gap between high-level privacy principles on the one hand, and compliance on the ground on the other, by developing a methodology for organisations to apply, calibrate and implement abstract privacy obligations based on the actual risks and benefits of the proposed data processing."[89]

Risk-based approaches seek to identify the threats to personal data and their causes in advance to protect privacy more effectively. According to the Centre for Information Leadership (UK), "the question should be whether there is a significant likelihood that an identified threat could lead to a recognised harm with a significant degree of seriousness."[90] Daniel Solove[91] provides a categorization of many socially-recognized (if not necessarily legally-recognized) privacy breaches and harms that can potentially result through a variety of inappropriate practices (Table 2).[92]

---

[86] Centre for Information Policy Leadership, "A Risk-Based Approach to Privacy."
[87] Gellert, "We Have Always Managed Risks in Data Protection Law," 17.
[88] Gellert, "We Have Always Managed Risks in Data Protection Law."
[89] Centre for Information Policy Leadership, "A Risk-Based Approach to Privacy," 1.
[90] Centre for Information Policy Leadership, 4.
[91] Solove, "A Taxonomy of Privacy."
[92] Kitchin, "Getting Smarter About Smart Cities: Improving Data Privacy and Data Security," 25–26.

Table 2: A taxonomy of privacy breaches and harms

Domain	Privacy breach	Description
Information collection	Surveillance	Watching, listening to, or recording of an individual's activities
	Interrogation	Various forms of questioning or probing for information
Information processing	Aggregation	The combination of various pieces of data about a person
	Identification	Linking information to particular individuals
	Insecurity	Carelessness in protecting stored information from leaks and improper access
	Secondary use	Use of information collected for one purpose for a different purpose without the data subject's consent
	Exclusion	Failure to allow the data subject to know about the data that others have about her and participate in its handling and use, including being barred from being able to access and correct errors in that data
Information dissemination	Breach of confidentiality	Breaking a promise to keep a person's information confidential
	Disclosure	Revelation of information about a person that impacts the way others judge her character
	Exposure	Revealing another's nudity, grief, or bodily functions
	Increased accessibility	Amplifying the accessibility of information
	Blackmail	Threat to disclose personal information
	Appropriation	The use of the data subject's identity to serve the aims and interests of another

	Distortion	Dissemination of false or misleading information about individuals
Invasion	Intrusion	Invasive acts that disturb one's tranquility or solitude
	Decisional interference	Incursion into the data subject's decisions regarding her private affairs

Source: Solove, Daniel J. "A Taxonomy of Privacy", as reproduced in Kitchin, Rob "Getting Smarter About Smart Cities", 25-6.

Risk-based approaches will continue to be an essential part of a municipal data governance toolkit in the context of the smart city. Although not all of the potential privacy breaches or harms listed above will be implicated with the use of a given smart city technology, cities must have a current understanding of the risks posed by new technologies.

3.1.3. Standardization

Standards are procedural data governance mechanisms that ensure that data is represented and treated consistently between organizations and across sectors.[93] They emanate from both domestic and international arenas, implicate multiple stakeholders, and act as a supplement to data protection legislation.[94] As Bennett and Raab note, "standards could fill important gaps in the enforcement regime, relieve regulators of compliance work and serve as credible methods of certification for transnational transfers of data." [95]

Standards can be defined internally or externally. For example, the Mobility Data Specification - a standard for e-scooter and private transportation company data - began as an internal project of the Los Angeles Department of Transportation before responsibility for maintaining the code base was transferred to the Open Mobility Foundation. External standards bodies such as the International Standards Organization continue to play a significant role in data governance, such as through the ISO 27000 family of information security management standards which, as of 2019, specifies requirements "for establishing, implementing, maintaining and continually improving a privacy-specific information security management system."[96]

[93] Abraham, Schneider, and vom Brocke, "Data Governance," 429–30.
[94] Bennett and Raab, "Revisiting the Governance of Privacy," 8.
[95] Bennett and Raab, 8.
[96] Naden, "Tackling Privacy Information Management Head On."

Data standardization initiatives are a key example of the variety of approaches that have been taken at multiple levels to govern data. For instance, regarding geospatial data in Canada, there are both internal and extra-territorial (i.e., foreign) influences on data standardization through organizations such as the International Standards Organization (ISO), public-private consortia such as the Open Geospatial Consortium (OGC), and bilateral collaborations such as the OGC Arctic Spatial Data Pilot (a partnership between the United States Geological Survey and Natural Resources Canada).

In the digital context, the Standards Council of Canada is currently working on data governance and artificial intelligence (AI) standards, such as through its Canadian Data Governance Standardization Collaborative, as well as the European Union's General Data Protection Regulation (GDPR) through its Canadian Advisory Committee on the General Data Protection Regulation. The CIO Strategy Council – composed of private-sector based Chief Information Officers (CIOs) – has also published a standard on the ethical design and use of automated decision systems. The Pan-Canadian Trust Framework, though not a standard as such, is a framework that relates different standards, guidelines, policies and practices in the area of digital identities and relationships.[97] Each of these bodies is investigating the impacts of current issues of data governance and their potential impacts on Canadian governments, businesses, and other stakeholders.

Accordingly, it is essential to keep existing Canadian involvement in various bodies and initiatives, such as those noted above, in mind when looking to develop data governance at any level of government to avoid duplication of efforts and maintain interoperability across all domains, including privacy protection.

3.2. Legislative context applicable in the City of Toronto

In Toronto, both federal and provincial privacy laws are applicable to the collection, sharing, and disclosure of personally identifiable information. However, precisely which rules apply depends on the actor involved in data collection (e.g. federal or provincial institution, municipality, private firm).

[97] Digital ID and Authentication Council of Canada (DIACC) Trust Framework Expert Committee, "Pan-Canadian Trust Framework Overview: A Collaborative Approach to Developing a Pan-Canadian Trust Framework."

3.2.1. Federal legislative context

Privacy is protected through a framework of laws that impose obligations on the public sector and the private sector. Individual privacy is protected constitutionally under the *Canadian Charter of Rights and Freedoms*.[98] *The Privacy Act* sets out how federal government institutions must deal with personal information.

For private companies engaged in commercial activities, the relevant data protection legislation is the *Personal Information Protection and Electronic Documents Act* (PIPEDA). This Act protects individual privacy by regulating the collection, use, and disclosure of personal information. The *Digital Privacy Act*, which came into force on November 1, 2018, updated PIPEDA by requiring additional record-keeping, reporting, and notification around data breaches.

As part of the overall direction set by Canada's Digital Charter, there has been movement on different issues within different areas of government. For example, as of the time of this writing, Innovation, Science and Economic Development Canada is advancing several proposals to modernize PIPEDA. These proposals cover four main areas: enhancing individuals' control, enabling responsible innovation, enhancing enforcement and oversight, and areas for ongoing assessment.[99]

- **Enhancing individuals' control**: This could include requiring organizations to provide plain-language explanations of the intended use of the information they collect and third parties with whom it will be shared, as well as prohibiting the bundling of consent into contracts. It could also include providing for data portability – the explicit right for individuals to direct that their data be moved from one organization to another. Finally, there could be provisions for individuals to maintain their online reputation, including the right to request deletion of personal information.
- **Enabling responsible innovation:** Possible options in this area include encouraging the responsible use of de-identified data for research and innovation without consent through data trusts, as well as incentivizing the adoption of technical standards and codes.
- **Enhancing enforcement and oversight:** Proposals to strengthen the Privacy Commissioner's mandate in areas of education and research, investigation and audit,

[98] The right to life, liberty and security of person as well as the right to be secure from unreasonable search and seizure. *Canadian Charter of Rights and Freedoms*, s 7-8, Part 1 of the *Constitution Act*, 1982, being Schedule B to the *Canada Act 1982* (UK), 1982, c 11.
[99] Canada. Innovation, Science and Economic Development Canada, "Strengthening Privacy for the Digital Age: Proposals to Modernize the Personal Information Protection and Electronic Documents Act."

tools to address non-compliance (such as the power to make cessation or records preservation orders or extend the range of fines), and ability to proactively engage stakeholders on technology issues.

- **Areas of ongoing assessment:** While seeking to maintain PIPEDA's basis in principles and technological neutrality, there is a recognition that its language and structure is difficult for individuals and organizations to understand. Furthermore, PIPEDA needs to be reviewed in light of emerging business models and organizations which do not act as data "processors" or "controllers," or are collecting data for non-commercial purposes.

The Department of Justice is also considering updates to the *Privacy Act* which would bring it into step with principles and rules in data protection laws such as the EU's GDPR.[100]

Other bodies of law, such as **intellectual property law** – including the *Copyright Act*, *Patent Act*, and *Trade-marks Act* – may also protect individual privacy by controlling the use of confidential information and appropriation of personality. However, as noted in Section 3.1.1.2., the application of these laws to data depends on the circumstances. **Tort law** allows claimants to seek a civil remedy in the event of a privacy breach or harm.[101] All of these laws are built on a rights-based model and recognize that privacy is a right held by individuals. The Supreme Court of Canada has emphasized that:

> *Privacy is at the heart of liberty in a modern state… essential for the well-being of the individual… [and] has profound significance for the public order. The restraints imposed on government to pry into the lives of the citizen [sic] go to the essence of a democratic state.*[102]

The Supreme Court of Canada has held that, even in public or semi-public places, individuals may still have a reasonable expectation of privacy – though the extent of this expectation of privacy is subject to different factors.[103] These factors include the location or manner in which the surveillance takes place, the subject matter or content to be recorded, the relationship between the parties, the degree of consent involved, and the vulnerability of the person to be observed. The need for these factors to be weighed in each individual case of a privacy breach

---

[100] Canada. Department of Justice, "Modernizing Canada's Privacy Act."
[101] See Fewer, "Open Smart Cities FAQ," for other bodies of law that may be applicable in the smart city context, including contract law, criminal law, intellectual property law, competition law, and environmental law.
[102] *R v Dyment,* [1988] 2 SCR 417, citing Alan F. Westin, *Privacy and Freedom* (1970), pp. 349-50, *per* La Forest J, concurring.
[103] *R v Jarvis*, [2019] 1 SCR 515. https://scc-csc.lexum.com/scc-csc/scc-csc/en/item/17515/index.do

affirms Chief Justice Wagner's statement in the Court's decision, that "'privacy,' as ordinarily understood, is not an all-or-nothing concept."[104]

3.2.2. Ontario legislative context

In Ontario, several pieces of data protection legislation are relevant to the public sector. Personal health information is regulated by the *Personal Health Information Protection Act* (PHIPA), which applies to personal health information held by hospitals, pharmacies and other health information custodians. The *Municipal Freedom of Information and Protection of Privacy Act* (MFIPPA) applies to municipalities and municipal institutions in Ontario, and the *Freedom of Information and Protection of Privacy Act* (FIPPA) applies to the Ontario provincial government as well as hospitals, universities, and other specified agencies. In these last two acts, personal information is defined as follows:[105]

> *"personal information" means recorded information about an identifiable individual, including,*
>
> *(a) information relating to the race, national or ethnic origin, colour, religion, age, sex, sexual orientation or marital or family status of the individual,*
>
> *(b) information relating to the education or the medical, psychiatric, psychological, criminal or employment history of the individual or information relating to financial transactions in which the individual has been involved,*
>
> *(c) any identifying number, symbol or other particular assigned to the individual,*
>
> *(d) the address, telephone number, fingerprints or blood type of the individual,*
>
> *(e) the personal opinions or views of the individual except where they relate to another individual,*
>
> *(f) correspondence sent to an institution by the individual that is implicitly or explicitly of a private or confidential nature, and replies to that correspondence that would reveal the contents of the original correspondence,*
>
> *(g) the views or opinions of another individual about the individual, and*

---

[104] *R v Jarvis*, [2019] 1 SCR 515 at para. 41.
[105] Municipal Freedom of Information and Protection of Privacy Act, R.S.O. 1990, c. M.56, sec. 2; Freedom of Information and Protection of Privacy Act, R.S.O. 1990, c. F.31, sec. 2.

