Data Governance and Digital Infrastructure:

Analysis and Key Considerations for the City of Toronto

Final Report (Appendices) - June 2020

OpenNorth

Table of Contents

List of Figures

List of Tables

Appendix A: Case study criteria

A list of criteria and clarifying questions were supplied by the City of Toronto (Table A:1) and used to guide the case study analysis.

Table A:1 Case study analysis criteria

Criterion	Clarifying questions
I. Regulatory Framework	What are the regulatory frameworks, statutes and case law on which the structure depends for its establishment and continued effectiveness (e.g. GDPR)?
ii. Democratic accountability and transparency	In operation, how is the structure or organization governed in a democratic way that allows for consultation and public participation?
iii. Supporting Policies, Principles, Frameworks and Risk Mitigation strategies	What are the identified guiding principles and what identified risks is it intended to mitigate?
iv. Consultation Process	What consultation process was used to inform the creation of the structure?
v. Link to Government Approvals	How is the organization linked to government approvals processes, if at all? Are there any specific obligations related to developing proposals for review and approval by the governance organization?
vi. Public Interest and End Goals	How is the Public Interest defined? What Social, Economic and Environmental goals is the data governance structure intended to advance? What outcomes have been reported?
vii. Equity and Human Rights	How are Equity and Human Rights accounted for in the governance structure? What outcomes have been reported?
viii. Collection Categorization and Storage	What types of data are collected and how are they categorized? What type of hardware is used to collect and store this data?
ix. Consent	How is consent for data collection received, if at all? Is it possible to opt-out of data collection?
x. Privacy	What tools and policies are used to protect privacy? How is privacy defined?
xi. Cybersecurity and Hardware Management	Who builds, manages, repairs and updates the physical infrastructure in the public realm and back office? How does the structure protect against service disruptions and breaches? This report is not intended to focus on broader cybersecurity/critical

Criterion	Clarifying questions
	infrastructure cybersecurity issues, but when describing the data governance structure, the supporting role of the relevant cybersecurity approach should be articulated.
xii. Data Residency / Data Sovereignty	How, if at all, is the physical and geographical location of the data regulated (e.g. must personally identifiable information be stored on servers within the jurisdiction)?
xiii. Ethical Use of Data and Technology	How is the use of the data regulated? Are there restrictions on the applications that can use it?
xiv. Transparency and Individual Control	How are individuals informed about the collection and storage of information and to what degree are individuals able to access their personal information and control its use?
xv. Intellectual Property Rights, Ownership and Equitable Distribution of Value	Who has intellectual property rights and how is the value and compensation for data use determined? How is ownership of the data defined?
xvi. Open Data	What is the role of Open Data?

Appendix B: Notes on research

B.1 Selection of case studies

We identified an initial list of 34 organizations from which case studies could be drawn by consulting recent research from several organizations (Table B:1). We supplemented this scan with extensive desk research and 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. We also established the following focuses:

- Governance models involving governments or other public-sector institutions as these were deemed to have the most relevance to the City of Toronto's context.
- Existing or are soon-to-be-implemented cases. We did not choose examples that had no current or intended implementers, as these would be theoretical and not constitute 'case studies.'

Table B:1 Major sources consulted

Source	How we used it
MaRS Discovery District. "A Primer on Civic Digital Trusts." https://marsdd.gitbook.io/datatrust/.	Source lists several examples: • Barcelona: DECODE project [not selected] • London: Smarter London Together, London Datastore [not selected] • Trūata: a private partnership between IBM and Mastercard [not selected] • US Nationwide Health Information Network [not selected] • We selected the Silicon Valley Regional Data Trust for further investigation.
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.	Assisted in framing our understanding of the emerging data trust landscape. However, it did not discuss any cases in-depth.
GovLab. "Data Collaboratives." http://datacollaboratives.org/explorer.html.	Source contains a longer list of approximately 150 potential examples. The majority of the examples were products (e.g., proprietary analytics platforms), software-as-services, APIs, campaigns, and technology competitions which were deemed to be outside the scope of a data governance

Source	How we used it
	model. In most cases, limited information was provided beyond a short summary paragraph (self-reported). Potentially interesting examples were investigated further such as Data Ventures (Stats NZ) and the Consumer Data Research Centre.
Hardinges, Jack. "What Is a Data Trust?" The Open Data Institute (ODI), July 10, 2018. https://theodi.org/article/what-is-a-data-trust/.	Source contains several resources relating to data trusts.

Table B:2 Cases not selected for investigation

Case	Source	Rationale
Amsterdam Data Exchange	Nesta	Insufficient documentation; looked at other Amsterdam examples; abundance of EU case studies
Truata	MaRS	Insufficient documentation (private sector example, private-to-private partnership). This is a data management platform and data analytics company selling a product, not a governance model
London: Smarter London Together, London Datastore	MaRS	Insufficient documentation available on data governance; abundance of EU case studies
Greater London Authority/Borough of Greenwich data trust pilot	ODI	Pilot project; mainly an exploration of public decision-making processes
US Nationwide Health Information Network	MaRS Civic Digital Trust Primer	Overlap with our existing health information case from Australia
Information Commissioner's Office regulatory beta sandbox (UK)	Nesta	Insufficient documentation (program participants selected in Summer 2019)
Personal data stores (Mydex, SOLID, Citizen-me, digi.me)	Nesta	Insufficient documentation about PDSs in action. The focus here is on personal data ownership and

Case	Source	Rationale
		portability between different proprietary internet platforms; relevance for government unclear.
India's Data Protection Bill and Privacy Protection Framework	City of Toronto	National-level legislation, not a governance structure. Unable to find detailed documentation on implementation.
Pakistan's Data Protection Bill	City of Toronto	National-level legislation, not a governance structure. Unable to find detailed documentation on implementation
Alternative Camden: Community Data Trust	Desk research	Not implemented; conceptual phase only - no implementation or chosen direction.
Structural Genomics Consortium	Desk research	Insufficient documentation about data governance; attempts to contact were unsuccessful.
Te Mana Raraunga: Māori Data Sovereignty Network	Desk research	Not governance structure, but rather principles.
Singapore Data Protection Act	ODI	National-level legislation; unable to locate detailed documentation of implementation.
Copenhagen City Data Exchange	Desk research	City and private sector initiative; however, limited documentation available.
Amsterdam Data Exchange	Nesta	Insufficient documentation; looked at other Amsterdam examples; abundance of EU case studies
Truata (Mastercard)	MaRS	Insufficient documentation (private sector example, private-to-private partnership). This is a data management platform and data analytics company selling a product, not a governance model
London: Smarter London Together, London Datastore	MaRS	Insufficient documentation available on data governance; abundance of EU case studies

Note: this table includes 14 cases not selected for further examination but excludes the GovLab Data Collaboratives database (which contained over 150 examples at the time of writing).

B.2 Evaluating case studies

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.

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,[1] 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.[2] We also drew upon current thinking among technology and legal scholars to inform our analysis.

B.3 Overview of case studies

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

[1] See for example “Data Governance Maturity Models - IBM.”

[2] Abraham, Schneider, and vom Brocke, “Data Governance.”

Table B:3: Description of all 20 case studies

Name	Maturity Level[3]	Documentation[4]	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 protection legislation after it released travel

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

Name	Maturity Level[3]	Documentation[4]	Organization Type	Location	Description
					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 Surveillance Impact Reports for new technology.

Name	Maturity Level[3]	Documentation[4]	Organization Type	Location	Description
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 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

Name	Maturity Level[3]	Documentation[4]	Organization Type	Location	Description
					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 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.

Name	Maturity Level[3]	Documentation[4]	Organization Type	Location	Description
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.
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

Name	Maturity Level[3]	Documentation[4]	Organization Type	Location	Description
					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 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

Name	Maturity Level[3]	Documentation[4]	Organization Type	Location	Description
					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 & public agencies	USA	New York City became one of the first jurisdictions to pass a law on automated decision systems (ADS). The ADS Task 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.

Name	Maturity Level[3]	Documentation[4]	Organization Type	Location	Description
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 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.

B.4 Case study research notes

B.4.1 Ontario's Smart Energy Metering Program

B.4.1.1 Profile

- Location: Ontario, Canada
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Established

B.4.1.2 Background

In Ontario, smart meter technology is part of a smart grid – tracking energy use in the home on an hourly basis and sending this information back to the Local Distribution Company that services and bills the customer. Smart metering was deployed as a means of managing demand for electricity more efficiently. 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.

B.4.1.3 Regulatory Framework

According to source materials, data governance complies with the following legislation:

- Section 53.7 of the Electricity Act, 1998,14 charges the Smart Metering Entity (SME) with carrying out the government's Smart Metering Initiative (SMI).
- Section 5(1) of Regulation 393/07 gives the SME authority over receiving smart metering data and providing all services performed on smart metering data to create billing quantity data, including validation, estimating and editing services.
- Freedom of Information and Protection of Privacy Act [FIPPA]. Private sector actors requesting access to data in the MDM/R would be governed by the federal Personal Information Protection and Electronic Documents Act (PIPEDA). Any collection, use or disclosure of personal data is subject to the oversight of the OIPC (Ontario Information and Privacy Commissioner).[5]

B.4.1.4 Democratic Accountability and Transparency

Nominally transparent due to institutions involved and reporting obligations. However, the public may find it difficult to follow Ontario Energy Board activities.

[5] Scassa and Vilain, "Governing Smart Data in the Public Interest Lessons from Ontario's Smart Metering Entity."

B.4.1.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

The Third Party Access Implementation Plan 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 re-identification risk.

B.4.1.6 Public Consultation

In 2015, the SME Working Group conducted broad outreach with experts, stakeholders, and the general public.

In an Ontario Energy Board decision, it was noted that, while the SME has attempted to consult on these issues, “a more comprehensive consumer engagement process should take place.”[6]

B.4.1.7 Government Approvals

As a public body that must act under terms set by the regulator (the Ontario Energy Board), it requires OEB approval of any data sharing plan.

B.4.1.8 Public Interest and End Goals

As a public body, the SME was charged with administering the data in the public interest defined as energy conservation and better energy management.

Sale of data was proposed to be done at “market price” – not just on a cost recovery basis. Profits would be passed onto consumers in the form of lowering the Smart Metering Charge, resulting in lower energy costs for consumers.

Finally, there is an additional goal of stimulating innovation by providing the private sector with useful data.

B.4.1.9 Equity and Human Rights

No information found.

B.4.1.10 Data Collection, Categorization and Storage

Central data hub is operated by IBM. 65 local distribution companies (LDCs) manage around five million smart meters, which transfer data hourly. Data travels from smart meters installed at a residence, a business or a sub-metering entity to neighbourhood collectors. These collectors transmit the data to control computers that transmit it to the meter data management repository.

[6] 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.

In its application to the Ontario Energy Board for a license,[7] the SME proposed to begin providing data access to third parties (researchers, governments and private firms). It proposed three categories of data:[8]

Public offers: monthly, seasonal or quarterly data aggregated by postal district (first digit of postal code), provided to users free of charge and subject to unspecified terms and conditions.