> *(h) the individual's name where it appears with other personal information relating to the individual or where the disclosure of the name would reveal other personal information about the individual.*

According to the Information and Privacy Commissioner of Ontario (IPC), "when municipalities contract with private sector organizations to carry out activities that involve the collection, use or disclosure of personal information, compliance with MFIPPA is of the utmost importance."[106] In this sense, public-private partnerships may be subject to more strict controls (MFIPPA does not allow the collection of personal information on the basis of consent), while private enterprises have more leeway under PIPEDA to collect personal information as long as there is consent.[107]

However, while non-personal information does not fall under these privacy regulations and is not covered by consent requirements, distinguishing between personal and non-personal information is a major issue, as noted in Section 2.4.1.

Separate privacy law statutes governing the collection of personal information for commercial and public purposes may produce confusion among the public. This confusion may lead to higher expectations around privacy and consent as they navigate what they perceive to be public space, but which are increasingly governed by public-private partnerships and other pseudo-public arrangements.

3.3. Contrasting Canadian and international legislative frameworks

While the Canadian and Ontario have functioning privacy legislation frameworks, it is a fragmented landscape, and the emergence of new data-driven business models has exposed gaps in current approaches. For example, PIPEDA has been criticized for its consent regime, enforcement model, and its lack of appropriate incentives to ensure that organizations comply. It has also been recommended that the legislation be updated to improve individual control of personal data and increase transparency within organizations.[108]

Looking to other jurisdictions can provide a sense of the direction in which the legal landscape is evolving. Furthermore, developments in the international arena are influential in terms of establishing global norms and best practices in the area of data protection. In this section, we

---

[106] Information and Privacy Commissioner of Ontario to Waterfront Toronto, "Re: Sidewalk Lab's Proposal," September 24, 2019.
[107] Information and Privacy Commissioner of Ontario to Waterfront Toronto.
[108] Canada. Innovation, Science and Economic Development Canada, "Strengthening Privacy for the Digital Age: Proposals to Modernize the Personal Information Protection and Electronic Documents Act."

explore two examples that illustrate current trends toward more clearly defined rights and obligations relating to data: the European Union's *General Data Protection Regulation* and the *California Consumer Protection Act.*

3.3.1. The European Union's General Data Protection Regulation (GDPR)

The European Union's General Data Protection Regulation (GDPR) is a risk-based legislative framework that updates the European Union's data protection framework to account for automated processing of personal data, such as that described in Section 2.2. Several key features of the GDPR are as follows:[109]

- Strengthened consent requirements;
- Monetary penalties for noncompliance;
- Required notification of data breaches;
- Privacy by design principles incorporated;
- New rights for data subjects;
    - The right to be forgotten (data erasure), and;
    - The right to data portability.

Notably, the GDPR has extraterritorial application in that it applies to non-EU organizations that target anyone residing in the EU (not only to EU citizens) by offering goods or services or by monitoring their behaviour. For example, a Canadian university recruiting international students from the EU may be subject to the GDPR insofar as they process students' personal information.[110]

The GDPR also adopts a risk-based compliance approach to its data security provisions, where the nature of the protective measures corresponds to the likelihood and severity of risk (high risk, risk, or minimal risk). For high-risk activities, data controllers must conduct a Data Protection Impact Assessment (DPIA). For example, the UK Information Commissioner's Office (ICO) lists the following steps for their DPIA:[111]

1. Describe the nature, scope, context and purposes of the processing.
2. Ask data processors to explain and document their processing activities and identify any associated risks.
3. Consider how best to consult individuals (or their representatives) and other relevant stakeholders.

---

[109] Gittens et al., "Understanding the GDPR: A Comparison Between the GDPR, PIPEDA and PIPA."
[110] Privacy Commissioner of Canada, "Privacy Fact Sheet: General Data Protection Regulation."
[111] UK Information Commissioner's Office, "Data Protection Impact Assessments."

4. Consult the data protection officer (DPO).
5. Check that the processing is necessary for and proportionate to the organization's stated purposes, and describe how to ensure compliance with data protection principles.
6. Conduct an objective assessment of the likelihood and severity of any risks to individuals' rights and interests.
7. Identify measures that can be put in place to eliminate or reduce high risks.
8. Record decision-making in the outcome of the DPIA, including any difference of opinion with the DPO or individuals consulted.
9. Implement the measures identified and integrate them into the project plan.
10. Consult the ICO before processing, if high risks cannot be mitigated.
11. Keep DPIAs under review and revisit them when necessary.

A well-written DPIA provides evidence that an organization has considered the risks related to the intended data processing as well as met its broader data protection obligations.

3.3.2. California Consumer Privacy Act (CCPA)

More specific considerations in privacy also have implications for data governance, such as limitations on applicable activities, the legal basis for which to make distinctions, and accountability obligations. The California Consumer Privacy Act has some similarities to the GDPR but has its own distinctions that can result in different governance mechanisms.

The California Consumer Privacy Act (CCPA) was enacted in 2018 and took effect as of January 1, 2020.[112] It applies to businesses where at least one of these conditions are true: 1) gross annual revenues of more than $25 million; 2) buys, sells or receives the personal information of 50,000 or more consumers, households or devices, and/or; 3) derives 50% or more of its annual revenue from the sale of consumers' personal information. The CCPA grants new rights to consumers in California, some of which echo provisions contained in the GDPR:

- Right to know what personal information is collected, used, shared or sold
- Right to delete personal information held by businesses
- Right to opt-out of sale of personal information
- Right to non-discrimination in terms of price or service if exercising a privacy right under CCPA

[112] California. Department of Justice, Office of the Attorney General, "California Consumer Privacy Act (CCPA) Fact Sheet."

3.3.3. Comparing the GDPR and the CCPA

As noted by California's Department of Justice, the CCPA and GDPR differ significantly in terms of who and what activities they cover as well as what requirements they place upon different actors.[113]

- While the GDPR stipulates "the requirement to have a "legal basis" for all processing of personal data," CCPA only requires businesses to obtain consent when there is financial gain involved.[114]
- For example, the GDPR requires that companies undertake a data inventory and map data flows to demonstrate compliance. The CCPA may require additional data mapping to reflect its different requirements.
- Under the GDPR, companies are required to develop processes or systems capable of responding to individual requests for access to personal information and for the deletion of personal information. While these same methods may be applied to handling CCPA consumer requests, businesses may need to reconcile the different definitions of personal information and applicable rules on verification of consumer requests.

A preliminary assessment by DataGuidance and the Future of Privacy Forum indicates that the CCPA may have a more limited scope than the GDPR.[115] While the CCPA incorporates certain rights afforded under GDPR, it primarily applies to organizations that share or sell data (the GDPR applies to all organizations that process data). It also does not establish clear roles (e.g., the roles of Data Controller and Data Processor).[116] However, as the CCPA has been in force for only a short time, further evaluation will be required in the future to assess its approach to privacy protection.

3.4. Conclusion

The legislative landscape for data governance may evolve significantly in the coming years. In response to more stringent privacy legislation introduced in other jurisdictions, Canada is reviewing its own laws to determine how they can better respond to the opportunities and challenges of a digital society. Recalling that the legislative and regulatory context is an

---

[113] California. Department of Justice, Office of the Attorney General.
[114] DataGuidance and Future of Privacy Forum, "Comparing Privacy Laws: GDPR vs. CCPA."
[115] DataGuidance and Future of Privacy Forum.
[116] DataGuidance and Future of Privacy Forum.

antecedent to data governance, changes to existing laws could have significant downstream impacts on organizations.

In this section, we have provided a summary of various laws governing the collection, sharing, and disclosure of personal information, both in the Canadian context and internationally. As we noted in our introduction to Abraham, Schneider and vom Brocke's conceptual framework, legislative context is a precursor to data governance, in that it influences both the scope (organizational, data, domain) and the nature of governance mechanisms (structural, procedural, and relational). An awareness of developments in the international legislative context is vital since these global networks of influence shape the discourse around data governance. This background will help frame the case study analysis, which follows in the next section.

4. Analysis of data governance mechanisms

In this section, we present and discuss selected examples and models of data governance in use in various organizational contexts around the world. We begin with an overview of our research process (more detail can be found in Appendix B), including the guiding criteria we employed to identify relevant cases. Then, based on Abraham, Schneider, and vom Brocke's conceptual framework (introduced in Section 2.1), we present our findings categorized according to the type of governance mechanism.

4.1. Methodology

We identified an initial list of 34 organizations from which case studies could be drawn by consulting recent research from several organizations.[117] We combined this scan with extensive desk research (see Appendix B for further details). We then narrowed the initial list to a final long-list of 20 case studies based on factors such as availability of information, geographical representation (scope requirements necessitated geographic spread across multiple continents) and type of organization.

Once we identified promising examples of data governance, we searched for any potential sources of information relating to these examples. We encountered varying degrees of available documentation, which had to be accounted for in our analysis. It should be noted that self-reported statements made in governmental documents, reports, and interviews were taken at face value due to time constraints. We grouped case studies into three categories according to the level of documentation available:

1. **Low-level documentation:** Data governance outcomes have not been reported.
2. **Documented (self-reported):** Organization has self-reported outcomes, but these have not been externally verified.
3. **Documented (independently):** Reputable sources have externally verified data governance outcomes.

[117] Bass, Sutherland, and Symons, "Reclaiming the Smart City: Personal Data, Trust and the New Commons"; Compute Ontario and ORION, "The Future of Ontario's Data: Fulfilling the Potential of Smart Cities"; Element AI and Nesta, "Data Trusts: A New Tool for Data Governance"; GovLab, "Data Collaboratives"; MaRS Discovery District, "Towards a Smart City Data Trust: Design Recommendations for a Personal Mobility Data Trust"; MaRS Discovery District, "A Primer on Civic Digital Trusts"; Mulgan and Straub, "The New Ecosystem of Trust"; Open Data Institute, "Data Trusts: Lessons from Three Pilots."

As measuring outcomes of data governance presents considerable challenges, we opted to place our focus on analyzing the details of the case studies rather than determining the level of success achieved by the data governance models. We recommend that further research be undertaken to follow up on these cases and evaluate their outcomes. Additionally, since failures can be as instructive as successes, some cases were chosen because they highlighted important issues.

Based on these various sources, we documented each case according to an extensive list of criteria supplied by the City of Toronto (Appendix A). These descriptions are incomplete in some cases due to several factors: lack of documentation in English and French; availability of information regarding outcomes, and; overall maturity level of each example.

The maturity level refers to the time a data governance model had been in operation at the time of documentation. We did not undertake a complete data governance maturity assessment,[118] but subjectively classified them by referring to their stage of implementation.

1. **Nascent:** Data governance model exists as high-level principles that have been adopted, but not operationalized.
2. **Emerging:** Data governance model has recently been operationalized through policy or programs but with little available documentation.
3. **Established:** Data governance model has been operationalized through policy or programs for a significant period, with documentation existing that describes processes and procedures.

Finally, we analyzed the different case studies to identify various data governance mechanisms based on Abraham, Schneider, and vom Brocke's conceptual framework.[119] We also drew upon current thinking among technology and legal scholars to inform our analysis.

4.2. Overview of case studies

In the following table (Table 3), we present a brief description of each of the case studies, organized with the most mature and well-documented cases at the top.

Table 3: Description of all 20 case studies

[118] See for example "Data Governance Maturity Models - IBM."
[119] Abraham, Schneider, and vom Brocke, "Data Governance."

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
Ontario's Smart Energy Metering Program	3	3	Government & public agencies	Canada	As part of Ontario's Smart Energy Metering Program Entity, the Independent Electricity System Operator tracks energy use in the home on an hourly basis and sends this information back to the Local Distribution Company that services and bills the customer. In 2013, smart metering data was identified as a valuable asset and efforts were made to promote innovation by using data while preserving security and privacy of customers.
Silicon Valley Regional Data Trust	3	3	Government & public agencies	USA	The Silicon Valley Regional Data Trust (SVRDT) brings together data from numerous public agencies serving children and families including public schools, health and human services organizations, and juvenile justice systems. SVRDT was established as an initiative of the Santa Clara County Office of Education in partnership with the University of California, Santa Cruz and three counties that comprise Silicon Valley–Santa Clara, San Mateo, and Santa Cruz.
Public Transport Victoria	3	3	Government & public agencies	Australia	Public Transport Victoria (PTV) is the system authority for public transport in Victoria, Australia. In 2019, PTV was found in breach of national privacy and data

[120] Maturity level is ranked as follows: Nascent – 1; Emerging – 2, and; Established – 3.
[121] Documentation is ranked as follows: Low-level documentation – 1; Documented (self-reported) – 2, and; Documented (independently) – 3.

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
					protection legislation after it released travel history data for 15 million smart fare cards, which would have allowed for individuals to be re-identified. A subsequent investigation revealed deficiencies in PTV's data governance and risk management processes.
Australian Institute of Health and Welfare	3	3	Government & public agencies	Australia	The Australian Institute of Health and Welfare is an independent statutory agency whose purpose is to work with states and territories to provide "reliable, regular and relevant information of Australia's health and welfare." As an information agency, AIHW relies upon strong data governance to perform its functions effectively and maintain a trusted reputation amongst its many data providers, data recipients and stakeholders. In 2014, AIHW established its Data Governance Framework.
Seattle Privacy Program	3	2	Government & public agencies	USA	The City of Seattle has established a Privacy Program in response to the privacy implications of smart city technologies and several criticisms of the city's data practices. The Privacy Program assesses how the city authorities collect, store and use data and to consider issues such as confidentiality, anonymity, archival procedures, deletion, sharing and publishing as open data. A notable feature of Seattle's privacy program is its creation of an inventory of all surveillance technologies and the preparation of

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
					Surveillance Impact Reports for new technology.
Chicago Array of Things	3	2	Government & public agencies; academic & research institutions	USA	Launched in 2016, Chicago's Array of Things (AoT) project collected real-time data on Chicago's environmental surroundings and urban activity using a network of sensor boxes mounted on light posts. A multi-stakeholder partnership, AoT was mainly led by the Argonne National Lab, but policy and oversight were driven by the City of Chicago and an Executive Committee composed of stakeholders from research institutions, universities, municipal government, and industry. Data collected from AoT was made accessible online, providing valuable information for researchers, urban planners, and the general public.
SAIL (Secure Anonymised Information Linkage) Databank	3	2	Academic & research institutions	United Kingdom	The SAIL Databank is a repository of person-based health and population records with 'data linkage and analysis toolsets' to help researchers. Researchers can access a range of data spanning up to 20 years from an entire population. An independent Information Governance Review Panel (IGRP), composed of representatives from various governmental organizations and sectors as well as the public, provides independent guidance and advice on policies, procedures and processes. The IGRP also reviews all proposals to use SAIL Databank to ensure

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
					that they are appropriate and in the public interest.
Consumer Data Research Centre	3	2	Academic & research institutions	United Kingdom	The Consumer Data Research Centre (CDRC) is an academic-led, multi-institution laboratory that brings together consumer-related datasets from around the UK and provides researchers with access to a broad range of consumer data to address many societal challenges. It uses several governance mechanisms, including an ethics review committee and a data sensitivity categorization scheme, to control access to the data sets it holds.
First Nations Data Centre	3	2	Research institution	Canada	The First Nations Data Centre is a limited access research site operated by the First Nations Information Governance Centre (FNIGC). Its purpose is to provide researcher access to individual-level data drawn from FNIGC's surveys that otherwise would not be available due to its sensitivity. The OCAP (Ownership, Control, Access and Possession) principles form the basis of its mission and researchers must adhere to them as a condition of access to their data.
Portland Smart City PDX Program	2	3	Government & public agencies	USA	The City of Portland, under its Smart City PDX program, is using sensors to understand how and when vehicles, pedestrians and bicycles use street infrastructure, monitor and analyze vehicle speeds; and track supply and demand of parking spaces to design better streets. To protect the privacy of residents, the City

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
					worked with the project vendors to ensure that photos are not saved and any information is anonymized. The plan is notable among smart city strategies for its explicit focus on marginalized and underrepresented communities.
Nantes Métropole Data Charter	2	3	Government & public agencies	France	Nantes Métropole is the first French region to have published a Data Charter articulating a set of principles relating to data produced by municipal administrations, private companies involved in the management of urban services (public transport, energy, water, waste) and private operators whose activity has an impact on the public space (such as Waze, Uber). While not a data governance program, this data charter is an essential antecedent to data governance in the region.
Argentina-Microsoft Partnership, AI Tools for Public Policy	2	3	Government & public agencies	Argentina	The Ministry of Early Childhood in the Argentinian Province of Salta entered into a public-private partnership with Microsoft to implement artificial intelligence tools using data provided by the ministry. The purpose of using these AI models was to understand better the factors contributing to school dropouts and teenage pregnancies. However, a lack of transparent communication on how data was used in the AI models fostered mistrust on the part of residents.

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
Barcelona Municipal Data Office	2	2	Government & public agencies	Spain	The City of Barcelona set up its Municipal Data Office (MDO) in 2017, based on direction from the City Council, to coordinate and support data activities across departments, as well as foster a city-wide data culture as part of the Digital City Plan. The Digital City Plan is notable for its focus on ethical digital standards and technological sovereignty, structured around three areas: the transition and use of free software, the interoperability of services and systems, and the use of free standards.
Los Angeles Department of Transportation	2	1	Government & public agencies	USA	The Los Angeles Department of Transportation (LADOT) implemented the Mobility Data Specification (MDS) to manage e-scooter and private transportation company data. Critics have raised concerns that the current MDS gives LADOT access to highly sensitive and potentially identifiable location information, which could pose significant risks for privacy and security.
Estonian Data Embassies	2	1	Government & public agencies	Estonia	Estonia is widely considered to be one of the most technologically integrated and advanced governments in the world. Due to its reliance on its ICT infrastructure, Estonia is testing what it calls “data embassies” to provide a measure of redundancy and continuity in the event of digital infrastructure failure. These are network servers which, although located outside of Estonia, are nonetheless

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
					governed by its laws. The first of these is located in Luxembourg with plans for others in the future.
Data Ventures	2	1	Government & public agencies	New Zealand	Data Ventures is a business unit of Stats NZ, New Zealand's official data agency, that functions as a trusted intermediary collecting datasets from various sectors for later re-distribution to the platform's customers. The platform collects statistical data, government data, and private sector data, such as that from telecommunications companies. Data Ventures operates under Stats NZ's social license, which it defines as the permission it has to make decisions about the management and use of the public's data, and ensuring it has the public's trust and confidence.
Liberian telecommunications authorities	1	3	Government & public agencies	Liberia	During the 2014 West African Ebola outbreak, a group of actors from the development sector called for the use of aggregated location data (Call Detail Records) collected from local cell phone towers as a means of contact tracing those who may have been exposed to disease. While many governments in West Africa agreed to release these records, the government of Liberia refused to release them due to concerns about managing and enforcing access.
NYC Automated Decision Systems Task Force	1	3	Government &	USA	New York City became one of the first jurisdictions to pass a law on automated decision systems (ADS). The ADS Task

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
			public agencies		Force was a consultative mechanism tasked with recommending a process for evaluating ADS proposed for implementation in city operations. It was concerned primarily with the most complex systems whose decisions would have the most significant impact on an individual's job prospects, financial outcomes, or similar opportunities. The Task Force final report was released in November 2019. However, an independent 'shadow report' cites a lack of transparency in the process as a significant hindrance to the work of the Task Force.
France's National Health Data Hub	1	1	Government & public agencies	France	France's National Health Data Hub is an instrument for sharing health data and securing access to it. It has been conceived as a "trusted third party" to ensure both ethical use and quality of data. It will connect data producers with public or private users, providing a one-stop-shop for all national health data. It is also intended to ensure transparency by providing a portal through which civil society and citizens can consult available data sources and their use.
Japanese Information Banks	1	1	Private sector	Japan	In Japan, information banks have a similar objective as data trusts to protect data but use a different mechanism. Through a contractual mechanism, individuals would be able to deposit their information with a trusted data intermediary, decide how the information is shared with third parties and

Name	Maturity Level[120]	Documentation[121]	Organization Type	Location	Description
					receive economic gains based on its value. A certification process for such an entity is currently being developed, and the initiative is still in the pilot phase.

4.3. Observing governance mechanisms in practice

We observed a wide variety of approaches to data governance. In this section, we highlight structures, procedures and relations that emerged through this analysis. These are grouped according to structural mechanisms, procedural mechanisms, and relational mechanisms. It is important to note that the case studies were selected to illustrate variations in the mechanisms listed; they are not intended to be an exhaustive list. They should not necessarily be considered as representative of their sector (i.e., government, research).

4.3.1. Structural mechanisms

Structural mechanisms exist to ensure a chain of accountability within a data governance program. They determine reporting structures and governance bodies, set out roles and responsibilities, and allocate decision-making authority.

Governance bodies include councils and boards, which provide strategic direction for data governance programs and align them with organizational goals. They also include data governance offices, where various functions are performed by traditional IT personnel as well as new emerging positions such as data stewards, to support the implementation of data governance. Various data governance committees can also come into play to provide guidance and to oversee compliance with policies and standards.[122]

The separation between strategic decision-making roles for the broader data governance program, and compliance monitoring (e.g., with ethics and privacy rules) for individual data use cases is important. Supporting bodies and roles are also essential to ensure that all parts of the organization are on board, and the data governance program is sustainable. In the case studies, we observed a variety of organizational structures at work, exhibiting different divisions of responsibilities.

4.3.1.1. Leadership mechanisms

Executive boards or councils were a common structural mechanism for providing overall strategic direction within a data governance program. For example, in the **Silicon Valley Regional Data Trust** (SVRDT), a Core Management Team, comprising members from participating agencies, oversees the development and approval of all core components of the

[122] Compliance monitoring practices (such as auditing, performance measurement, corrective and preventive action plans) are procedural mechanisms put in place by structural bodies. Abraham, Schneider, and vom Brocke, 430.

data-sharing system. The respective agencies and service providers provide access to one another to facilitate information sharing and comprehensive case management for minors (i.e., the beneficiaries) involved in different education, health, and social service systems.[123]

In the case of **Chicago's Array of Things**, the program was governed by an executive oversight council consisting of representation from all partners. This council was responsible for setting policy and establishing processes and procedures related to system operation, configuration, and capabilities, access to data and other resources, and communication and interactions with the City and community.

Analysis: Broad representation on a board or committee may increase the range of stakeholder interests considered and may increase the public legitimacy of the governance structure.

Accountability can manifest itself in a transparent chain of authority, with increasing power vested in specific positions. In the **Array of Things** project, for instance**,** final approval authority rested with the Commissioner of the City's Department of Innovation and Technology.[124] Similarly, the **Barcelona Municipal Data Office** has an Executive Committee of Data, which sets the high-level strategic and tactical direction. Still, the MDO is ultimately accountable to the City Council through its Chief Data Officer. A chain of authority ending with a democratically elected body may be perceived as more legitimate than if it leads to an appointed official or entity.

Analysis: Clarity around who has the final say appears to be vital to establishing accountability and trust. If a nominally representative board is perceived to have no actual power, this could affect the perceived legitimacy of the governance model.

4.3.1.2. Compliance mechanisms

Structural compliance mechanisms are more narrowly focused structures, which monitor and "enforce conformance with regulatory requirements and organizational policies, standards, procedures" relating to data protection.[125] We observed two primary forms of structural

---

[123] Allison-Jacobs, "IDS Case Study: Silicon Valley Regional Data Trust: Supporting Students through Integrated Data and Research-Practice Partnerships."
[124] University of Chicago, Argonne National Labs, and City of Chicago, "Array of Things Governance Policy and Process."
[125] Abraham, Schneider, and vom Brocke, "Data Governance," 430.

compliance mechanisms in our case studies: ethics review committees and data protection functions.