Standard private offerings: “pre-designed extracts based on popular data requests.” such as “Hourly or daily consumption data aggregated by 6, 5, 4- or 3-digit Postal Code at the municipal level, specifying the Distributor Rate Class and Commodity Rate Class”, as well as different types of visualizations. Access to this category of data would be provided subject to a Data Use Agreement and at “market prices.”

Custom private offerings: data sets customized to meet the requirements of specific clients, subject to a Data Use Agreement and sold at “market price.”

B.4.1.11 Consent

While customers cannot opt out of having a smart meter installed on their property, the Green Button Initiative makes it possible for customers to view their own energy use data from meter readings. The standard allows customers to consent and control with regards to whom they share data about their energy use.[9]

B.4.1.12 Privacy

The SME collected the following information associated with each meter (modified where necessary to sufficiently render it non-personal information): postal code, distributor rate class, commodity rate class and occupant change data.

Data was required to be depersonalized. Other personal information, such as consumers’ names, addresses or phone numbers are not available to the SME.

However, according to the consultant retained to address privacy issues linked to de-identification and data aggregation, there was a small risk that, despite the recommended risk mitigation techniques, information contained in the meter data management repository could be used by a recipient to identify the home linked to the data.

[7] 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.

[8] Scassa, “Plans to Sell Smart Meter Data Nixed over Privacy and Consultation Concerns.”

[9] Tracey P. Lauriault, Rachel Bloom, and Jean-Noé Landry, “Open Smart Cities in Canada: Assessment Report,” April 11, 2018, SocArXiv, https://doi.org/10.31235/osf.io/qbyzj.

Indeed, a recent decision by the Ontario Energy Board denied the SME's license to sell de-identified data to third parties on the grounds that "it is 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."[10]

B.4.1.13 Cybersecurity and Hardware Management

The SME elected not to allow third-party access via APIs until greater experience with third-party data sharing was acquired

B.4.1.14 Data Residency and Data Sovereignty

No information found.

B.4.1.15 Ethical Use of Data and Technology

Data sharing with third parties would be subject to data-sharing agreements that would place limits on how third parties could use the data provided (no direct access via API).

B.4.1.16 Transparency and Individual Control

The smart meter (which remains on customers' property once installed) collects consumption data, and the consumer receives a monthly bill from their local distribution company based upon this consumption.

Electricity consumers have access to a dashboard which details their personal consumption.

B.4.1.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

The SME is considering charging fees for access to and reuse of data, however, has not moved forward with this proposal to date.

Assuming inelastic demand for data (due to lack of alternatives), this would put the SME in considerable price-setting power

B.4.1.18 Open Data

No information found.

[10] 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.

B.4.2 Silicon Valley Regional Data Trust

B.4.2.1 Profile

- Location: California, USA
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Established

B.4.2.2 Background

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. SVRDT brings together data from numerous public agencies serving children and families including public schools, health and human services organizations, and juvenile justice systems.[11]

B.4.2.3 Regulatory Framework

According to their website, data governance complies with confidentiality and privacy rules and regulations, both at the federal and state levels are applicable, including:

- The Family Educational Rights and Privacy Act (FERPA)
- The Health Insurance Portability and Accountability Act (HIPAA) and;
- The Children's Online Privacy Protection Act (COPPA)

B.4.2.4 Democratic Accountability and Transparency

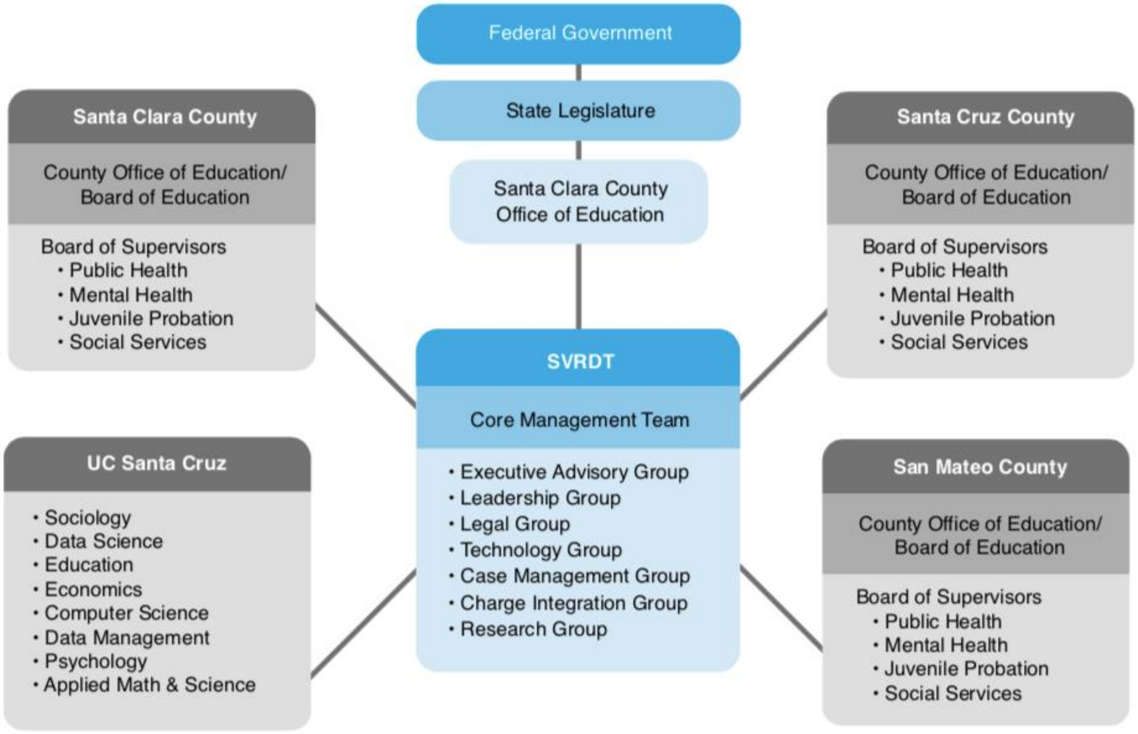
Santa Clara County is the official organizational 'home' of the trust; the County is ultimately accountable to the State legislature.

SVRDT governance structure comprises seven working groups: executive, administrative leadership, legal, technology, research, change integration, and case management.

These working groups include representation from data-contributing stakeholder groups from each of the three counties: the three county offices of education, Behavioral Health, Juvenile Probation, Child Welfare, and UCSC researchers.

[11] Allison-Jacobs, "IDS Case Study: Silicon Valley Regional Data Trust: Supporting Students through Integrated Data and Research-Practice Partnerships."

Figure B-1 Silicon Valley Regional Data Trust organizational chart



B.4.2.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

SVRDT is governed internally by the following mechanisms:

Multi-Agency Agreement: the three counties and the individual agencies within the three counties have agreed to a high-level, policy-based document (Multi-Agency Agreement) setting out the purpose for the SVRDT

Enterprise Memorandum of Understanding (E-MOU): defines what data will be shared, access rights and credentials, credentialing processes, the process for identifying the clients served by the different systems and possibly by the different counties, and the hardware/software criteria and processes.

B.4.2.6 Public Consultation

No public consultation. However, significant face-to-face consultations required between the original four-member planning team, agency members, and policymakers across the three counties in order to build the necessary level of trust to move forward with the project.

Internal consultation included development of client personas and 30 real-world scenarios in order to discuss applications and potential legal issues with data sharing.

B.4.2.7 Government Approvals

County leadership approves data access requests.

B.4.2.8 Public Interest and End Goals

Public interest is defined in terms of:

Data sharing to research contributing factors of student success and failure

Sharing across previously siloed institutions: public school districts, public health, child and family services, mental health, juvenile justice and Education Technology companies

B.4.2.9 Equity and Human Rights

Source report states that data access will be used to improve equity but does not substantiate claims.

“Data will be accessed to improve equity for persons served by the counties. All of the parties agree to implement processes to overcome institutional, structural, and individual racism that exists in society; to enable the children, youth and families to lead more prosperous and rewarding lives; and for the government to provide services that are the most effective possible.”[12]

B.4.2.10 Data Collection, Categorization and Storage

Users can access existing case information residing in different agencies’ case management systems. These access requests are managed by the SVRDT Data Portal.

B.4.2.11 Consent

A universal consent form was developed that will be applicable for all four systems. A client or participant will be able to sign one consent in any of the four systems, which will apply if that individual is involved with more than one system.

B.4.2.12 Privacy

All confidentiality and privacy rules and regulations, federal and state levels are to be followed

Unique identifier: California Department of Education Statewide Student Identifier (SSID) serves as a unique identifier

- Note: SSIDs are managed centrally by state government (California Longitudinal Pupil Achievement Data System (CALPADS)) - schools obtain SSIDs via CALPADS

[12] Allison-Jacobs, “IDS Case Study: Silicon Valley Regional Data Trust: Supporting Students through Integrated Data and Research-Practice Partnerships,” 9.

- This suggests that unique identifiers within this closed system are subject to external influence - stakeholders of the system have no control over SSIDs
- Unclear how privacy protection is maintained without hashing of SSIDs

B.4.2.13 Cybersecurity and Hardware Management

SVRDT Secure Data Environment - application platform to allow certified users query access to case management data across agencies

Data Portal - data exchange platform to make and respond to data requests from certified users

B.4.2.14 Data Residency and Data Sovereignty

- DataZone (DZ) - educational data platform (dashboards, cohort tracking, student assessment data, data warehousing) used as data storage
- DataZone owner: Santa Clara County Office of Education (SCCOE)

DataZone developer (probable): Hoonuit (private sector analytics company focused on the education sector)

Hoonuit is referred to as a partner of SCCOE. Based on lack of back-end developers amongst SCCOE DataZone team, platform development is likely outsourced to Hoonuit, but relationship is referred to as a 'collaboration'[13]

  - Unclear status of Hoonuit and any other outsourced vendors within the data governance framework
  - SCCOE hosts data from public school in three counties: Santa Clara, Santa Cruz, San Mateo
    - 30 school districts on over 280,000 students (more than 60% of the three-county student population)

B.4.2.15 Ethical Use of Data and Technology

"Do No Harm" philosophy is embedded into the SVRDT governance framework with the aim of addressing issues of disproportionality and systemic inequities. However, it is unclear how this is put into practice.

Academic research conducted by UC Santa Cruz based on SVRDT data is subject to institutional review board ethics approval.

[13] Benson, "Customer Spotlight: Santa Clara County Office of Education Is Using Data to Change Outcomes."

B.4.2.16 Transparency and Individual Control

Individuals have a right to know which data are being shared and with whom. Clients must consent to their data being utilized.

B.4.2.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

Data is the property of contributing agencies and governments. No statement regarding value distribution

B.4.2.18 Open Data

No open data is provided.

B.4.3 Public Transport Victoria

B.4.3.1 Profile

- Location: Victoria, Australia
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Established

B.4.3.2 Background

Public Transport Victoria (PTV) was the system authority for public transport in Victoria, Australia up until July 2019 when its functions were transferred to the Department of Transport. One of its functions was to administer and operate ticketing systems, including the myki smart fare card. These smart cards collect tap-on / tap-off data to record payment for public transport. PTV was found in breach of national privacy and data protection legislation after an incident in July 2018 where it released travel history data for 15 million smart fare cards to an open data hackathon event. This data breach could have allowed for individuals to be re-identified. A subsequent investigation revealed deficiencies in PTV's data governance and risk management processes.