Ethics review committees most commonly exist in organizations with close ties to research and academic sectors. For example, the **Consumer Data Research Centre**'s Research Approvals Group (RAG) is responsible for reviewing and approving each project and is drawn from the UK social science academic community. The RAG operates independently from the Centre's Senior Management Teams, implying at least a degree of separation from business considerations in its decision-making.

Similarly, in the **Australian Institute of Health and Welfare**, a separate ethics committee makes substantive decisions about data use and access, which takes place under a strict legislative framework and with researchers similarly bound by institutional ethics frameworks. AIHW also identifies 'data custodians' – staff members delegated to exercise overall responsibility for a specified data collection, including the power to release the data to other bodies or individuals.[126]

Analysis: Data ethics boards, tasked with reviewing data use proposals for potential risk, may have more credibility if they remain independent from strategic and business functions of organization-wide data governance.

Data protection functions can be exercised through various models. In the EU, the GDPR introduces the duty to appoint a data protection officer (DPO) for public authorities or bodies. We found such roles in the cases of **Barcelona** and **Nantes**, in which compliance functions were separate from other data governance and management functions. In Nantes' new data advisory roles[127] are responsible for coordinating with the city's DPO concerning GDPR compliance. The city also carefully set up internal rules defining how the Chief Data Officer and the DPO work together. In Barcelona, officials went further and created a Data Protection Table with a broader mandate regarding data protection policy coordination, data protection internal training and personal data confidentiality and protection.

External compliance mechanisms also exist. In the case of **France's National Health Data Hub** - which connects data producers with public or private users - national data protection laws are

[126] Australian Institute of Health and Welfare, "AIHW Data Governance Framework 2019."
[127] *Référents data*

enforced through an external agency, the national data protection agency, *la Commission nationale de l'informatique et des libertés* (CNIL).[128]

Analysis: The GDPR has established a strong role for DPOs in data processing organizations. However, the data protection function does not stop at the DPO role – other actors have responsibilities for implementing data protection measures.

4.3.2. Procedural mechanisms

Procedural mechanisms ensure accurate, secure, and effective data collection and appropriate sharing.[129] As outlined in Section 2.1, while structural mechanisms are primarily about "who" makes the decisions, the procedural mechanisms are "how" they are made to ensure consistency, accountability and transparency. Mechanisms with reproducible steps and processes build confidence in the data governance program. Procedural mechanisms provide assurance that issues and challenges that arise can be mitigated and dealt with. We observed several varieties of these mechanisms throughout the case study examples, including risk assessment tools and processes and data access controls.

The data lifecycle is one way to operationalize procedural mechanisms and has been observed in open data implementation in Canada already. A data lifecycle is a management framework that identifies critical stages and transformations of a dataset.[130] It includes basic phases such as the planning and acquisition of data, processing and analysis, and post-analysis activities such as archiving and sharing, with variations to include other aspects such as data quality or privacy depending on the context. Below, we refer data management lifecycles when framing procedural mechanisms governing data.

4.3.2.1. Planning

Procedural mechanisms supporting the planning stages of data collection and sharing processes minimize uncertainties, facilitate coordination among stakeholders, improve the allocation of limited resources, and make it easier to assess when project objectives have been achieved. The importance of planning was underscored during our interview with the **City of Portland** Smart City PDX office, where we learned that its data governance framework was

---

[128] Cuggia and Combes, "The French Health Data Hub and the German Medical Informatics Initiatives," 430.
[129] Abraham, Schneider, and vom Brocke, "Data Governance," 430.
[130] See for example Faundeen et al., "The United States Geological Survey Science Data Lifecycle Model: U.S. Geological Survey Open-File Report 2013–1265," fig. 1.

created from a realization that proposed data collection activities were more complicated than originally anticipated.[131]

A standard planning tool in the data governance field is the risk assessment, usually in the form of privacy impact assessments (PIAs). Risk assessments - as the name suggests - are an important tool in implementing a risk-based approach to data protection.

Municipal government by-laws or ordinances often lead to the creation of policies and procedures whose aim is to control and supervise technology. For example, in 2017, the **City of Seattle** introduced an ordinance outlining a range of procedures designed to increase transparency around the city's use of surveillance technologies. Under that policy, before the council approving a surveillance technology, the relevant department must produce a Surveillance Impact Report on its privacy implications. An initial report is prepared by City staff, who then host one or more public meetings to receive feedback. Similarly, the **City of Portland's** Smart City PDX Priorities Framework stipulates that any policy, plan or project receiving Smart City PDX support must provide a detailed PIA.

At the national level, we find a similar approach in **Ontario's Smart Energy Metering Program,** which adopted a Third Party Access Implementation Plan that anticipated using Privacy Analytics Inc.'s specialized software to conduct a risk assessment with each request for data. These assessments consider the context and intended use of the data in evaluating the re-identification risk.

Another example of a risk-based guideline is the Methodology for Privacy Risk Management, which was published by France's data protection authority CNIL as guidance for **France's** Data Protection Act.[132] The Methodology is intended to help data controller stakeholders improve their data processing practices. The CNIL methodology is based on five factors: context, feared events, threats, risks, and measures.

- Context includes the main regulatory guidelines and the benefits that data processing offers;
- Feared events include illegitimate access to personal data, unwanted change in personal data, the disappearance of personal data, and unavailability of legal processes;
- Threats include function creep, espionage, theft, and damage;
- Risks are assessed according to severity and likelihood, and;

---

[131] Kevin Martin (Smart City PDX Manager at City of Portland), in discussion with Open North, October 2019.

[132] Commission Nationale de l'Informatique et des Libertés, "Methodology for Privacy Risk Management: How to Implement the Data Protection Act."

- Measures are used to treat risks in a proportionate manner.

The **Australian Institute of Health and Welfare** uses a risk assessment framework - the Five Safes framework - to ensure data access results in "safe people, safe projects, safe settings, safe data and safe outputs."[133] AIHW separately evaluates its data linkage, confidentialization, data security and data access and release practices, then assesses them all together to determine the overall safety of the data collection or sharing project.

However, having a risk assessment process in place does not always ensure that risks will be avoided altogether. **Public Transport Victoria** – which mainly serves the metropolitan Melbourne area – also used the Five Safes framework. However, following a data breach involving over 15 million 'Myki' smart cards, an investigation concluded that a flawed PIA was a contributing factor.[134] In effect, the PIA form in use encouraged staff to make a judgment without detailed justification about whether or not the data in question appeared to be personal information. In this case, it resulted in an ill-considered approval for data release. Furthermore, it was concluded that the completed PIA document was treated as an authorizing document for data release, rather than as it was intended: as a tool for helping the organization to identify and address privacy risks.

On the other hand, we also found a case in which the sale and release of sensitive data were preemptively halted based solely on the absence of public trust. A recent decision by the Ontario Energy Board denied the Smart Metering Entity – which oversees their **Smart Energy Metering Program** – the license to sell de-identified data to third parties. This decision was reached because it was "not clear from the evidence that consumers support the notion that consumption data (even if de-identified) should be offered for sale to third parties."[135]

Freedom of Information (FOI) legislation, which exists in almost every jurisdiction, also needs to be accounted for in privacy impact assessments. In the case of **Portland**, staff expressed concerns that data collected would be subject to existing FOI legislation – which is particularly strong in the State of Oregon and could be subsequently exploited by a bad actor towards public harm.[136] The concern was that, if people could be reidentified through sensed data, it

---

[133] Australian Institute of Health and Welfare, "The Five Safes Framework."
[134] State of Victoria (Australia). Office of the Victorian Information Commissioner, "Disclosure of Myki Travel Information: Investigation under Section 8C(2)(e) of the Privacy and Data Protection Act 2014 (Vic)."
[135] Ontario Energy Board, Independent Electricity System Operator (in its capacity as the Smart Metering Entity): Application for approval to provide access to certain non-personal data to third parties at market prices at 14.
[136] Lempert, "Shared Mobility Data Sharing: Opportunities for Public-Private Partnerships," 8.

would have the effect of violating their right to privacy. As we learned from our interview with a staff member of the City of Portland's smart city office,[137] the City's proposed approach for the future, in response to these concerns, was not to collect sensitive information in the first place.

In Ontario, MFIPPA sets out both the FOI obligations as well as privacy protection obligations to which municipal governments and bodies are subject. Data collected in the course of providing regular services may be subject to access to information requests (under MFIPPA). Existing risk assessments under current legislation need to evolve to keep pace with the changing legislative and technological environment.

Analysis: Risk assessment processes may be more useful if they are used to support broader accountability processes and encourage careful reflection of the intended uses and potential consequences, rather than serving as perfunctory checklists.

4.3.2.2. Data acquisition

Municipal officials face the dilemma of making sure that their data collection activities – especially where sensitive data is concerned – balance existing legislative requirements, specific planning purposes, and public interest objectives. Considerations around privacy and informed consent should be at the heart of data collection decisions.

Determining how much sensitive data to acquire is challenging, from both a usefulness and privacy protection standpoint. We observed this in the implementation of the **Los Angeles Department of Transportation** (LADOT) Dockless Mobility Program, in which service providers were required to provide their data - including the precise location of each vehicle - to LADOT using the Mobility Data Specification (MDS). This requirement prompted critics to raise questions about whether this level of granularity was required for planning purposes and, more importantly, how the privacy and security of this data would be maintained once in LADOT's possession.[138]

Anonymization, de-identification, and obfuscation are often cited as privacy-enhancing techniques (PETs), which strip away the personally-identifiable elements within a data set. However, even where data is de-identified, significant privacy risks remain. For example, a 2019 study showed that anonymized data could successfully be re-identified and associated with an

---

[137] Kevin Martin (Smart City PDX Manager at City of Portland), in discussion with Open North, October 2019.

[138] Electronic Frontier Foundation and Open Technology Institute to Los Angeles City Council and Los Angeles Department of Transportation, "Urgent Concerns Regarding the Lack of Privacy Protections for Sensitive Personal Data Collected Via LADOT's Mobility Data Specification," April 3, 2019.

identifiable individual at a rate of 99.98 percent using fifteen demographic factors.[139] An earlier study found that, in a dataset of 1.5 million people collected over six months using location points triangulated from cellphone towers, 95 percent of individuals could be uniquely identified based on just four time-stamped and geo-located points.[140]

As noted in a discussion paper from the Office of the Privacy Commissioner of Canada, it can be challenging to quantify the potential for re-identification of individuals because:

> *The purpose of big data algorithms is to draw correlations between individual pieces of data. While each disparate piece of data on its own may be non-personal, by amassing, combining and analyzing the pieces, the processing of non-personal information may result in information about an identifiable individual. Big data analytics has the ability to reconstitute identities that have been stripped away. It is difficult if not impossible to know in advance when an algorithm will re-identify an individual or what pieces of data will allow it to do so.*[141]

The above quotation highlights the tension between protecting privacy and maintaining the usefulness of data. The addition of interests other than the public interest - a profit motive, for example - may create incentives for an organization that has expended resources to collect data to maximize its usefulness at the expense of privacy.[142]

Analysis: While privacy-enhancing techniques may reduce the risk of re-identification of individuals, they do not eliminate it. The collection of sensitive data should be limited to minimize privacy risks.

4.3.2.3. Data security and access controls

The main objective of data security policies and procedures is to ensure that data is protected across all of its forms and storage media, throughout every phase of its lifecycle from

---

[139] Rocher, Hendrickx, and de Montjoye, “Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models.”
[140] de Montjoye et al., “Unique in the Crowd.”
[141] Office of the Privacy Commissioner of Canada, Policy and Research Group, “Consent and Privacy: A Discussion Paper Exploring Potential Enhancements to Consent Under the Personal Information Protection and Electronic Documents Act,” 7.
[142] Grieman, “Pedestrian Curiosity: A Brief Examination of Consent and Privacy in Swath Section Smart City Spaces,” 2.

unauthorized or inappropriate access, to usage, modification, disclosure, and destruction.[143] As more and more personal data is shared among various stakeholders, ensuring both physical and computer security has become critical. We observed different procedural mechanisms deployed to secure personal data throughout its lifecycle. These included limiting access to sensitive data through review panels, role-based credentials, and metadata.

Review panels can control data access, such as in the case of the Secure Anonymised Information Linkage **(SAIL) Databank** in the United Kingdom. The SAIL Databank is a repository of anonymized person-based health and population records with "data linkage and analysis tool sets" to support research. Researchers can access a range of data spanning up to 20 years from an entire population.[144] An independent Information Governance Review Panel (IGRP), composed of representatives from various governmental organizations and sectors as well as the public, provides independent guidance and advice on policies, procedures and processes. The IGRP also reviews all proposals to use SAIL Databank to ensure that they are appropriate and in the public interest. The Panel assesses each request to use the data to ensure they are compliant with the Information Governance policy.

Moreover, once a researcher has completed their analysis using data from the SAIL Databank, they can only remove their results from the secured platform following scrutiny by a SAIL Data Guardian to ensure that any risk of disclosure has been mitigated.[145] However, since review panel determinations are based on human judgment, they may only be a solution for relatively small numbers of data access requests. For high-volume, near-real-time queries, secure digital protocols and policies are needed.

One way to implement secure digital protocols and data access policies is through agreements or memoranda of understanding. In the case of **Chicago's Array of Things**, all individuals with access to its data were subject to strict contractual confidentiality obligations and subject to discipline or termination if they failed to meet these obligations. The **Silicon Valley Regional Data Trust** maintains a Secure Data Environment (SDE) - an application platform that allows certified users to query access to case management data across agencies. An Enterprise MOU sets out several data access provisions for the SDE, including what data will be shared, determining the level of access based on role, and the process for credentialing users.[146] Such

[143] Cloud Security Alliance, "Security Guidance for Critical Areas of Focus in Cloud Computing V. 3.0."
[144] SAIL Databank, "SAIL Databank - The Secure Anonymised Information Linkage Databank."
[145] SAIL Databank, "Data Privacy & Security: Ensuring Secure Access to the Data."
[146] Allison-Jacobs, "IDS Case Study: Silicon Valley Regional Data Trust: Supporting Students through Integrated Data and Research-Practice Partnerships."

an agreement is a potentially reproducible, technologically agnostic framework that could be used in other situations where there is data sharing between different agencies.

Data classification schemes describe the degree of sensitivity associated with particular data sets and can provide an additional layer of security. The **Consumer Research Data Centre** in the UK uses metadata to classify data according to three levels of sensitivity: open, safeguarded or controlled.[147] Open data is freely available to all for any purpose. Safeguarded data has restricted access because of license conditions, but where data are not considered personally identifiable information (PII or otherwise sensitive). Controlled data needs to be held under the most secure conditions with highly restricted access. This approach aligns with the Fair Information Principle of purpose limitation; access is granted to researchers only when appropriate justification is provided.

Analysis: Sensitive data can be protected using several procedural mechanisms, including approvals processes, assignment of role-based credentials, and classification of data by sensitivity level. Determining the appropriate mechanisms will depend on the nature and volume of data access requests.

4.3.2.4. Data storage

The issue of data storage must also be considered in a global context in which organizations increasingly rely on third-party, cloud-based data infrastructure and software platforms, potentially leaving their data exposed to novel risks.[148] For this reason, data residency – referring to the physical or geographical location of an organization's digital information while at rest – emerges as an essential consideration in their data governance programs. [149] We saw data residency incorporated as a key consideration in several cases.

Knowing what and where data is stored is a critical initial step in assessing exposure to data residency-related risks. However, undertaking a complete inventory of all public data infrastructure is a significant task. Municipal officials in **Nantes**, France conducted an audit of data storage devices (e.g., servers) used for its applications and determined that the majority were in France or the EU (both covered under the GDPR). Only one software application exported data outside the EU; however, as the GDPR's provisions apply extraterritorially if the data concerns EU citizens, this was not deemed to be a risk. **Estonia** has taken this concept further in their **data embassies** – which are servers located out-of-country (in Luxembourg),

[147] Consumer Data Research Centre, "About Our Data."
[148] Bohaker et al., "Seeing Through the Cloud."
[149] Canada. Treasury Board of Canada Secretariat, "Government of Canada White Paper."

granted diplomatic status (i.e. subject to Estonian law). By hosting redundant data in networked servers outside its borders but still within the EU, the Estonian government is attempting to mitigate risks associated with equipment failures or cybersecurity attacks on a centralized infrastructure.

Indigenous data sovereignty – exemplified in the OCAP principles described in Section 3.2.2 – has been operationalized at the **First Nations Data Centre** (FNDC), operated by the First Nations Information Governance Centre (FNIGC). FNDC is a secure research site whose purpose is to provide researcher access to individual-level data drawn from FNIGC's surveys that otherwise would not be available due to its sensitivity. Researchers must be physically present on-site to access data, and all research outputs must be vetted before being taken out of the facility.[150]

Data sovereignty and localization[151] may be challenging to achieve in the broader Canadian context as data flows can still be routed through extraterritorial jurisdictions, especially when retrieving data hosted by third parties outside the country.[152] Moreover, the US government, via the US Foreign Intelligence Surveillance Act (FISA), can "compel an organization subject to US law to turn over data under its control, regardless of the data's location and without notifying Canada." [153] Furthermore, Article 19.12 of the United States–Mexico–Canada Agreement (USMCA) restricts the ability of Party countries to legislate data localization as a condition for entities wishing to do business.[154]

Analysis: Data sovereignty – while a promising approach to mitigating data security risks – may be challenging to fully implement in the context of current international trade agreements.

4.3.2.5. Data sharing and publishing

Data sharing between parties generally requires a contractual agreement. In Japan, **information banks** function as trusted data intermediaries (certified by an industry group) with

---

[150] First Nations Information Governance Centre, "Data Access at the First Nations Data Centre | FNIGC."
[151] Data sovereignty refers to data not only being stored in a particular location, but also being subject to the laws of the country in which it is stored. Data localization refers to a requirement that data created within a country's borders stay within them.
[152] Bohaker et al., "Seeing Through the Cloud," 2.
[153] Canada. Treasury Board of Canada Secretariat, "Government of Canada White Paper."
[154] "No Party shall require a covered person to use or locate computing facilities in that Party's territory as a condition for conducting business in that territory." Agreement between the United States of America, the United Mexican States, and Canada 12/13/19 Text [USMCA] Article 19.12.

whom individuals can enter into a contractual agreement and 'deposit' their personal information. The information bank then sells their data to private sector companies, generating dividends for clients. For example, customers can choose to share their consumer preferences or credit scores with third-party companies so that those companies can better target their advertising.[155] This model is predicated on the notion of giving people a degree of control over their data, though this control does not constitute an ownership right. Information banks have a fiduciary duty to their customers and would assume liability in the event of a data breach.[156]

Analysis: Cities need to keep looking for ways to ensure personal data is protected in the context of public-private partnerships.

4.3.3. Relational mechanisms

Relational mechanisms are strategies or practices that facilitate collaboration between stakeholders. They encompass communication and training, as well as collaborative approaches that leverage formal and informal coordination mechanisms for decision-making.[157]

We observed the presence of various relational mechanisms - including communication, education, and stakeholder engagement - as some of the most crucial elements in setting up a data governance framework in cases such as the **Silicon Valley Regional Data Trust (SVRDT), Chicago's Array of Things,** and **New York City's Automated Decision Systems Task Force**. While both structural and procedural mechanisms are essential, these communicative practices play an ongoing role in building a robust, ethical data culture within and between organizations.

4.3.3.1. Communication and education

Building capacity with internal stakeholders is an important step, both in the early stages of a data governance program as well as throughout its lifecycle. For example, establishing the **Silicon Valley Regional Data Trust (SVRDT)** relied heavily on the efforts of champions (several retired individuals in this case) to build the necessary level of trust between stakeholders, navigating legal issues, and developing systems for secure data sharing.[158] In

---

[155] Heimi, "Japan's 'information Banks' to Let Users Cash in on Personal Data."
[156] The Asahi Shimbun, "Seeking a Bite of Big Data, Banks to Act as Brokers of Personal Info."
[157] These include "working groups, committees, task forces, and integrator roles, but also informal coordination mechanisms such as interdepartmental events, performance reviews across business units, and job rotation". Abraham, Schneider and vom Brocke, 430.
[158] Moore and Reeves Bracco, "Scaling Goodwill: The Challenges of Implementing Robust Education Data Sharing Through Regional Partnerships," 5.

Barcelona, the MDO saw its role as fostering an internal data culture across municipal departments, in addition to providing support on specific projects. Similarly, in the **City of Seattle** designated departmental 'Privacy Champions' across different agencies handle basic inquiries, oversee low-risk privacy assessments, and escalate issues to a manager whose role is to coordinate and develop a community of practice among these champions.[159]

External capacity building in the form of public education can also increase trust in smart cities initiatives. Chicago's **Array of Things** (AoT) program operators maintained a public website with current information on the project, including educational materials regarding the hardware and software technologies and capabilities associated with AoT, a directory with detailed information on all components, experiments, and projects supported by AoT, all policies and procedures for AoT operation, governance body meeting minutes, and reports. The operators also noted that, while some members of the public called for less technical language in AoT policies, others called for more technical details. Therefore, it may be challenging to strike the right balance between complete transparency while communicating in an accessible manner when dealing with complex technological systems.

This tension is especially relevant when we talk about artificial intelligence models or algorithmic decision systems that many public administrations are beginning to experiment with, such as in the **Province of Salta, Argentina**. There, officials partnered with Microsoft to implement artificial intelligence tools to research factors contributing to school dropouts and teenage pregnancies to identify at-risk youth and coordinate interventions with the appropriate Social Service Agency.[160] The artificial intelligence models used data provided by the Ministry of Early Childhood, which was collected from predominantly low-income areas in the city of Salta in 2016 and 2017. However, as the collection methodology for the training data was not made publicly available, it was difficult to evaluate the validity of the model adequately.[161] Consequently, researchers at the University of Buenos Aires raised concerns that the initiative may have led to predicting disproportionately higher numbers of cases of teen pregnancy or school dropouts compared to other groups or areas. This is especially concerning considering that the Ministry "does not collect or consolidate information on the impact of these tools."[162][163]

There is a real risk of bias and discrimination that may enter systems via inputs from human developers. Introduced in Section 3.1.1., open-source software and standards could play a role

---

[159] Armbruster, "The City of Seattle Privacy Program."
[160] Freuler and Iglesias, "Algorithms and Artificial Intelligence in Latin America: A Study of Implementation by Governments in Argentina and Uruguay," 17.
[161] Freuler and Iglesias, 19.
[162] Freuler and Iglesias, 22.
[163] Freuler and Iglesias, 18.

in being able to identify potential bias early, as the source code would be open for outside researchers and experts to 'look under the hood.' Without intellectual property considerations as barriers, it may be easier to explain to the public how a given input leads to a given output, thereby increasing confidence that AI models are being used to support ethical purposes and sound policy objectives. Overall, this case suggests that openness is key to fostering public trust in algorithmic models used for public policy.

Analysis: Fostering data literacy and building capacity for meaningful participation, among internal and external stakeholders and the public, appears to be essential in public-sector data governance

4.3.3.2. Coordination

Coordinating functions to ensure alignment across functions can occur in a top-down or a horizontal, cooperative manner. While more hierarchical coordination is focused on steering and control, more horizontal models make use of collaborative behaviours to clarify differences and solve problems. These can use formal mechanisms (e.g., task groups) or informal mechanisms (departmental 'privacy champions,' learning opportunities).

In **Barcelona's** case, a municipal directive served as the foundation for the establishment of several data governance bodies – including the Municipal Data Office (MDO) mentioned above – and the division of powers between them. In our interview with MDO Program Manager,[164] we learned that it acts as a service unit,[165] providing support across the data lifecycle for all other departments, such as data structuration and analysis. However, when the MDO was established, its staff complement was limited to the existing positions, with no provision for hiring new staff with specialized data skill sets.