B.4.3.3 Regulatory Framework

Privacy and Data Protection Act 2014 (Vic) (PDP Act) and the Health Records Act 2001 (Vic).[14]

B.4.3.4 Democratic Accountability and Transparency

PTV is subject to the oversight of the Office of the Victorian Information Commissioner (OVIC).

B.4.3.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

DataVic Access Policy guides public access to government data to "support research and promote innovation"

PTV also uses the Five Safes Framework to assess risks. The five elements of the framework are as follows:

- Safe project: Is the use of the data appropriate?
  - Interpretation: Use of the data is legal, ethical and the project is expected to deliver public benefit.
- Safe people: Can the users be trusted to use it in an appropriate manner?

[14] Public Transport Victoria. "Information Privacy Policy." Public Transport Victoria. Accessed March 18, 2020. https://www.ptv.vic.gov.au/footer/legal-and-policies/information-privacy-policy/.

    - Interpretation: Researchers have the knowledge, skills and incentives to act in accordance with required standards of behaviour.
- Safe data: Is there a disclosure risk in the data itself?
    - Interpretation: Data has been treated appropriately to minimize the potential for identification of individuals or organizations.
- Safe settings: Does the access facility prevent unauthorized use?
    - Interpretation: There are practical controls on the way the data is accessed – both from a technology perspective and considering the physical environment.
- Safe output: Are the statistical results non-disclosive?
    - Interpretation: A final check can be required to minimize risk when releasing the findings of the project.

However, OVIC's investigation into the above-mentioned data breach concluded that a flawed privacy impact assessment (PIA) process contributed to the breach. The PIA template consisted of four parts:

- Part 1, which asks the user to describe the project or program in question and identify whether or not personal information will be handled;
- Parts 2 and 3, which requires the user to identify privacy risks and potential mitigation strategies, and;
- Part 4, which is simply a summary and sign-off section.

However, if the user answered in Part 1 that no personal information would be used, they were not required to provide any further explanation of their reasoning and could proceed to the summary and sign-off section. It was further concluded that the completed PIA document was treated as an authorizing document for data release, rather than as it was intended: as an aid for organizations to identify and address privacy risks.

B.4.3.6 Public Consultation

Not specified.

B.4.3.7 Government Approvals

Not specified.

B.4.3.8 Public Interest and End Goals

PTV's mission statement is to provide an integrated public transportation network. Its data governance policies are represented as furthering this goal.

B.4.3.9 Equity and Human Rights

Not specified.

B.4.3.10 Data Collection, Categorization and Storage

PTV's data security policies state that it will take reasonable steps to protect it from misuse and loss and unauthorised access, modification or disclosure.

B.4.3.11 Consent

Not specified.

B.4.3.12 Privacy

PTV states that it upholds privacy by collecting, using, storing and disclosing personal information and health information in accordance with Information Privacy Principles (IPPs) set out in the Privacy and Data Protection Act 2014 (Vic) (PDP Act) and the Health Privacy Principles (HPPs) and the Health Records Act 2001 (Vic).

B.4.3.13 Cybersecurity and Hardware Management

Not specified.

B.4.3.14 Data Residency and Data Sovereignty

Not specified.

B.4.3.15 Ethical Use of Data and Technology

PTV states that it complies with the purpose limitation principle, only collecting information if it is absolutely necessary for the provision of services.

B.4.3.16 Transparency and Individual Control

People are entitled to contact the PTV Information Privacy Officer and request access to, and correction of any of their personal information or health information held by PTV.

B.4.3.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

Not specified.

B.4.3.18 Open Data

Not specified.

B.4.4 Australian Institute of Health and Welfare

B.4.4.1 Profile

- Location: Australian Capital Territory, Australia
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Established

B.4.4.2 Background

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.

B.4.4.3 Regulatory Framework

AIHW was established under the Australian Institute of Health and Welfare Act 1987

The Privacy Act 1988 (Privacy Act) establishes obligations on private and public sector organisations for collecting, using or disclosing personal information.

B.4.4.4 Democratic Accountability and Transparency

Several structures and roles:

- AIHW Board: have knowledge or experience relevant to the work of AIHW and come from a wide range of Commonwealth, State and other organizations. The Board has approved the Data Governance Framework.
- Ethics Committee: forms an opinion, on ethical grounds, about the acceptability of, and to impose any conditions that it considers appropriate.
- AIHW Director: with the power to manage the affairs of the Institute, subject to the directions of, and in accordance with policies determined by the Board.
- Audit and Finance Committee: comprised of three non-executive members of the Board and one independent member. It authorizes and oversees the audit program and reports to the Board on strategic, financial and data audit matters.
- Executive Committee: which comprises the Director and AIHW Group Heads, supports the Director in managing the day-to-day affairs of AIHW.
- Data Governance Committee: reports to the Executive Committee and make recommendations in relation to data governance and data-related matters.
- Data Custodians: are staff members with delegation from the AIHW Director to exercise overall responsibility for a specific data collection, in accordance with policies, guidelines and any specific condition for use data collection.

- Security roles: several key security roles such as Security Executive and an Agency Security Adviser based on the Protective Security Policy Framework.
- Privacy Officer: first point of contact for advice on applying the Privacy, handling privacy complaints, and training staff.
- Privacy Champion: a role mandated by the Privacy Act for all agencies. The Privacy Champion provides high-level leadership on strategic privacy issues and promotes a culture of privacy within the organization. They also monitor progress with respect to the Privacy Management Plan.

B.4.4.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

Current AIHW practices in data linkage, confidentialization, data security and data access and release practices are being mapped to the Five Safes framework (described in B.4.3.5). These five dimensions are assessed separately, then all together to determine the overall safety of the data collection or sharing project.

Guidelines for the custody of AIHW data describes roles of staff and contractors in relation to that data.

B.4.4.6 Public Consultation

No information found.

B.4.4.7 Government Approvals

AIHW requires no government approvals as its mandate is established through legislation (AIHW Act).

B.4.4.8 Public Interest and End Goals

Public interest is defined as providing “reliable, regular and relevant information of Australia’s health and welfare.”

B.4.4.9 Equity and Human Rights

No information found.

B.4.4.10 Data Collection, Categorization and Storage

AIHW collects personal information either directly from members of the public or organizations, or indirectly through Australian government agencies, state or territory government bodies or other organizations.

B.4.4.11 Consent

Where necessary, consent is obtained at the point of data collection. No further consent is required for AIHW to release data to researchers.

B.4.4.12 Privacy

Privacy is defined by reference to the Privacy Act 1988.

B.4.4.13 Cybersecurity and Hardware Management

An ICT Steering Committee oversees information architecture, systems architecture and technology platforms. It also uses a number of systems and tools to support AIHW's data governance, including separately restricted access at the network, server, and database level requiring individual authorization.

AIHW's data catalogue identifies the data custodians responsible for each data collection, as well as describing the data collections.

B.4.4.14 Data Residency and Data Sovereignty

No information found.

B.4.4.15 Ethical Use of Data and Technology

AIHW uses EthOS, a web-based application to manage researcher ethics review requests.

AIHW is an accredited integrating authority, which means it can perform data linkages – bringing together information from different sources – to create new insights. However, because of the potential risk of reidentifying individuals, strict linkage protocols must be followed to ensure that the risks have been assessed, managed and mitigated throughout the project.

B.4.4.16 Transparency and Individual Control

Privacy Champion is required to report regularly to the Executive Committee about any privacy issues.

B.4.4.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

No information found.

B.4.4.18 Open Data

AIHW holds many open data sets at the aggregate level: social security payments, health and welfare services, population groups, health conditions, disabilities and deaths, and behavioural and risk factors.

B.4.5 Seattle Privacy Program

B.4.5.1 Profile

- Location: Washington (State), USA
- Organization type: Government & public agencies
- Level of documentation: Documented (self-reported)
- Maturity level: Established

B.4.5.2 Background

In response to the privacy implications of smart city technologies and a number of criticisms of the city's data practices, Seattle has established a Privacy Office to assess the ways in which the city authorities generate, store and use data, and to consider issues such as confidentiality, anonymity, archival procedures, deletion, sharing and publishing as open data, and the ability to conduct forensic internal audits.

A notable feature of Seattle's privacy program is its creation of an inventory of all surveillance technologies and the preparation of Surveillance Impact Reports for new technology.

B.4.5.3 Regulatory Framework

Federal and state privacy protection laws apply.

Additionally, on February 23, 2015, Seattle's City Council unanimously approved a resolution to provide a framework for dealing with current and future technologies that impact privacy.[15] The resolution adopted six privacy principles to guide the City in collecting and using information from the public. The Council also established an August 2015 reporting deadline for City departments to create a "Privacy Toolkit."

The City of Seattle Surveillance Ordinance 125376 took effect on September 1, 2017 and requires that:

1. *For each new technology that meets the criteria for surveillance, a City department must prepare a Surveillance Impact Report ("SIR"). These reports include an in-depth review of privacy implications, especially relating to equity and community impact.*

2. *At least one community meeting with comments collected from that meeting submitted to Council via the SIR. Council may require departments to conduct additional community engagement on the technology.*

3. *Council review and vote about the acquisition and deployment of all new and currently-used surveillance technologies.*

[15] City of Seattle, "City of Seattle Privacy Principles."

4. *Regular, detailed reports on surveillance technology use, community equity impact, and non-surveillance technology acquisitions.*

B.4.5.4 Democratic Accountability and Transparency

Chief Privacy Officer provides overall leadership and direction to the Privacy Program, including working with the City Auditor to assess compliance with the city's Privacy Principles. The Chief Privacy Officer is ultimately accountable to the Mayor and City Council.

Departmental 'Privacy Champions' across different agencies handle basic enquiries, conduct and sign-off low-risk privacy reviews, and escalate or reporting issues to the Privacy Program Manager, who is responsible for coordinating the Privacy Champions and 'cultivating a community of practice to share knowledge and best practices'.[16]

B.4.5.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

As noted in B.4.5.3, City Council issued an Ordinance which outlines a range of procedures designed to increase transparency around the city's use of surveillance technologies.

Before the City Council will consider approval of a surveillance technology, the relevant department must host public meetings and invite feedback on the technology via a web tool.

Six privacy principles provide an ethical framework and guide the actions the City takes when collecting and using residents' data:

- Conduct a Privacy Impact Assessment prior to collecting, using and disclosing personal information
- Collect only the information needed to deliver City services and keep it as long as legally required.
- Inform residents about the ways their personal information will be used at the time of collection and requiring their consent whenever possible about how it is used.
- Protecting personal information from unauthorized access and cybersecurity threats
- Follow federal and state laws about information disclosure when working with outside governmental agencies and in answering Public Disclosure Requests. In addition, external partners and contracted vendors must agree to the City's privacy requirements.
- Commit to correcting inaccurate personal information.

B.4.5.6 Public Consultation

Seattle's privacy program was developed in 2015 in collaboration with community activist groups and 'privacy thought leaders from academia, local companies, and private legal practice.'

[16] Bass, Sutherland, and Symons, "Reclaiming the Smart City: Personal Data, Trust and the New Commons."

B.4.5.7 Government Approvals

No information found.