In contrast to this centralized approach, other cities embedded roles within their organization to ensure that the strategic vision was understood and communicated. For instance, in **Seattle** "privacy champions" act as a point of contact for the City's Privacy Program within each department, "cultivating a community of practice to share knowledge and best practices."[166] According to a recent annual report, the City was preparing to expand the privacy champions

[164] Pau Balcells (Program Manager, City of Barcelona Municipal Data Office), in discussion with Open North, December 2019.
[165] Barcelona City Council, "Directive Concerning Municipal Data Governance and the Municipal Data Offices."
[166] Bass, Sutherland, and Symons, "Reclaiming the Smart City: Personal Data, Trust and the New Commons," 23.

program by adding more external speakers to leverage the program city-wide.[167] In **Nantes**, new data advisory roles were created in all the city departments who communicate internal and external governance directives throughout the organization, as well as support data management.[168]

Analysis: Empowering key individuals who provide technical support and cultivate organizational culture can lead to more internal alignment on implementation of data governance

4.3.3.3. Stakeholder engagement

Consultation with external stakeholders should also take place proactively. In the case of the **Chicago Array of Things**, draft governance and privacy policies were created in 2015 and reviewed by privacy, technology, and legal experts in early 2016.[169] Later, the project leaders partnered with the Smart Chicago Collaborative to organize public meetings and promote interaction using an online "policy co-creation" platform. During the implementation phase, officials conducted outreach in individual neighbourhoods to seek resident approval in situating the sensor clusters.[170] Consultation processes are more than an opportunity for residents to learn from and ask questions of city staff; they are also a critical point at which the community can grant or refuse social license for the project. For example, communities that have historically experienced high levels of policing may not be receptive to data collection activities (e.g., gunshot-detection technologies), which may result in further stigmatization.[171] Municipal staff should, therefore, be aware of important community history and context before entering a community to seek social license for new data collection.

Data Ventures has selected a business model that allows them to leverage their "social license" – the permission it has to make decisions about the management and use of public data held by Stats NZ.[172] This social license allows them to position themselves as a trusted data

---

[167] City of Seattle, "City of Seattle Privacy Program: 2019 Annual Report: Transforming Privacy."
[168] Nathalie Hopp (Directrice générale adjointe, Nantes Métropole), in discussion with CIVITEO, October 2019.
[169] University of Chicago, Argonne National Labs, and City of Chicago, "Array of Things Governance Policy and Process."
[170] Smart Chicago Collaborative, "Array of Things Civic Engagement Report: A Summary of Public Feedback & the Civic Engagement Process."
[171] Electronic Frontier Foundation, "Acoustic Gunshot Detection."
[172] Nielsen, "A Social License Approach to Trust - A Close-Up on Trust."

intermediary and extract public value from otherwise privately-held data sets (e.g., cell tower location data).

Engagement with a broad range of stakeholders is vital. In **Seattle's** case, governance was created in collaboration with community activist groups, leaders from academia, local companies, and private legal practice as a response to public concerns relating to a pilot program for police body-worn cameras in 2014.[173] To ensure community representation, **Portland** officials issued a call for two to three 'Equity Consulting Advisors' having lived or leadership experience in marginalized communities, who would be paid to work with them and act as liaisons with the local community.[174] Unlike with traditional public consultations, in which anyone can show up, it is important to make sure that the selected individuals truly represent their neighbourhood interests, can legitimately speak on behalf of other residents and do not have a conflict of interest.

Analysis: Plan on investing significant time into developing the relationships, trust, and social license necessary for successful data governance - ideally before project implementation.

In order to address discrimination and bias resulting from the use of automated decision systems (ADS), **New York City**'s City Council passed Local Law 49, which called for officials to form an ADS Task Force. Its purpose was to recommend a process for reviewing the use of algorithms in City agencies and departments. The Task Force was chaired by municipal officials and included representatives from the private sector, nonprofit, advocacy, and research communities.[175] However, while the Task Force was given a mandate by law, they were unable to gain access to information they needed - namely, a list of ADS currently in use.[176] As such, some members felt their recommendations reflected only the dominant viewpoint and left out dissenting voices. Ultimately, rather than making specific policy changes, the report offered only broad recommendations, which some members felt lacked force and could have been produced without convening a task force.[177] This case illustrates that, while stakeholder engagement is important, if carried out poorly it can lead to reduced trust.

---

[173] Armbruster, "How Seattle Is Tackling Privacy Problems in Today's Digital Age."
[174] City of Portland, "RFQ: Equity Consulting Advisors."
[175] New York City Automated Decision Systems (ADS) Task Force, "New York City Automated Decision Systems Task Force Report."
[176] "Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force."
[177] Lecher, "NYC's Algorithm Task Force Was 'a Waste,' Member Says."

Analysis: Without a specific mandate backed up by high-level support, task forces or working groups may lack the tools to implement their coordination role effectively.

Even with public engagement or using city planners as neighbourhood liaisons in developing a community engagement plan for a sensor deployment project,[178] defining the acceptable boundaries for data practices is a challenge. Since the norms around data collection are subjective and dynamic, it may not necessarily be achievable to establish these boundaries before a given smart city technology is implemented.

These boundaries are influenced by public opinion, which in turn reflects broader discourses on privacy and surveillance. It may be challenging to anticipate use cases, risks, and public sentiment associated with specific data collection activities. These variables may change over time.

The **City of Portland** chose to switch off the microphones embedded in the GE CityIQ sensors installed on 200 streetlights while leaving video recording turned on (though no video was to be retained). Our conversation with the City of Portland's Smart City PDX Manager confirmed that public opinion influenced the City of Portland's decision to turn off microphones. In this case, public opinion may have helped shape the perceived acceptability and boundaries of surveillance. Specific technologies – such as facial recognition – may be viewed as socially unacceptable. Portland is currently exploring an outright ban on this technology.[179] This also appears to be in reaction to public opinion rather than foresight or planning. In other US cities, facial recognition technology is experiencing similar public backlash (e.g., San Francisco and Oakland, California, and Somerville, Massachusetts).[180]

Planning and public engagement processes themselves may not be adequate for anticipating all issues. The full array of end uses may not be identified at the start of implementation, and public opinion may change once the project is operational. It may thus be impossible to identify the issues and challenges that a city may face until after implementation has started. In Portland, operators made the decision to turn off a specific hardware feature after deployment, suggesting that their purpose and the evaluated risk were unclear and potentially unknowable during planning and procurement. It may be difficult to judge which technologies (or specific use cases)

---

[178] Kendrick and Rodgers, "Recommendations for the Development and Implementation of Distributed Sensor Networks."
[179] City of Portland, Oregon, "What Is Facial Recognition and Why Is the City of Portland Trying to Regulate It?"
[180] Jee, "A Facial Recognition Ban Is Coming to the US, Says an AI Policy Advisor."

the public is comfortable with until they are encountered in real life, and opinions may change after a highly-visible incident brings mainstream media attention to the issue.

Analysis: Establishing multi-stakeholder feedback loops for continuous learning should help anticipate data governance needs, challenges and concerns.

4.4. Conclusion

Through our case study analysis, we observed how organizations used a variety of structural, procedural, and relational mechanisms to unlock the value of their data while minimizing risk. The analysis of structural mechanisms showed that trust, representativeness and accountability are at the center of data governance and are supported by ethical and other compliance instruments. Many procedural mechanisms we were able to identify had strong limits, reflecting a similar need to embed them in principled governance. As for relational mechanisms, they illustrated how stakeholders' capacity building and engagement are an integral part of data governance.

While the framework proposed by Abraham, Schneider, and vom Brocke was useful in providing a high-level taxonomy of data governance mechanisms, we noticed several limitations. First, it was challenging to separate real-world structures, procedures, and relations from one another. It is perhaps more helpful to think of the different mechanisms as lenses through which to view governance practices, rather than discrete categories. It was also a challenge to disentangle data governance from data management practices, and difficult to accurately assess the latter without fully understanding their technical details. Second, while impacts are the final component of this conceptual framework, tracing effects back to specific mechanisms or antecedents – especially from an outside vantage point – is a difficult undertaking.

This exercise pointed to the value for organizations – especially municipal governments – to evaluate and objectively document their efforts in implementing data governance programs. Forming a community of practice of cities facing similar issues and challenges, and reflecting on each other's successes and failures, will help to develop more effective data governance mechanisms.

5. Key considerations and next steps

As digital infrastructure becomes a part of our cities at an unprecedented scale, the City of Toronto will increasingly have to reckon with the potential risks and impacts involving the data that accompanies it. In this report, we explored a conceptual framework that we used as a lens into examples of data governance and introduced some relevant schools of thought in current data governance discourse. We examined legislation and policy that serve as the basis for data governance practices in Ontario and detailed our observations of various data governance mechanisms identified in our scan of 20 case studies. We conclude here with some key considerations derived from our research, and suggestions on future research and next steps for the City of Toronto.

5.1. Key considerations

Define a clear set of guiding values for data governance

- The public interest needs to be clearly articulated as a starting point.
- Privacy and commercial values may conflict, especially in the case of public-private partnerships.
- Further research is needed to document outcomes and successful resolutions of competing interests.

Lead with governance, not technology

- Digital technologies are a means of implementing a governance model, but should not be confused with the model itself.
- Governance of data within a municipality remains imbricated with internal relationships and political considerations

Build trust and social license through collaboration and transparent communication

- Ongoing public engagement can increase the perceived legitimacy of a data governance model.
- Plan on investing significant time into developing the relationships, trust, and social license necessary for a successful data governance model - ideally before project implementation.
- Clear communication of mandates, combined with political support, can help ensure that a mandate given to any trusted body can be implemented.

Anticipate new risks for individuals created by new data sources

- Consider risk assessment protocols, which start from the assumption that all data is potentially personally identifiable. Risk assessment ideally results in a careful and deliberate reflection of the intended use(s) of data, and potential consequences.
- Limit data collection and provide a clear rationale for data that are collected.

5.2. Future research and next steps

Uncertainties remain regarding potential data use cases, the evolution of local, national, and international discourses and legislative context around critical issues such as privacy, and future implementations of future smart city projects.

For this reason, a flexible approach should be taken by the City of Toronto as it develops its data governance framework. We suggest a set of activities that may aid in research and internal alignment and support external collaboration on data governance.

Engage internal stakeholders

- As we have outlined in this report, different stakeholders may have differing priorities and needs around the governance of data. Internal engagement, through quantitative and qualitative research methods, can help the City understand the needs, test knowledge, and solicit opinions of internal stakeholders.
- Development of a shared vocabulary for understanding issues among internal City of Toronto stakeholders can facilitate further collaboration.

Develop a research agenda

- Target specific data governance themes in future case study research. This study represents an initial exploration and snapshot of examples gathered from a broad set of requirements and global scope - future research would benefit from methods scoped from research questions.
- Track the evolution of case studies outlined in this report. However, be aware that the outcomes of data governance models may be difficult to capture.
- Engage outside academic partners to undertake empirical, longitudinal research into specific issue areas of concern within data governance in Toronto.

Establish and reinforce feedback loops

- The City of Toronto already engages external stakeholders through multiple venues. Coordination on questions and data collection across the City's various engagements with multi-stakeholder groups (e.g. through consultations, projects, events) can aid in researching external perceptions and needs for data governance.
- Use existing channels of engagement to triangulate public perceptions of legitimacy and trust around data practices in a given project.

Stay connected to national and international conversations

- Active participation in networks and communities of practice will allow the City to respond more quickly to local demands and issues that emerge from evolving global conversations.

References

Abraham, Rene, Johannes Schneider, and Jan vom Brocke. “Data Governance: A Conceptual Framework, Structured Review, and Research Agenda.” *International Journal of Information Management* 49 (December 2019): 424–38. https://doi.org/10.1016/j.ijinfomgt.2019.07.008.

Allison-Jacobs, Rosalyn. “IDS Case Study: Silicon Valley Regional Data Trust: Supporting Students through Integrated Data and Research-Practice Partnerships.” Actionable Intelligence for Social Policy, December 2018. https://www.aisp.upenn.edu/wp-content/uploads/2018/12/SVRDT-.pdf.

Armbruster, Ginger. “How Seattle Is Tackling Privacy Problems in Today’s Digital Age.” Government Technology, September 11, 2018. https://www.govtech.com/opinion/How-Seattle-Is-Tackling-Privacy-Problems-in-Todays-Digital-Age.html.

———. “The City of Seattle Privacy Program.” PowerPoint, 2015. https://www.seattle.gov/Documents/Departments/InformationTechnology/privacy/PrivacyProgramIntroductionE-TeamBriefing.pdf.

Australian Institute of Health and Welfare. “AIHW Data Governance Framework 2019.” Australian Government, 2019. https://www.aihw.gov.au/getmedia/a10b8148-ef65-4c37-945a-bb3effaa96e3/AIHW-Data-Governance-Framework.pdf.aspx.

———. “The Five Safes Framework.” Australian Institute of Health and Welfare, December 3, 2019. https://www.aihw.gov.au/about-our-data/data-governance/the-five-safes-framework.

Bannerman, Sara, and Angela Orasch. “Privacy and Smart Cities: A Canadian Survey.” Report for the Office of the Privacy Commissioner of Canada (OPC). McMaster University, January 2019. https://smartcityprivacy.ca/wp-content/uploads/2019/01/Bannerman-Orasch-Privacy-and-Smart-Cities-A-Canadian-Survey-v1-2019.pdf.

Barcelona City Council. “Directive Concerning Municipal Data Governance and the Municipal Data Offices.” Ajuntament de Barcelona, April 18, 2018. https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/108746/2/GM_Circular_OMD_2018.pdf.

Barcelona City Council. Office for Technology and Digital Innovation. “Barcelona City Council Technological Sovereignty Guide.” Ajuntament de Barcelona, September 2017. https://ajuntament.barcelona.cat/digital/sites/default/files/guia_adt_4_guia_sobre_sobirania_tecnologica_en_2017_af_9en_2.pdf.

Bass, Theo, Emma Sutherland, and Tom Symons. “Reclaiming the Smart City: Personal Data, Trust and the New Commons,” July 2018.

https://decodeproject.eu/publications/reclaiming-smart-city-personal-data-trust-and-new-commons.

Bennett, Colin J., and Charles D. Raab. “Revisiting the Governance of Privacy: Contemporary Policy Instruments in Global Perspective: Revisiting the Governance of Privacy.” *Regulation & Governance*, September 27, 2018. https://doi.org/10.1111/rego.12222.

Bernholz, Lucy. “Workshop Summary: Trusted Data Intermediaries.” Digital Civil Society Lab at the Stanford Center on Philanthropy and Civil Society, December 2016. https://pacscenter.stanford.edu/wp-content/uploads/2018/05/TDI-Workshop-Summary.pdf.

Bohaker, Heidi, Lisa M. Austin, Austin Clement, and Stephanie Perrin. “Seeing Through the Cloud: National Jurisdiction and Location of Data, Servers, and Networks Still Matter in a Digitally Interconnected World.” The University of Toronto, 2015. http://ecommoutsourcing.ischool.utoronto.ca/wp-content/uploads/BohakerAustinClementPerrin_SeeingThroughTheCloud-PublicReport-15Sept2015.pdf.

Bria, Francesca, and Malcolm Bain. “Manifesto in Favour of Technological Sovereignty and Digital Rights for Cities :: Ethical Digital Standards.” Accessed June 14, 2020. https://www.barcelona.cat/digitalstandards/manifesto/0.2/.

Bruhn, Jodi. “Identifying Useful Approaches to the Governance of Indigenous Data.” *International Indigenous Policy Journal* 5, no. 2 (April 7, 2014). https://doi.org/10.18584/iipj.2014.5.2.5.

California. Department of Justice, Office of the Attorney General. “California Consumer Privacy Act (CCPA) Fact Sheet.” Accessed December 9, 2019. https://oag.ca.gov/system/files/attachments/press_releases/CCPA%20Fact%20Sheet%20%2800000002%29.pdf.

Canada. Department of Justice. “Modernizing Canada’s Privacy Act,” August 20, 2019. https://www.justice.gc.ca/eng/csj-sjc/pa-lprp/modern.html.

Canada. Innovation, Science and Economic Development Canada. “Strengthening Privacy for the Digital Age: Proposals to Modernize the Personal Information Protection and Electronic Documents Act.” Innovation for a Better Canada. Innovation, Science and Economic Development Canada, May 21, 2019. https://www.ic.gc.ca/eic/site/062.nsf/eng/h_00107.html.

Canada, Office of the Privacy Commissioner of. “A Data Privacy Day Conversation with Canada’s Privacy Commissioner,” February 5, 2020. https://www.priv.gc.ca/en/opc-news/speeches/2020/sp-d_20200128/.

Canada. Treasury Board of Canada Secretariat. “Government of Canada White Paper: Data Sovereignty and Public Cloud.” Report on plans and priorities, June 25, 2018.

https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/cloud-services/gc-white-paper-data-sovereignty-public-cloud.html.

Centre for Information Policy Leadership. “A Risk-Based Approach to Privacy: Improving Effectiveness in Practice,” June 19, 2014. https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/white_paper_1-a_risk_based_approach_to_privacy_improving_effectiveness_in_practice.pdf.

Chen, Min, David Ebert, Hans Hagen, Robert S. Laramee, Robert van Liere, Kwan-Liu Ma, William Ribarsky, Gerik Scheuermann, and Deborah Silver. “Data, Information, and Knowledge in Visualization.” *IEEE Computer Graphics and Applications* 29, no. 1 (January 2009): 12–19. https://doi.org/10.1109/MCG.2009.6.

Cities Coalition for Digital Rights. “Declaration of Cities Coalition for Digital Rights.” Cities for Digital Rights. Accessed December 9, 2019. https://citiesfordigitalrights.org/home.

City of Portland. “RFQ: Equity Consulting Advisors,” 2019. https://static1.squarespace.com/static/5967c18bff7c50a0244ff42c/t/5d3f7848d7cba60001942a6e/1564440648818/RFQ+Equity+Consulting+Advisors+for+SC+PDX+Submissions+Due+Aug+15+2019.pdf.

City of Portland, Oregon. “What Is Facial Recognition and Why Is the City of Portland Trying to Regulate It?” Smart City PDX, December 4, 2019. https://www.smartcitypdx.com/news/2019/12/4/what-is-facial-recognition-and-why-is-the-city-of-portland-trying-to-regulate-it.

City of Seattle. “City of Seattle Privacy Program: 2019 Annual Report: Transforming Privacy,” 2019. http://www.seattle.gov/Documents/Departments/Tech/Privacy/2019%20Privacy%20Program%20Annual%20Report.pdf.

Cloud Security Alliance. “Security Guidance for Critical Areas of Focus in Cloud Computing V. 3.0,” 2009. https://downloads.cloudsecurityalliance.org/assets/research/security-guidance/csaguide.v3.0.pdf.

Commission Nationale de l’Informatique et des Libertés. “Methodology for Privacy Risk Management: How to Implement the Data Protection Act.” Paris, France, 2012. https://www.cnil.fr/sites/default/files/typo/document/CNIL-ManagingPrivacyRisks-Methodology.pdf.

Compute Ontario, and ORION. “The Future of Ontario’s Data: Fulfilling the Potential of Smart Cities.” Building Ontario’s Next-Generation Smart Cities Through Data Governance, 2019. https://computeontario.ca/wp-content/uploads/2019/11/Smart-Cities_The-future-of-Ontario%E2%80%99s-data.pdf.

"Confronting Black Boxes: A Shadow Report of the New York City Automated Decision System Task Force." AI Now Institute, December 2019. https://ainowinstitute.org/ads-shadowreport-2019.pdf.

Consumer Data Research Centre. "About Our Data." CDRC. Accessed September 26, 2019. https://www.cdrc.ac.uk/about-data/.

Cuggia, Marc, and Stéphanie Combes. "The French Health Data Hub and the German Medical Informatics Initiatives: Two National Projects to Promote Data Sharing in Healthcare." *Yearbook of Medical Informatics* 28, no. 01 (August 2019): 195–202. https://doi.org/10.1055/s-0039-1677917.

LightsOnData. "Data Governance Maturity Models - IBM," August 8, 2018. https://www.lightsondata.com/data-governance-maturity-models-ibm/.

DataGuidance, and Future of Privacy Forum. "Comparing Privacy Laws: GDPR vs. CCPA," 2018. https://fpf.org/wp-content/uploads/2018/11/GDPR_CCPA_Comparison-Guide.pdf.

Delacroix, Sylvie, and Neil D Lawrence. "Bottom-up Data Trusts: Disturbing the 'One Size Fits All' Approach to Data Governance." *International Data Privacy Law*, October 1, 2019, ipz014. https://doi.org/10.1093/idpl/ipz014.

Digital ID and Authentication Council of Canada (DIACC) Trust Framework Expert Committee. "Pan-Canadian Trust Framework Overview: A Collaborative Approach to Developing a Pan-Canadian Trust Framework," August 2016. https://diacc.ca/wp-content/uploads/2016/08/PCTF-Overview-FINAL.pdf.

Electronic Frontier Foundation. "Acoustic Gunshot Detection." Electronic Frontier Foundation, November 4, 2019. https://www.eff.org/pages/gunshot-detection.

Electronic Frontier Foundation, and Open Technology Institute. Letter to Los Angeles City Council and Los Angeles Department of Transportation. "Urgent Concerns Regarding the Lack of Privacy Protections for Sensitive Personal Data Collected Via LADOT's Mobility Data Specification," April 3, 2019.

Element AI, and Nesta. "Data Trusts: A New Tool for Data Governance," 2019. https://hello.elementai.com/rs/024-OAQ-547/images/Data_Trusts_EN_201914.pdf.

Faundeen, John L., Thomas E. Burley, Jennifer A. Carlino, David L. Govoni, Heather S. Henkel, Sally L. Holl, Vivian B. Hutchison, et al. "The United States Geological Survey Science Data Lifecycle Model: U.S. Geological Survey Open-File Report 2013–1265." Open-File Report, 2014. http://dx.doi.org/10.3133/ofr20131265.

Fewer, David. "Open Smart Cities FAQ." Samuelson-Glushko Canadian Internet Policy and Public Interest Clinic (CIPPIC), November 2017. https://cippic.ca/en/Open_Smart_Cities.

First Nations Information Governance Centre. “Data Access at the First Nations Data Centre | FNIGC.” Accessed October 1, 2019. https://fnigc.ca/first-nations-data-centre/data-access-first-nations-data-centre.html.

———. “The First Nations Principles of OCAP®.” Accessed June 11, 2020. https://fnigc.ca/ocap.

———. “Understanding the Basics of OCAP®,” 2019. https://www.fnhma.ca/wp-content/uploads/2019/05/Understanding-the-Basics-of-OCAP.pdf.

Freuler, Juan Ortiz, and Carlos Iglesias. “Algorithms and Artificial Intelligence in Latin America: A Study of Implementation by Governments in Argentina and Uruguay.” World Wide Web Foundation, 2018. http://webfoundation.org/docs/2018/09/WF_AI-in-LA_Report_Screen_AW.pdf.

Gellert, Raphaël. “We Have Always Managed Risks in Data Protection Law: Understanding the Similarities and Differences Between the Rights-Based and the Risk-Based Approaches to Data Protection.” *European Data Protection Law Review* 2, no. 4 (2016): 481–92. https://doi.org/10.21552/EDPL/2016/4/7.

Gittens, Sebastien A, Stephen D Burns, Martin P J Kratz, and Kees de Ridder. “Understanding the GDPR: A Comparison Between the GDPR, PIPEDA and PIPA.” Bennett Jones, May 14, 2018. https://www.bennettjones.com/en/Blogs-Section/Understanding-the-GDPR.

GovLab. “Data Collaboratives.” Accessed August 15, 2019. http://datacollaboratives.org/explorer.html.