B.4.5.8 Public Interest and End Goals

Public interest is framed, via the Privacy Principles, as a balance between collecting information to provide essential municipal services and safeguarding the privacy of residents.

B.4.5.9 Equity and Human Rights

The City has committed to protecting vulnerable populations from improper data use. To this end, their Privacy Principles were developed to be consistent with their Race and Social Justice Initiative.

One method for operationalizing these principles is through the City's "Data and Survey Demographic Data Collection Playbook". The Playbook teaches City staff about the privacy implications of municipal surveys and provides a guide for how to design a survey that collects demographic data in alignment with the City's Privacy Principles.

B.4.5.10 Data Collection, Categorization and Storage

The City collects multiple types of data, which are stored on various departmental and agency systems:

Personally identifiable information: Name, address, age, birthdate, social security number, driver's license number

Website information: Information passively gathered from visitors to their website (including visits from mobile devices)

Financial information and payment card information: Bank account number, credit or debit card numbers, or other billing information, such as when residents pay utilities, pay taxes, or sign up program membership or classes

Health records: Medical information collected during emergency response, vaccination records, health program participation

Digital images: Facility security cameras, City sponsored event photos, traffic camera video

Utility use: Consumption data about electricity, water and waste management services

Permitting information: New construction, reconstruction and remodeling, land use, events, utilities

Public safety: Violations, court records, emergency calls

Traffic movement: Traffic flows, event monitoring

Demographic information: Income bracket, gender, race or ethnicity, vocation

It categorizes these according to their level of sensitivity: public, sensitive, confidential, and confidential requiring special handling.

B.4.5.11 Consent

See Transparency and Individual Control

B.4.5.12 Privacy

- The City commits to take the following measures when collecting data:

“Minimize Data Collection. Minimizing data means only collecting what is necessary to get done the job at hand.

Provide Notice. Clearly communicating about our data collection and use and provide access to our Privacy Statement.

Review Obligations. Understanding and follow and legal, contractual, and other obligations.

Review Data and Systems Security. Taking steps to secure adequately stored data.

  - Delete or De-identify Data. Follow City data retention schedules and dispose of data as required.”[17]

B.4.5.13 Cybersecurity and Hardware Management

The City uses “physical, administrative and technological techniques to protect data including but not limited to access control, monitoring, auditing, and encryption to secure data.”

- Uses the OneTrust platform[18] to create specific review workflows for Open Datasets and Contract reviews. This helps to minimize the assessment form length and time to complete, while improving the quality of information received.

B.4.5.14 Data Residency and Data Sovereignty

No information found

[17] City of Seattle, Information Technology, “Data the City Collects.”

[18] OneTrust is a data privacy management compliance platform which helps businesses adhere to the growing array of regulations, including GDPR and California Consumer Privacy Act.

B.4.5.15 Ethical Use of Data and Technology

Privacy Principles (See Supporting Policies section) provide guidance for ethical use of data and technology.

As consultation is only carried out via web tool, there is a risk of excluding individuals without Internet access from public engagement activities.

B.4.5.16 Transparency and Individual Control

Where it is possible, the City commits to presenting information about what they are collecting and provide an opportunity to accept or decline to provide it to them, such as follow-up communications not directly related to the service being requested.

However, according to an external report, there appear to be few coordinated efforts to provide specific notices to individuals at the time of data collection about the possibility of their data being released publicly.[19]

B.4.5.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

Data remains the property of the collecting department, agency or firm, but are all subject to the City's privacy policy.

B.4.5.18 Open Data

Seattle's Privacy Program also includes an Open Data Policy, which was developed in collaboration with various partners, including the University of Washington.

This policy stipulates that government data should be 'open by preference', meaning that the city reserves the right to withhold data if it has the potential to cause privacy harms. Datasets must be reviewed for potential privacy harms prior to release, and an annual risk assessment must be performed of both the Open Data Program and Open Data Portal.

[19] Future of Privacy Forum, "City of Seattle: Open Data Risk Assessment" (Washington, D.C.: Future of Privacy Forum, January 2018), https://www.seattle.gov/Documents/Departments/SeattleIT/DigitalEngagement/OpenData/FPF-Open-Data-Risk-Assessment-for-City-of-Seattle.pdf.

B.4.6 Chicago Array of Things

B.4.6.1 Profile

- Location: Illinois, USA
- Organization type: Government & public agencies; academic & research institutions
- Level of documentation: Documented (self-reported)
- Maturity level: Established

B.4.6.2 Background

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. Implementation began in 2016, with the installation of a small number of sensors downtown and elsewhere. By late 2019, approximately 130 sensor nodes had been installed across the city, with the city's open data portal listed locations for all of AoT's active and yet-to-be installed sensors. Data collected from AoT was made accessible online, providing valuable information for researchers, urban planners, and the general public.[20]

B.4.6.3 Regulatory Framework

Array of Things Operating Policies (including privacy policy)

Federal and state privacy protection laws apply.

An issue raised during public consultation was to what extent data collected would be subject to Freedom of Information Act disclosure requests.

B.4.6.4 Democratic Accountability and Transparency

The 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.

[20] Sean Thornton, "A Guide to Chicago's Array of Things Initiative," Data-Smart City Solutions, January 2, 2018, https://datasmart.ash.harvard.edu/news/article/a-guide-to-chicagos-array-of-things-initiative-1190.

B.4.6.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

The governance structure was implemented in three bodies: (a) Executive Oversight Council (EOC), (b) Scientific Review Group (SRG), and (c) Technical Security and Privacy Group (TSPG). The AoT project team regularly updated these groups regarding progress and new challenges or opportunities.[21]

The EOC was co-chaired by the AoT project director and Chicago's Commissioner for Innovation and Technology

Also included individuals with varying perspectives and backgrounds, including academia, industry, policymakers, and community organizations. The EOC provided guidance to the project regarding policy and public engagement, along with an approval process for policies and major operational changes such as those related to privacy, data access, or installation and location selection strategies.

The SRC provided feedback and guidance to the project team and the EOC regarding the scientific and technical directions and services provided by AoT. The SRC provided evaluation of major technical changes in terms of their merit in enhancing the scientific utility of the instrument.

The TSPG reviewed AoT privacy policies and any proposed changes, advising the EOC on their findings. It also had the ability to audit projects as needed.

B.4.6.6 Public Consultation

Draft governance and privacy policies were created in 2015 and reviewed by privacy, technology, and legal experts in early 2016.

Partnered with Smart Chicago Collaborative to organize public meetings and online interaction via the OpenGov Madison "policy co-creation platform."

This led to a set of consensus-driven policies and governance structures, a report detailing community input, suggestions, questions, and responses from the AoT team, and ongoing engagement processes.

B.4.6.7 Government Approvals

None apparent. Approvals appear to occur within project governance framework.

[21] Charles E. Catlett et al., "Array of Things: A Scientific Research Instrument in the Public Way: Platform Design and Early Lessons Learned," in *Proceedings of the 2nd International Workshop on Science of Smart City Operations and Platforms Engineering - SCOPE '17* (the 2nd International Workshop, Pittsburgh, Pennsylvania: ACM Press, 2017), 26–33, https://doi.org/10.1145/3063386.3063771.

B.4.6.8 Public Interest and End Goals

The AoT project aimed to create a community asset, where residents and businesses saw the instrument as a public resource. Although Chicago had thousands of public cameras that garnered little attention, the AoT project aspired to go beyond “accepted” to being “embraced.”

B.4.6.9 Equity and Human Rights

No information found.

B.4.6.10 Data Collection, Categorization and Storage

Data collected included: environmental (temperature, humidity, barometric pressure, vibration, air quality, cloud cover), and traffic (pedestrian and vehicle counts).

Cameras for pedestrian and vehicle counts could collect PII through images or sound recordings, but this data was not made publicly available. Inclusion of cameras in the AoT sensor nodes was intended for detection of specific environmental conditions such as street flooding, car/bicycle traffic, storm conditions, or poor visibility. To support such capabilities, images were analyzed using an image processing computer within the node, after which the images were deleted (‘anonymization at the source’).

B.4.6.11 Consent

Commenters at final public engagement recommended that the privacy policy include a clear process for when residents believe their PII has been publicly shared accidently and would like it removed.[22]

B.4.6.12 Privacy

As the majority of data from AoT was environmental, data was published openly. Data such as weather or air quality are unlikely to have personal privacy implications. A privacy policy was, however, central to the use of cameras and microphones. The AoT privacy strategy involves three components: technical architecture and operation, transparency, and accountability

B.4.6.13 Cybersecurity and Hardware Management

No information found.

[22] Smart City 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.

B.4.6.14 Data Residency and Data Sovereignty

Further research needed, but raw calibration data would be stored in a secure facility (Argonne National Laboratory) for processing only by authorized researchers during the course of the project.

B.4.6.15 Ethical Use of Data and Technology

Access to the limited volume of data that may have included non-sensitive PII was restricted to operator employees, contractors and approved scientific partners who needed to process the data for instrument design and calibration purposes. All individuals with access to this data were subject to strict contractual confidentiality obligations and subject to discipline and/or termination if they failed to meet these obligations.

B.4.6.16 Transparency and Individual Control

Participants at public consultation identified the selection of future sensor node locations as an opportunity to involve residents and community organizations in the AoT project. There were several questions about how and if residents could be involved in selecting or informing sensor node placement. The AoT operators' online form was changed to ask for resident ideas and suggestions for node locations.

B.4.6.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

No information found.

B.4.6.18 Open Data

Data in compliance with the AoT privacy policy was made publicly available via the City's Data Portal to support application development and data analysis. All sensor data was publicly available as open data, under the stewardship of the University of Chicago.

B.4.7 SAIL (Secure Anonymised Information Linkage) Databank

B.4.7.1 Profile

- Location: Wales, United Kingdom
- Organization type: Academic & research institutions
- Level of documentation: Documented (self-reported)
- Maturity level: Established

B.4.7.2 Background

The SAIL Databank is a repository of person-based health and population records with ‘data linkage and analysis tool sets’ to help researchers. Researchers can access a range of data spanning up to 20 years from an entire population.[23]

B.4.7.3 Regulatory Framework

According to source materials, data governance complies with the following legislation:

The GDPR;

- The UK Data Protection Act;
- Also processes non-health datasets for research purposes under the provisions of the Digital Economy Act 2017.

B.4.7.4 Democratic Accountability and Transparency

The SAIL Databank receives core funding from the Welsh Government’s Health and Care Research Wales.

In 2011, SAIL Databank established a Consumer Panel. Its sixteen members are recruited from the public. Panel members are involved in all elements of the SAIL Databank process, from developing ideas, advising on bids through approval processes (via the independent Information Governance Review Panel), to disseminating research findings. The panel’s role includes:

- Acting as advisors on issues in research
- Advising on how best to engage with the public
- Offering guidance on how to recruit people to study steering groups
- Providing views on data protection issues
- Discussing proposals for research
- Reviewing information designed for a lay audience
- Acting as advocates for data linkage research

B.4.7.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

Information Governance Review Panel (IGRP): purportedly provides independent guidance and advice on Information Governance policies, procedures and processes for SAIL Databank.

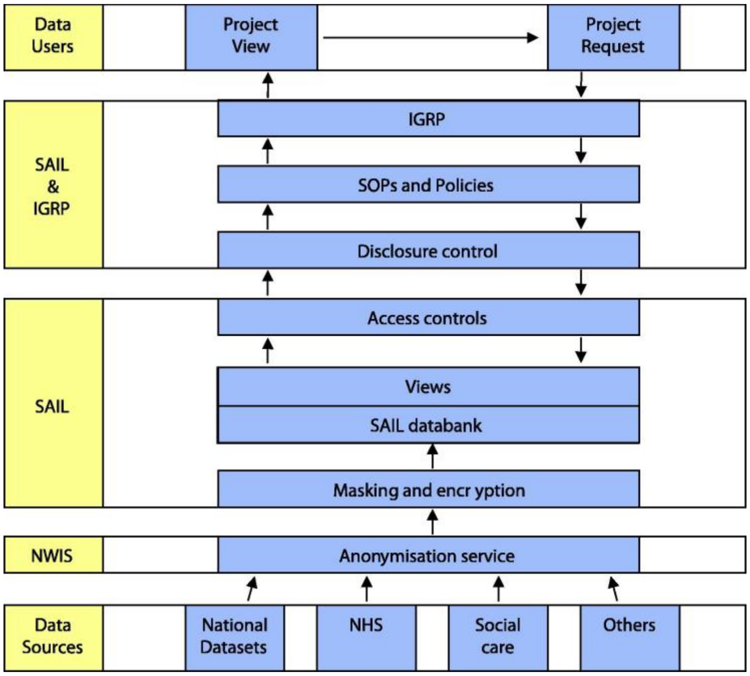
IGRP reviews all proposals to use SAIL Databank. IGRP comprises representatives from various organisations and sectors including:

[23] SAIL Databank, “SAIL Databank - The Secure Anonymised Information Linkage Databank.”

- British Medical Association (BMA) Cymru Wales
- The Welsh Government
- Public Health Wales
- National Research Ethics Service
- NHS Wales Information Service
- Swansea Bay University Health Board (formerly known as the Abertawe Bro Morgannwg University Health Board)
- The public

All access to SAIL Databank is monitored closely. Approval must be given by the independent IGRP prior to data access.

Figure B-2 SAIL Databank structure



“This diagram shows the SAIL databank system and the controls in place for data acquisition and utilisation, with an indication of the roles carried out by each party.”[24]

Beginning at the base of the diagram, SAIL has formal agreements with data providers to provide their data to the databank in accordance with Information Governance.

The commonly-recognised identifiers are anonymised at NWIS, who provide a trusted third party service to SAIL.

[24] Jones et al., “A Case Study of the Secure Anonymous Information Linkage (SAIL) Gateway.”

- Further processes of masking and encryption are carried out at SAIL, and the SAIL databank is constructed.

From the top of the diagram, requests to use the data are reviewed by SAIL and an independent Information Governance Review Panel (IGRP) to assess compliance with Information Governance before access can be allowed.

Once this is agreed, a data view is created by SAIL staff, and access to this view can be made available via the SAIL Gateway.

For this to happen, further data transformations are carried out to control the risk of disclosure, and the data user signs an access agreement for responsible data utilisation, in accordance with the specifications of the IGRP to comply with Information Governance

B.4.7.6 Public Consultation

Stakeholders (unspecified) were engaged prior to the launch of SAIL Databank in 2006.[25] It also considers its Consumer Panel (described above) as an ongoing public engagement mechanism.

B.4.7.7 Government Approvals

No information found.

B.4.7.8 Public Interest and End Goals

Its mission is to “[harness] the power of linked de-identified data to bring demonstrable improvements to people’s lives through research, evaluation, planning and policymaking.”[26]

B.4.7.9 Equity and Human Rights

No information found.

B.4.7.10 Data Collection, Categorization and Storage

Organisations provide data to SAIL Databank via the NHS Wales Informatics Service (NWIS), which acts as a Trusted Third Party which carries out anonymization and encryption.

[25] Jones, Ford, and Lyons, “The SAIL Databank: 10 Years of Spearheading Data Privacy and Research Utility, 2007-2017.”

[26] SAIl Databank, “Our Mission.”

B.4.7.11 Consent

Not required as SAIL Databank holds only anonymized data.

B.4.7.12 Privacy

SAIL Databank does not receive or handle identifiable data. They make anonymized data available for legitimate research purposes only where there is a potential for benefit.

Commonly recognised identifying details are removed before datasets come to SAIL Databank and once anonymized, they cannot be reconstructed. Because SAIL holds only anonymized data, researchers carry out their work without knowing the identities of the individuals represented in the data.

B.4.7.13 Cybersecurity and Hardware Management

Has implemented an ISO 27001 Information Security Management System (ISMS).

Implements 'residential anonymous linking fields', an identifier created from a mapping and two-step encryption process

Intent of the approach is to prevent re-identification by linking data with other databases

Efficacy of this approach remains unclear as there is no research replicating the method

Initial version required dedicated terminals to access data. 'SAIL Gateway' (online access portal) was then created to facilitate easier access

B.4.7.14 Data Residency and Data Sovereignty

Data will be held only on secure servers owned and administered by Swansea University which are subject to appropriate physical, electronic, and managerial procedures to safeguard and secure the information.

B.4.7.15 Ethical Use of Data and Technology

Process in place whereby researchers can only remove their results from the digital platform following scrutiny by a SAIL Data Guardian (i.e. an analyst). The SAIL Data Guardian assesses the proposed outputs (e.g. results tables, charts) to ensure that any risk of disclosure has been mitigated.

Researchers must undergo training before being allowed to make data requests, must have ethics approval, and permission to use health/administrative records

This suggests that part of the validation process relies upon university research standards monitoring and approval processes, including research ethics boards

B.4.7.16 Transparency and Individual Control

Because SAIL Databank holds only de-identified data and is not able to identify individuals, they are not able to process opt-out requests from members of the public, and they retain the data for long-term use.

Anyone wishing to opt out of de-identified data related to them being sent to SAIL or used for other secondary purposes can contact the relevant data provider(s) listed on our website about what options they may provide for allowing individuals to opt out.

B.4.7.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

No information found.

B.4.7.18 Open Data

No, all data access requests must be approved.

B.4.8 Consumer Data Research Centre

B.4.8.1 Profile

- Location: England, United Kingdom
- Organization type: Academic & research institutions
- Level of documentation: Documented (self-reported)
- Maturity level: Established

B.4.8.2 Background

The Consumer Data Research Centre (CDRC) is an academic led, multi-institution laboratory which brings together consumer-related datasets from around the UK. The CDRC forms part of the ESRC-funded Big Data Network and offers a data service aimed at providing researchers with access to a broad range of consumer data to address many societal challenges.

Consumer-related data are data generated by retailers and other service organizations as part of their business process. They can be used to monitor the needs, preferences and behaviours of customers. Examples of consumer-related data include:

- Sales data from till receipts
- Loyalty card and reward scheme data
- Market research data
- Travel records
- Retail turnover by store or product category
- Energy consumption meter data

The CDRC provides data with three different levels of access: open data, safeguarded data and controlled data. Access to both safeguarded and controlled data requires a process by which individuals submit project proposals for assessment and approval.[27]

B.4.8.3 Regulatory Framework

Operates under Research Data Policy of the Economic and Social Research Council

B.4.8.4 Democratic Accountability and Transparency

No information found.

[27] Consumer Data Research Centre, “CDRC - About Our Data,” CDRC, accessed September 26, 2019, https://www.cdrc.ac.uk/about-data/.

B.4.8.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

In addition to researcher registration and training, each proposed project undergoes an individual approval process. The CDRC 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 Senior Management Teams (SMT) implying at least a degree of separation from business considerations in its decision-making.

Also follows the UK Information Commissioner's Officer (ICO) Data Sharing Code of Practice.

B.4.8.6 Public Consultation

No information found.

B.4.8.7 Government Approvals

No information found.

B.4.8.8 Public Interest and End Goals

Linking consumer data with other data sources allows researchers to undertake innovative research projects that provide fresh perspectives on the dynamics of everyday life, economic well-being and social interactions in cities.

B.4.8.9 Equity and Human Rights

No information found.

B.4.8.10 Data Collection, Categorization and Storage

CDRC data is available for research purposes to a broad range of users, both internal and external to academia. Access to data is governed by a committee (Research Approvals Group/RAG) that approves requests according to a set of criteria. Some data have restrictions on access due to the data licence agreements with data providers.

B.4.8.11 Consent

No information found.

B.4.8.12 Privacy

Data sets are categorized as open, safeguarded or controlled.

Open data is freely available to all for any purpose.

Safeguarded data: data to which access is restricted because of license conditions, but where data are not considered ‘personally-identifiable’ or otherwise sensitive.

Controlled data: data which needs to be held under the most secure conditions with highly restricted access.

B.4.8.13 Cybersecurity and Hardware Management

No information found.

B.4.8.14 Data Residency and Data Sovereignty

No information found.

B.4.8.15 Ethical Use of Data and Technology

No information found.

B.4.8.16 Transparency and Individual Control

No information found.

B.4.8.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

Safeguarded or controlled data can not be used for purposes other than those originally approved by the Research Approvals Group. Open data can be used for any purpose.

B.4.8.18 Open Data

Some data sets are available as open data.

B.4.9 First Nations Data Centre

B.4.9.1 Profile

- Location: Ontario, Canada
- Organization type: Academic & research institutions
- Level of documentation: Documented (self-reported)
- Maturity level: Established

B.4.9.2 Background

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. It is notable for the OCAP (Ownership, Control, Access and Possession) principles which form the basis of its mission and to which researchers must adhere as a condition of access to their data.[28]

B.4.9.3 Regulatory Framework

Historically, researchers and government officials have entered First Nations communities and collected data (including biological samples). 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*.[29]

B.4.9.4 Democratic Accountability and Transparency

No information found.

B.4.9.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

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.**

[28] 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.

[29] 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.[30] These are articulated through the following four principles:[31]

> ***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.*

B.4.9.6 Public Consultation

See Ethical Use of Data and Technology section.

B.4.9.7 Government Approvals

No information found.

B.4.9.8 Public Interest and End Goals

The OCAP principles emphasize the group interest of the Indigenous peoples.

B.4.9.9 Equity and Human Rights

The OCAP principles are premised on the rights of the First Nations. They exist to prevent the recurrence of historical injustices, ensure benefits to the First Nations and to assert indigenous data sovereignty.

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

B.4.9.10 Data Collection, Categorization and Storage

FNIGC holds unpublished and record-level data from FNIGC's respected survey work, including the First Nations Regional Health Survey (FNRHS, or RHS) and the First Nations Regional Early Childhood, Education and Employment Survey (FNREEES, or REEES).

Data is stored in on-site servers and provided to researchers in the form of general purpose tables, special purpose tables, or record level data.

B.4.9.11 Consent

Initial consent by the indigenous community occurs at time of data collection, prior to its arrival at FNIGC. Consent to use data provided by the centre is obtained through a vetting process. Prior to any data being allowed to leave the centre the program manager will review all research outputs to ensure compliance with confidentiality policies.