Grieman, Keri. “Pedestrian Curiosity: A Brief Examination of Consent and Privacy in Swath Section Smart City Spaces.” *Spatial Knowledge and Information Canada* 7, no. 5 (2019): 1. http://ceur-ws.org/Vol-2323/SKI-Canada-2019-7-5-1.pdf.

———. “Smart City Privacy in Canada.” Report for the Office of the Privacy Commissioner of Canada. Samuelson-Glushko Canadian Internet Policy and Public Interest Clinic (CIPPIC), January 2019. https://smartcityprivacy.ca/wp-content/uploads/2019/03/Greiman-OPC-Report_Final-2019.pdf.

Hardinges, Jack. “Defining a ‘Data Trust.’” *The Open Data Institute* (blog), October 19, 2018. https://theodi.org/article/defining-a-data-trust/.

Heimi, Junya. “Japan’s ‘information Banks’ to Let Users Cash in on Personal Data.” Nikkei Asian Review, May 9, 2019. https://asia.nikkei.com/Business/Business-trends/Japan-s-information-banks-to-let-users-cash-in-on-personal-data.

Information and Privacy Commissioner of Ontario. “Fact Sheet: What Is Personal Information?,” October 2016. https://www.ipc.on.ca/wp-content/uploads/2016/10/what-is-personal-information.pdf.

———. Letter to Waterfront Toronto. “Re: Sidewalk Lab’s Proposal,” September 24, 2019. https://www.ipc.on.ca/wp-content/uploads/2019/09/2019-09-24-ltr-stephen-diamond-waterfront_toronto-residewalk-proposal.pdf.

Jee, Charlotte. “A Facial Recognition Ban Is Coming to the US, Says an AI Policy Advisor.” MIT Technology Review, September 18, 2019. https://www.technologyreview.com/s/614362/a-facial-recognition-ban-is-coming-to-the-us-says-ai-policy-advisor/.

Kendrick, Christine, and Andrew Rodgers. “Recommendations for the Development and Implementation of Distributed Sensor Networks,” 2017. https://www.theenterprisectr.org/uploads/Distributed_Sensor_Networks.pdf.

Kitchin, Rob. “Big Data.” In *International Encyclopedia of Geography: People, the Earth, Environment and Technology*, edited by Douglas Richardson, Noel Castree, Michael F. Goodchild, Audrey Kobayashi, Weidong Liu, and Richard A. Marston, 1–3. Oxford, UK: John Wiley & Sons, Ltd, 2016. https://doi.org/10.1002/9781118786352.wbieg0145.

———. “Getting Smarter About Smart Cities: Improving Data Privacy and Data Security.” Dublin, Ireland: Data Protection Unit, Department of the Taoiseach, 2016. http://smartdublin.ie/wp-content/uploads/2016/12/Smart_Cities_Report_January_2016-Department-of-Taoiseach-1.pdf.

———. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. Los Angeles, California: SAGE Publications, 2014.

Kitchin, Rob, Tracey P. Lauriault, and Gavin McArdle. “Knowing and Governing Cities through Urban Indicators, City Benchmarking and Real-Time Dashboards.” *Regional Studies, Regional Science* 2, no. 1 (January 2015): 6–28. https://doi.org/10.1080/21681376.2014.983149.

Kooper, M. N., R. Maes, and E. E. O. Roos Lindgreen. “On the Governance of Information: Introducing a New Concept of Governance to Support the Management of Information.” *International Journal of Information Management* 31, no. 3 (June 1, 2011): 195–200. https://doi.org/10.1016/j.ijinfomgt.2010.05.009.

Lecher, Colin. “NYC’s Algorithm Task Force Was ‘a Waste,’ Member Says.” The Verge, November 20, 2019. https://www.theverge.com/2019/11/20/20974379/nyc-algorithm-task-force-report-de-blasio.

Lempert, Rainer. “Shared Mobility Data Sharing: Opportunities for Public-Private Partnerships.” TransLink New Mobility Lab, April 23, 2019. https://sustain.ubc.ca/sites/default/files/Sustainability%20Scholars/2018_Sustainability_Scholars/Reports/2018-70%20Shared%20Mobility%20Data%20Sharing%20Opportunities_Lempert.pdf.

MaRS Discovery District. “A Primer on Civic Digital Trusts.” Accessed August 15, 2019. https://marsdd.gitbook.io/datatrust/.

———. “Towards a Smart City Data Trust: Design Recommendations for a Personal Mobility Data Trust.” Building Ontario’s Next-Generation Smart Cities Through Data Governance, 2019. https://computeontario.ca/wp-content/uploads/2019/11/Smart-Cities_MaRS_Towards-a-Smart-City-Data-Trust-1-1.pdf.

McDonald, Sean. “Reclaiming Data Trusts.” Centre for International Governance Innovation. Accessed March 4, 2020. https://www.cigionline.org/articles/reclaiming-data-trusts.

Montjoye, Yves-Alexandre de, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel. “Unique in the Crowd: The Privacy Bounds of Human Mobility.” *Scientific Reports* 3, no. 1 (December 2013): 1376. https://doi.org/10.1038/srep01376.

Moore, Colleen, and Kathy Reeves Bracco. “Scaling Goodwill: The Challenges of Implementing Robust Education Data Sharing Through Regional Partnerships.” California Education Policy, Student Data, and the Quest to Improve Student Progress. Education Insights Center, 2018. https://eric.ed.gov/?id=ED584699.

Mosley, Mark, Michael Brackett, Susan Earley, Deborah Henderson, and Data Administration Management Association, eds. *The DAMA Guide to the Data Management Body of Knowledge*. Bradley Beach, N.J.: Technics Publications, 2009.

Mulgan, Geoff, and Vincent Straub. “The New Ecosystem of Trust.” *Nesta* (blog), February 21, 2019. https://media.nesta.org.uk/documents/nesta.org.uk-The_new_ecosystem_of_trust_-_printable.pdf.

Naden, Clare. “Tackling Privacy Information Management Head on: First International Standard Just Published.” International Standards Organization, August 6, 2019. https://www.iso.org/cms/render/live/en/sites/isoorg/contents/news/2019/08/Ref2419.html.

New York City Automated Decision Systems (ADS) Task Force. “New York City Automated Decision Systems Task Force Report,” November 2019. https://www1.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf.

Nielsen. “A Social License Approach to Trust - A Close-Up on Trust.” Stats NZ, August 2018. https://www.stats.govt.nz/assets/Uploads/Corporate/Measuring-Stats-NZs-social-licence/a-social-licence-approach-to-trust.pdf.

OECD. “Guidelines Governing the Protection of Privacy and Transborder Flows (1980),” 1980. http://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm.

———. “Guidelines Governing the Protection of Privacy and Transborder Flows (2013),” 2013. https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf.

Office of the Privacy Commissioner of Canada. “PIPEDA Fair Information Principles,” September 16, 2011. https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/p_principle/.

Office of the Privacy Commissioner of Canada, Policy and Research Group. “Consent and Privacy: A Discussion Paper Exploring Potential Enhancements to Consent Under the Personal Information Protection and Electronic Documents Act.” Gatineau, May 2016. https://www.priv.gc.ca/en/opc-actions-and-decisions/research/explore-privacy-research/2016/consent_201605/.

Ontario Energy Board. Independent Electricity System Operator (in its capacity as the Smart Metering Entity): Application for approval to provide access to certain non-personal data to third parties at market prices, No. EB-2018-0316 (October 24, 2019).

Open Data Institute. “Data Trusts: Lessons from Three Pilots,” 2019.

Privacy Commissioner of Canada. “Privacy Fact Sheet: General Data Protection Regulation,” July 2018. https://www.ipc.on.ca/wp-content/uploads/2018/07/fs-privacy-gdpr.pdf.

Privacy International. “The Keys to Data Protection: A Guide for Policy Engagement on Data Protection,” August 2018. https://privacyinternational.org/sites/default/files/2018-09/Data%20Protection%20COMPLETE.pdf.

Rocher, Luc, Julien M. Hendrickx, and Yves-Alexandre de Montjoye. “Estimating the Success of Re-Identifications in Incomplete Datasets Using Generative Models.” *Nature Communications* 10, no. 1 (December 2019): 3069. https://doi.org/10.1038/s41467-019-10933-3.

Ruttan, Craig, Raly Chakarova, Natasha Apollonova, Patrick Gill, and Brian Kelcey. “Bibliotech: Beyond Quayside: A City-Building Proposal for the Toronto Public Library to Establish a Civic Data Hub.” Toronto Region Board of Trade, January 2019. https://www.bot.com/Portals/0/Bibliotech%20-%20Final%20-%20Jan%208.pdf?timestamp=1546987861621.

SAIL Databank. “Data Privacy & Security: Ensuring Secure Access to the Data,” 2019. https://saildatabank.com/saildata/data-privacy-security/#secure-access.

———. “SAIL Databank - The Secure Anonymised Information Linkage Databank.” Accessed September 20, 2019. https://saildatabank.com/.

Scassa, Teresa. “Data Ownership.” CIGI Papers, September 2018.

———. “Enforcement Powers Key to PIPEDA Reform.” Policy Options, June 7, 2018. https://policyoptions.irpp.org/magazines/june-2018/enforcement-powers-key-pipeda-reform/.

Sidewalk Labs. “Digital Governance Proposals for DSAP Consultation,” 2018. https://www.waterfrontoronto.ca/nbe/wcm/connect/waterfront/41979265-8044-442a-9351-e28ef6c76d70/18.10.16_SWT_Draft+Proposals+Regarding+Data+Use+and+Governance_Tuesday_730pm.pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=41979265-8044-442a-9351-e28ef6c76d70.

Smart Chicago Collaborative. “Array of Things Civic Engagement Report: A Summary of Public Feedback & the Civic Engagement Process,” August 2016. https://arrayofthings.github.io/engagement-report.html.

Solove, Daniel J. “A Taxonomy of Privacy.” *University of Pennsylvania Law Review* 154, no. 3 (January 1, 2006): 477. https://doi.org/10.2307/40041279.

State of Victoria (Australia). Office of the Victorian Information Commissioner. “Disclosure of Myki Travel Information: Investigation under Section 8C(2)(e) of the Privacy and Data Protection Act 2014 (Vic),” August 15, 2019. https://ovic.vic.gov.au/wp-content/uploads/2019/08/Report-of-investigation_disclosure-of-myki-travel-information.pdf.

Stinson, Catherine. “Healthy Data: Policy Solutions for Big Data and AI Innovation in Health.” Mowat Research. Toronto: Mowat Centre, 2018.

The Asahi Shimbun. “Seeking a Bite of Big Data, Banks to Act as Brokers of Personal Info.” The Asahi Shimbun. Accessed June 13, 2020. http://www.asahi.com/ajw/articles/AJ201906270060.html.

The British Academy, techUK, and The Royal Society. “Data Ownership, Rights and Controls: Reaching a Common Understanding.” Discussions at a British Academy, Royal Society and techUK seminar on 3 October 2018, 2018. https://royalsociety.org/~/media/policy/projects/data-governance/data-ownership-rights-and-controls-October-2018.pdf.

UK Information Commissioner’s Office. “Data Protection Impact Assessments,” June 24, 2019. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/.

University of Chicago, Argonne National Labs, and City of Chicago. “Array of Things Governance Policy and Process,” August 2016. https://arrayofthings.github.io/final-policies.html.

Vincent, Donovan, and David Rider. “Sidewalk Labs Pulls out of Toronto’s Quayside Project, Blaming COVID-19.” *The Toronto Star*, May 7, 2020, sec. City Hall. https://www.thestar.com/news/city_hall/2020/05/07/sidewalk-labs-pulling-out-of-quayside-project.html.

Weill, Peter. “Don’t Just Lead, Govern: How Top-Performing Firms Govern IT.” *MIS Quarterly Executive* 3, no. 1 (2004): 1–17. https://aisel.aisnet.org/misqe/vol3/iss1/3.

Wylie, Bianca, and Sean McDonald. “What Is a Data Trust?” Centre for International Governance Innovation, October 9, 2018. https://www.cigionline.org/articles/what-data-trust.