B.4.9.12 Privacy

See Consent

B.4.9.13 Cybersecurity and Hardware Management

Physical measures limiting access to hardware help ensure security. Record level data are only provided on isolated computer workstations that have no internet access, storage, or the ability to plug in peripheral devices such as USB sticks. Printers are also not provided to ensure data do not leave the facility.

B.4.9.14 Data Residency and Data Sovereignty

Data is located on servers which are physically present within the FNIGC offices

B.4.9.15 Ethical Use of Data and Technology

All individual record level research requests to use the data are subject to a rigorous approval process. This includes:

Demonstrated need for the data in question

The research may not be harmful to either the survey respondents or to the First Nations communities.

Research must have demonstrable benefits to First Nations interests

B.4.9.16 Transparency and Individual Control

The FNIGC maintains control throughout a research project. In addition to raw data being located on physical servers in FNIGC offices, contractual agreements between the FNIGC and the researcher specify that, in the event of non-compliance with rules or restrictions, all unpublished research outputs (for example, papers or tables of data) must be destroyed.

The incentive for compliance with this requirement is that individuals who are not authorized to access the research output will not be able to reference it in their research legitimately.

B.4.9.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

As noted above, the OCAP Principles were developed to articulate a set of collective or community data ownership rights (through contractual agreements with parties seeking to use First Nations data).

B.4.9.18 Open Data

Open data is not available.

B.4.10 Portland Smart City PDX Program

B.4.10.1 Profile

- Location: Oregon, USA
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Emerging

B.4.10.2 Background

Under its broader Smart City PDX program, the City of Portland 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, in order to design better streets. In order to protect the privacy of residents, the City worked with the project vendors to ensure that photos are not saved and are recorded over, and that any information is anonymized. Portland City Council must approve any changes to these terms.[32]

The Smart City PDX plan is notable among smart city strategies for its explicit focus on marginalized and underrepresented communities.

B.4.10.3 Regulatory Framework

The overall framework is provided by the City of Portland Privacy and Information Protection Principles, adopted on June 19, 2019, which provide guidelines for protecting private and sensitive data managed by the City of Portland or those working on behalf of the City of Portland. These focus on:

- Transparency and accountability
- Full lifecycle stewardship
- Equitable data management
- Ethical and non-discriminatory use of data
- Data openness
- Automated Decision Systems
- Data Utility

Federal and state privacy protection laws also apply.

[32] GovEx, "First Things First: Laying the Foundation for a Smart City," May 2018, https://govex.jhu.edu/wp-content/uploads/2018/05/SMARTCITIES_GUIDE_FINAL-1.pdf.

B.4.10.4 Democratic Accountability and Transparency

Any changes to the contract terms must go through Portland City Council.

B.4.10.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

The Smart City PDX program addresses disparities using technology and improved information management. This goal is established in the Smart City PDX Priorities Framework. Portland City Council adopted this framework on June 21, 2018.[33]

This framework also stipulates that any policy, plan or project receiving Smart City PDX support must provide a detailed report on data use and privacy protection.

B.4.10.6 Public Consultation

According to the City webpage, the project principles were developed through internal and external consultations, including two public forums.

B.4.10.7 Government Approvals

The Smart City PDX Priorities Framework stipulates that (among other criteria) any policy, plan or project receiving Smart City PDX support must provide a detailed report on data use and privacy protection.

B.4.10.8 Public Interest and End Goals

The public interest is defined as making opportunities from technology available and accessible to all. This requires addressing existing disparities based on race, socioeconomic factors, and ability.

B.4.10.9 Equity and Human Rights

Its Smart City PDX Initiative Priorities Framework explicitly identifies a focus on communities of color as well as those living with disabilities.

Additionally, the Office of Equity and Human Rights co-led the development of the privacy principles.

B.4.10.10 Data Collection, Categorization and Storage

All images are deleted immediately after processing.

According to the City, data that is taken from the images is completely anonymized and only contains information about the number and speed of motor vehicles, pedestrians and bicyclists. Each sensor

[33] City of Portland, Oregon, “Exhibit A: City of Portland’s Smart City PDX Initiative Priorities Framework,” June 21, 2018.

contains a microphone to record sound. The City of Portland has elected not to turn these microphones on as another measure to preserve privacy.[34]

B.4.10.11 Consent

Has not been at issue in the Traffic Safety Sensor project, but the Privacy Principles suggest informed consent is a priority.

B.4.10.12 Privacy

No specific definition given, except that neither images nor sound will be retained by sensors.

B.4.10.13 Cybersecurity and Hardware Management

No information found.

B.4.10.14 Data Residency and Data Sovereignty

AT&T provided telecom infrastructure for the project. All the sensors are powered by the AT&T LTE network.

Current by GE supplied the 200 sensors used in this project. Current by GE's application program interfaces (APIs) are critical elements of the sensors. These APIs make it possible for City of Portland staff to make use of the sensor data.

Intel supplied processors, platform sensors and security software in each sensor.

B.4.10.15 Ethical Use of Data and Technology

No information found.

B.4.10.16 Transparency and Individual Control

No information found.

B.4.10.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

Under the agreement, all rights to the data, including all intellectual property rights, belong solely to the City, and the City retains that ownership at the end of the project.

[34] City of Portland, Oregon, "Traffic Safety Sensor Project: How Privacy Is Protected," 2019, https://www.portlandoregon.gov/transportation/76740.

B.4.10.18 Open Data

No information found.

B.4.11 Nantes Métropole Data Charter

B.4.11.1 Profile

- Location: France
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Emerging

B.4.11.2 Background

Nantes Métropole, the 6th largest metropolitan area in France (23 municipalities around Nantes and 640,000 inhabitants), is the first French region to have published a Data Charter ("Charte métropolitaine de la donnée") concerning data produced by municipal administrations, but also by private companies involved in the management of urban services (public transport, energy, water, waste, etc.) or private operators whose activity has an impact on the public space (such as Waze, Uber, etc.).[35] While not a data governance program, this data charter is an essential antecedent to data governance in the region.

This charter is part of the context of the implementation of the GDPR since May 2018.

It was developed through a three-part process:

- An internal process was conducted to disseminate a real data culture to managers and to build the community's data management doctrine.
- A citizen consultation process was organized with the establishment of a panel.
- A consultation was also conducted with private stakeholders involved in smart city projects in Nantes. These private actors were invited to sign the charter at the end of the process (nearly 50 signatories).

B.4.11.3 Regulatory Framework

GDPR; French Digital Republic Act

B.4.11.4 Democratic Accountability and Transparency

- Three groups of stakeholders: public administration, private companies delivering public services and other companies
- Internal rules define how the Chief Data Officer and the Data Protection Officer work together
- Citizen involvement is limited to the initial participation for the charter
- Accountability mechanisms are not yet implemented except for an annual public report

[35] Nantes Métropole, “La charte de la donnée métropolitaine,” accessed September 13, 2019, https://metropole.nantes.fr/charte-donnee.

- A planned public education campaign will help citizens to better understand their privacy rights.

B.4.11.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

No information found.

B.4.11.6 Public Consultation

A citizen panel that met several times to define the conditions of acceptability of data collection and data management in the public space (in the context of an experimental district).

B.4.11.7 Government Approvals

While the Nantes system is not subject to regional or national control, its implementation of the GDPR is monitored by the "Commission nationale de l'informatique et des libertés" (CNIL), the French independent supervisory authority in charge of privacy rules.

B.4.11.8 Public Interest and End Goals

The Nantes charter defines the notion of "data of metropolitan interest" and concerns all data that have an impact on the effectiveness of public policies and public services?.

B.4.11.9 Equity and Human Rights

No information found.

B.4.11.10 Data Collection, Categorization and Storage

Within the framework of the GDPR, a directory of personal data is established, from which private impact assessments are organized.

B.4.11.11 Consent

Governed via GDPR

B.4.11.12 Privacy

Governed via GDPR

B.4.11.13 Cybersecurity and Hardware Management

For services directly operated by public servants: data hosting and security control are provided internally through a certified IT security process.

For services operated by private companies: Nantes has initiated a discussion to obligate companies to apply the principles of the charter. Work in progress e.g. with energy providers and mobility companies.

B.4.11.14 Data Residency and Data Sovereignty

The objective of data sovereignty is claimed. An audit of all storage devices was carried out for more than 450 applications.

- More than 400 devices (mainly servers) are located in Nantes.
- Three are in another European country under the GDPR rules.
- Only one application exports personal data outside the European Union. The transfer is governed by the GDPR conventions (in this case the data is in Canada).

B.4.11.15 Ethical Use of Data and Technology

The GDPR limits the use of data to applications for which consent has been given.

The Nantes Charter limits use to the interests of public action. This principle is not currently the subject of a review process

B.4.11.16 Transparency and Individual Control

In the GDPR context individuals can be informed by asking directly the DPO directly:

These rules also concern the data of employees of the city. The integration of the GDPR into human resources management is a particular challenge.

B.4.11.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

No information provided.

B.4.11.18 Open Data

Public ownership of data produced for urban management.

B.4.12 Argentina-Microsoft Partnership, AI Tools for Public Policy

B.4.12.1 Profile

- Location: Argentina
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Emerging

B.4.12.2 Background

In Argentina, the government of the Province of Salta partnered with Microsoft to implement artificial intelligence tools in order to better understand the factors contributing to school dropouts and teenage pregnancies. The artificial intelligence models use data provided by the Ministry of Early Childhood. However, a lack of transparent communication led to mistrust on the part of residents.

B.4.12.3 Regulatory Framework

The data used in the AI model were protected from being publicly available by Personal Data Protection Law 25,326

Argentina did not have an open data law at the time.

B.4.12.4 Democratic Accountability and Transparency

Accountability and transparency were seriously inhibited as the inner workings of the artificial intelligence model would tend to be unintelligible to most individuals.

B.4.12.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

No information found.

B.4.12.6 Public Consultation

No apparent public consultation.

B.4.12.7 Government Approvals

No information found.

B.4.12.8 Public Interest and End Goals

Public interest in this case is defined as being able to identify at-risk youth and coordinate interventions with the appropriate Social Service Agency.

B.4.12.9 Equity and Human Rights

Significant equity considerations were raised in the case of the training data used in the model which came from individuals residing in low-income areas. This might predict disproportionately higher numbers of cases of teen pregnancy or school dropouts among other groups or areas.

B.4.12.10 Data Collection, Categorization and Storage

The training data for the algorithm was collected from predominantly low-income areas in the city of Salta in 2016 and 2017. According to the Web Foundation report, the methodology for the collection of these data was not made publicly available.

B.4.12.11 Consent

No information found.

B.4.12.12 Privacy

No information found.

B.4.12.13 Cybersecurity and Hardware Management

While not explicitly stated, it is possible that the data and AI model were hosted in cloud infrastructure operated by Microsoft, with servers potentially outside Argentina's borders.

B.4.12.14 Data Residency and Data Sovereignty

No information found.

B.4.12.15 Ethical Use of Data and Technology

According to the Web Foundation, the black box nature of the algorithm processes have the effect of reducing public trust that data are being used for ethical purposes, since there is no way to trace a given output to its corresponding input.

B.4.12.16 Transparency and Individual Control

According to the report, transparent communication on this project was lacking. The authors recommended that the government provide quarterly reports on the impact of the AI model which would help residents better understand how their data was being used and what the outcomes were.

B.4.12.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

No information found.

B.4.12.18 Open Data

No information found.

B.4.13 Barcelona Municipal Data Office

B.4.13.1 Profile

- Location: Spain
- Organization type: Government & public agencies
- Level of documentation: Documented (self-reported)
- Maturity level: Emerging

B.4.13.2 Background

The City of Barcelona set up its Municipal Data Office (MDO) in 2017, based on direction from the City Council, in order 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.

B.4.13.3 Regulatory Framework

According to their website, data governance complies with the following legislation:

The GDPR

- The Organic Law 3/2018, of December 5, on Data Protection and Guarantee of Digital Rights

Spanish Law 39/2015 on Common Administrative Procedures for Public Authorities which states that by 2020, digital channels must take priority in the provision of public services in the Spanish State.

B.4.13.4 Democratic Accountability and Transparency

There are several entities having various responsibilities with the data governance framework:

- Executive Committee of Data: sets high level strategic and tactical direction for the MDO
- Transversal Data Coordination Board: acts as a venue for coordinating data projects across City departments as well as establishes, applies, and monitors the work carried out by the MDO
- Data Protection Table: ensures compliance with GDPR

The Municipal Data Office is accountable to City Council through the Chief Data Officer. The MDO is responsible for implementing the data governance model, based on the Executive Committee’s overall direction. Its main objectives are:

- Define and coordinate a municipal data governance model, from the definition of a corporate data management strategy.
- Ensure compliance with the standards and regulations outlined in the data strategy
- Facilitate the alignment of technological tools to needs of use.

- Introduce data analytics (data science) to City Council as a tool aimed at deepening the knowledge of citizens and their needs, city management, internal management, and risk management.
- Strengthening the sovereignty of data, dissemination, availability and transparency of the set of municipal data and the promotion of open data.
- The exchange of knowledge about policies, standards and good practices for data management through active dialogue with citizens and consumers of data, other producers, public / private corporations and centers of research.
- The promotion of maintenance, reuse, improvement of accessibility and enrichment of data.

B.4.13.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

The Ethical Digital Standards Toolkit provides multiple guiding documents, including ICT Procurement, Technology Sovereignty, and Ethical Data Management.

B.4.13.6 Public Consultation

Consultation took place during development of the Digital City Plan. The City of Barcelona also uses a citizen engagement platform called DECIDIM on an ongoing basis, though not specifically for data-related issues.

B.4.13.7 Government Approvals

No information found.

B.4.13.8 Public Interest and End Goals

The MDO’s mission is to transform the internal culture of the Barcelona civil service, to promote standardization in data management protocols, and promote good data sharing practices. The public interest is articulated as providing support for data-driven decision making within Barcelona City Council and its associated entities.[36]

B.4.13.9 Equity and Human Rights

Equity and human rights are not specifically named, however elements of these concepts could be addressed through Barcelona’s focus on Digital Commons and Technological Sovereignty.

B.4.13.10 Data Collection, Categorization and Storage

These policies are not explicitly addressed, possibly due to the fact that this is a high-level strategy.

[36] Ajuntament de Barcelona, “Directive Concerning Municipal Data Governance and the Municipal Data Offices.”

B.4.13.11 Consent

Subject to consent mechanisms provided for by GDPR.

Through the DECODE project, a pilot tested the concept of granular data sharing permissions with privacy enhancing technologies (PETs) and IoT devices, in order to give citizens control over which data is shared from their device.[37] However, this project has not been rolled out more widely yet.

B.4.13.12 Privacy

Privacy is defined by reference to the legal framework under which it operates (i.e., GDPR, Organic Law 3/2018).

B.4.13.13 Cybersecurity and Hardware Management

The City proposes to use hardware resources controlled by the City itself, adopting appropriate technical and organizational measures to ensure the protection of their citizens' and visitors' data and privacy.

B.4.13.14 Data Residency and Data Sovereignty

Data sovereignty not explicitly named. Barcelona has included the principle of technological sovereignty in its Digital City Plan, conceptualized in three categories: transition to and use of free software, interoperability of services and systems, and use of free standards.

B.4.13.15 Ethical Use of Data and Technology

As mentioned above (Supporting Policies), Barcelona has a set of Ethical Data Standards.

B.4.13.16 Transparency and Individual Control

Individual control of personal data is provided for in Barcelona's conception of digital rights.

B.4.13.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

Data commons approach: "City residents and the common good have to be the central focus of the Municipality of Barcelona's plans and technological platforms. Data is a source of wealth that empowers people who have access to it. Making it possible for city residents to control the data, minimising the digital gap and preventing discriminatory or unethical practices is the essence of municipal technological sovereignty."

[37] Bria et al., "Deployments of Pilots in Barcelona."

B.4.13.18 Open Data

Open data is available through the BCN Portal. Open source is also an important principle of the Digital Plan. Barcelona has plans to publish all non-confidential or private data generated by municipal ICTs as open data and provide platforms for other entities to do the same.

B.4.14 Los Angeles Department of Transportation

B.4.14.1 Profile

- Location: California, USA
- Organization type: Government & public agencies
- Level of documentation: Low-level documentation
- Maturity level: Emerging

B.4.14.2 Background

The Mobility Data Specification (MDS) – a standard for e-scooter and private transportation company data – began as an internal project of the Los Angeles Department of Transportation (LADOT) before responsibility for maintaining the code base was transferred to an open-source software foundation called the Open Mobility Foundation.[38] It has since been adopted by several other cities. However, 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.[39]

B.4.14.3 Regulatory Framework

According to source material, the MDS complies with the California Consumer Protection Act.

B.4.14.4 Democratic Accountability and Transparency

Not specified. There does not appear to be any democratic mechanism through which residents can hold LADOT directly accountable. However, LADOT has stated that after completion of the Dockless Mobility Pilot, they will create a publicly accessible transparency report discussing the types of third party requests for Dockless Mobility data that LADOT has received and how they have responded to those requests.

B.4.14.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

LADOT claims its implementation of MDS also complies with the City of Los Angeles Information Handling Guidelines and its own Data Protection Principles.

B.4.14.6 Public Consultation

Not specified.

[38] Los Angeles Department of Transportation. “Mobility Data Specification,” October 31, 2018. https://ladot.io/wp-content/uploads/2018/12/What-is-MDS-Cities.pdf.

[39] 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.

B.4.14.7 Government Approvals

Not specified.

B.4.14.8 Public Interest and End Goals

LADOT articulates its public interest in terms of delivering “a safe, livable, and well-run transportation system throughout the region.”

B.4.14.9 Equity and Human Rights

LADOT states that it is committed to preventing discrimination and protecting personal mobility data of its users.

B.4.14.10 Data Collection, Categorization and Storage

LADOT designates raw trip data as Confidential Information under the City of Los Angeles Information Handling Guidelines.

B.4.14.11 Consent

No consent mechanism specified. As LADOT collects data from dockless mobility providers through MDS, presumably they are relying on users of those services having consented by agreeing to their terms of service.

B.4.14.12 Privacy

LADOT states that they employ a variety of privacy enhancing techniques to protect privacy throughout the data lifecycle, including aggregation, obfuscation, de-identification, and destruction.

B.4.14.13 Cybersecurity and Hardware Management

Security practices are not specified, except that they comply with the City of Los Angeles Information Security Policy Manual (not found).

B.4.14.14 Data Residency and Data Sovereignty

Not specified

B.4.14.15 Ethical Use of Data and Technology

Not specified. However, LADOT has stated that they will not make personal mobility data available to law enforcement except in the case of a court order, subpoena or other legal process.

B.4.14.16 Transparency and Individual Control

LADOT publishes a list of the data types collected via the MDS and the length of time that data is retained.

B.4.14.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

Not specified.

B.4.14.18 Open Data

De-identified data is shared on the City’s Open Data Portal.

B.4.15 Estonian Data Embassies

B.4.15.1 Profile

- Location: Estonia
- Organization type: Government & public agencies
- Level of documentation: Low-level documentation
- Maturity level: Emerging

B.4.15.2 Background

Estonia is widely considered to be one of the most technologically integrated and advanced governments in the world. Under its e-Estonia initiative, digital services that have been implemented include i-Voting, e-Tax Board, e-Business, e-Banking, e-Ticket, e-School, University via internet, the e-Governance Academy. This digital shift also means that Estonia is very reliant on its ICT infrastructure, the main component of which is its decentralized X-Road data platform that links up different services. In order to provide a measure of redundancy and continuity in the event of digital infrastructure failure, Estonia is testing what it calls “data embassies.” These are network servers which, although located outside of Estonia, are nonetheless governed by its laws. The first of these is located in Luxembourg with plans for others in the future.

B.4.15.3 Regulatory Framework

GDPR; Bilateral agreement between Estonia and Luxembourg which guarantees immunity for the data embassy.

B.4.15.4 Democratic Accountability and Transparency

No information found.

B.4.15.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

No information found.

B.4.15.6 Public Consultation

No information found.

B.4.15.7 Government Approvals

In order to maintain a database, even if one does not share data, an organization (data service provider) must register with Estonia's Information System Authority, which maintains a list of data service providers and assigns codes, numbers and certificates for identification and authentication.

B.4.15.8 Public Interest and End Goals

The public interest is defined in terms of integrated information systems which enable easy access to public services.

B.4.15.9 Equity and Human Rights

No information found.

B.4.15.10 Data Collection, Categorization and Storage

Data are collected or generated in the course of state functions by data service providers. All databases must be registered with the government.

The key feature of Estonia's Data Embassy is that storage is decentralized. KSI Blockchain technology is being explored to maintain data integrity while keeping data in sync. All data manipulations are tracked and data cannot be overwritten.[40]

B.4.15.11 Consent

No information found.

B.4.15.12 Privacy

Data confidentiality is built into the system protocols.

B.4.15.13 Cybersecurity and Hardware Management

X-tee is the data exchange layer used in Estonia’s digital systems.[41] It employs cryptographic certificates to prevent unauthorized access and also allows data owners to enforce their own access conditions.

B.4.15.14 Data Residency and Data Sovereignty

By hosting redundant data in networked servers out-of-country, Estonia is attempting to mitigate risks associated with equipment failures or cybersecurity attacks on a centralized infrastructure.

B.4.15.15 Ethical Use of Data and Technology

No information found

[40] e-Estonia, “Security and Safety,” e-Estonia, accessed October 1, 2019, https://e-estonia.com/solutions/security-and-safety/.

[41] “Introduction of X-Tee | Estonian Information System Authority,” accessed October 1, 2019, introduction-x-tee.html.

B.4.15.16 Transparency and Individual Control

No information found

B.4.15.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

No information found.

B.4.15.18 Open Data

No information found.

B.4.16 Data Ventures

B.4.16.1 Profile

- Location: New Zealand
- Organization type: Government & public agencies
- Level of documentation: Low-level documentation
- Maturity level: Emerging

B.4.16.2 Background

Launched in 2018, Data Ventures is a business unit of Stats NZ, New Zealand's official data agency (explicitly naming itself as a commercial arm of Stats NZ, a data trust, and a data brokerage), which functions as a trusted intermediary that pulls 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.

After investigating several potential ventures, Data Ventures chose to focus on its Population Density venture, which collects anonymized data-sets from the large telecom companies in New Zealand and then provides snapshots of population density at given locations in New Zealand.

Stats NZ has taken the novel step of measuring its social licence, 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.

B.4.16.3 Regulatory Framework

Stats NZ is required to protect the information it collects by the Statistics Act 1975 (individual, household and business data) and the Privacy Act 1993 (personal information) which has twelve information privacy principles.

The supply of data from private sector companies to Data Ventures is governed by a memorandum of understanding.

B.4.16.4 Democratic Accountability and Transparency

Ultimately accountable to national government through the Minister of Statistics. Complaints can be directed to the Office of the Privacy Commissioner.

B.4.16.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

Stats NZ has a Chief Privacy Officer who is supported by a Senior Advisor, Privacy. The Chief Privacy Officer chairs the Information Privacy, Security, and Confidentiality governance group whose role is to ensure full compliance with the NZ Privacy Act in all activities. This includes monitoring, evaluating, and reporting on compliance. The Senior Advisor, Privacy handles any complaints and works closely with the Office of the Privacy Commissioner, on behalf of the Chief Privacy Officer, to ensure best practices are adhered to in Stats NZ's work.

Conducts privacy impact assessments (PIA) which focus on identifying the ways a new proposal or operating system, or changes to an existing process may affect personal privacy.

B.4.16.6 Public Consultation

Public and businesses were consulted (e.g., through raising awareness of the new venture).

B.4.16.7 Government Approvals

No information found.

B.4.16.8 Public Interest and End Goals

From their website: "For each product we bring [multiple datasets] together while making that data confidential, complete and ready to sell to our customers to help them make fully informed, data-driven decisions."

B.4.16.9 Equity and Human Rights

No information found.

B.4.16.10 Data Collection, Categorization and Storage

Stats NZ regularly collects information through censuses and other survey instruments and also links this information to data sets received from other agencies.

This data is generally stored in integrated research databases.

Initially for the population density venture, the data provided by telecom companies will be hosted in a cloud environment with password access. After being aggregated by Data Ventures, the data will be stored in a separate protected facility.

B.4.16.11 Consent

As the data is routinely collected by telecom companies over the course of service provision to customers, no explicit consent is required for its transfer to Data Ventures.

B.4.16.12 Privacy

According to their website, Stats NZ commits to protecting New Zealanders' privacy by:

Collecting and keeping only the information needed for a specific purpose

Guarding against unauthorized access

Withholding personally identifiable information

According to a PIA for its population density data venture, the privacy impact will be estimated by the following attributes: hourly time range, statistical area or suburb, and count.

Data will not be allowed to be personally-identifiable.

B.4.16.13 Cybersecurity and Hardware Management

No information found.

B.4.16.14 Data Residency and Data Sovereignty

Data are stored securely on Stats NZ servers

B.4.16.15 Ethical Use of Data and Technology

Since Data Ventures is a business unit at a public agency, data collected is required to be used for the public benefit.

B.4.16.16 Transparency and Individual Control

In the Data Ventures partner model, each party negotiates its contributions, invests accordingly and receives a share of returned value.

Price setting method remains unclear

B.4.16.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

No information found.

B.4.16.18 Open Data

No information found.

B.4.17 Liberian telecommunications authorities

B.4.17.1 Profile

- Location: Liberia
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Nascent

B.4.17.2 Background

During the 2014 West African Ebola outbreak, humanitarian organizations attempted to coordinate their responses and predict the spread of disease. A group of actors from the humanitarian aid sector called for the use of aggregated location data (Call Detail Records) collected from local cell phone towers as a means of expediting the laborious process of tracking down everyone who had come into contact with a sick person. Critics claimed that this method was only suitable for tracking vector-borne illnesses such as malaria, and not effective for understanding the complexities of human movement. 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. The primary source for this case study is a Centre for Internet and Society report by Sean McDonald.[42]

B.4.17.3 Regulatory Framework

In Liberia, the use of call detail records is illegal for non-governmental actors. The Telecommunications Act of 2007 provides for punishment of privacy violations.

B.4.17.4 Democratic Accountability and Transparency

No information found.

B.4.17.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

No information found.

B.4.17.6 Public Consultation

Two conceptions of the public interest were at odds with one another: containing the spread of disease and protecting personal privacy of individuals.

[42] Sean McDonald, “Ebola: A Big Data Disaster: Privacy, Property, and the Law of Disaster Experimentation,” CIS Papers (Center for Internet and Society, n.d.), http://cis-india.org/papers/ebola-a-big-data-disaster.

B.4.17.7 Government Approvals

No information found.

B.4.17.8 Public Interest and End Goals

No information found.

B.4.17.9 Equity and Human Rights

According to McDonald, this use of data may be in breach of international privacy law, e.g., Article 12 of the Universal Declaration of Human Rights which gives people the ability to challenge "arbitrary interference with their privacy"

B.4.17.10 Data Collection, Categorization and Storage

Call detail records can be collected in two ways.

By triangulating a person's location based on their proximity to the closest mobile network towers. The cell site location information method is more reliable where many towers are located close to one another.

By tracking location data via GPS

B.4.17.11 Consent

Consent was at issue in the attempt to use call detail records for tracing sick individuals. Liberian law requires telecommunications providers to explicitly communicate which purposes customer data will be used for.

B.4.17.12 Privacy

No information found.

B.4.17.13 Cybersecurity and Hardware Management

No information found.

B.4.17.14 Data Residency and Data Sovereignty

No information found.

B.4.17.15 Ethical Use of Data and Technology

In this case response organizations pushed for the release of data for use in untested models which likely would not have been useful for dealing with the outbreak.

For comparison, McDonald points to the case of the 2015 South Korean outbreak of Middle East Respiratory Syndrome, which killed 36 people. The government of South Korea preemptively used algorithms to identify – without direct evidence – potentially infected people and impose quarantines on them.

B.4.17.16 Transparency and Individual Control

No information found.

B.4.17.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

No information found.

B.4.17.18 Open Data

No information found.

B.4.18 New York City Automated Decision Systems Task Force

B.4.18.1 Profile

- Location: New York, USA
- Organization type: Government & public agencies
- Level of documentation: Documented (independently)
- Maturity level: Nascent

B.4.18.2 Background

New York City became one of the first jurisdictions to pass a law on automated decision systems (ADS).[43] Automated decision systems are technical systems that aim to aid or replace human decision making and are increasingly being deployed in areas such as criminal justice, education, social work, and policing. The ADS Task Force was concerned primarily with the most complex systems whose decisions would have the greatest impact on an individual's job prospects, financial outcomes, or similar opportunities. The ADS Task Force final report was released in November 2019. However, an independent 'shadow report' compiled by the AI Now Institute - which provides a comprehensive alternative record of events over a two-year period - cites the City's reluctance to provide a list of known automated systems in use as a major hindrance to the work of the Task Force.

As such the work of the ADS Task Force stalled and did not go beyond the broad recommendations made in its report.

B.4.18.3 Regulatory Framework

In 2018, NYC passed a law which required the creation of a task force that provides recommendations on how information on agency automated decision systems may be shared with the public and how agencies may address instances where people are harmed by agency automated decision systems.

B.4.18.4 Democratic Accountability and Transparency

An independent 'shadow report' compiled by the AI Now Institute - which provides a comprehensive alternative record of events over a two-year period - cites the City's reluctance to provide a list of known automated systems in use as a major hindrance to the work of the Task Force.

B.4.18.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

Risk would be assessed according to the following criteria:

[43] https://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0

Whether the proposed system would be “continual and operationalized” or would be utilized on a “one-off” basis;

The number of data subjects who would potentially be affected by the system;

The set of negative impacts that could result from the system (e.g. loss of job opportunity, negative credit impacts, etc.); and

The duration of these negative impacts

B.4.18.6 Public Consultation

According to the ADS Task Force report, public engagement activities took place in April 2019. Recommendations were provided by invited panelists, independent organizations, and the general public.

However, the AI Now Institute’s shadow report criticized the consultation process for only taking place after concerns were raised by a coalition of civil rights advocates, researchers, community organizers, and concerned residents around the lack of public engagement. They further assessed the public engagement efforts were inadequate, and charging that “information about the meetings [was] not provided to the public on the ADS Task Force website, and avenues for public participation [were] never made clear.”

B.4.18.7 Government Approvals

The ADS Task Force was charged with recommending criteria for identifying which agency automated decision systems should be subject to one or more of the procedures recommended by them.

B.4.18.8 Public Interest and End Goals

Based on the ADS Task Force Report, the public interest is implicitly defined as understanding how the City uses, manages, retains information about, and answers public questions about automated decision systems in order to reduce biases and ensure equitable outcomes.

B.4.18.9 Equity and Human Rights

Not specified.

B.4.18.10 Data Collection, Categorization and Storage

The ADS Task Force recognized the need for standards and protocols in this area but did not specify precisely which ones.

B.4.18.11 Consent

Not specified.

B.4.18.12 Privacy

Not specified.

B.4.18.13 Cybersecurity and Hardware Management

Not specified.

B.4.18.14 Data Residency and Data Sovereignty

Not specified.

B.4.18.15 Ethical Use of Data and Technology

Not specified.

B.4.18.16 Transparency and Individual Control

Not specified.

B.4.18.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

Not specified.

B.4.18.18 Open Data

No information found.

B.4.19 France’s National Health Data Hub

B.4.19.1 Profile

- Location: France
- Organization type: Government & public agencies
- Level of documentation: Low-level documentation
- Maturity level: Nascent

B.4.19.2 Background

The "Health Data Hub" is the French government's instrument for sharing health data and securing access to it. It is in the process of being created and will take one of two legal forms: a public institution or a public interest group.

The Health Data Hub is conceived as a "trusted third party," and it will ensure both ethical uses and quality of the data.

It will connect data producers with public or private users. It will provide a one-stop shop for all national health data supported by national solidarity, support accreditation procedures and carry out matching operations to make documented data sets available in a timely manner.

It is intended to ensure transparency towards civil society and citizens through a portal to consult available data sources and their use.

B.4.19.3 Regulatory Framework

GDPR and complementary national laws specifically protecting health data.

France operates a security reference system applicable to the National Health Data System. All public or private bodies that host and process health data must be accredited.

B.4.19.4 Democratic Accountability and Transparency

Three groups of actors are defined: data producers, data users and citizens. The roles will be clarified by Charters that have not yet been drafted.

Data producers make data available according to principles defined by a "Producer Charter". The "Producer" charter will describe in particular the governance of data access and the requirements common to all databases in terms of data quality and structuring.

Users use Hub services according to the principles defined by a "User Charter". It will cover the conditions and rules for access to the service, as well as the modalities for sharing research results. These principles will be based on clear rules on the protection of personal data (GDPR) on the one hand, and on intellectual and industrial property on the other.

Citizens contribute to ensuring transparency in accordance with the Hub's commitments under the "Citizen Charter". The "Citizen" Charter will formalise the commitments of transparency, ethics and respect for fundamental rights that all Hub stakeholders make to citizens and civil society.

B.4.19.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

No information found.

B.4.19.6 Public Consultation

A working group led by three experts and composed of about twenty experts in health data was organized.

Consultation with more than 100 health and data experts (academics, companies, laboratories, interest groups, etc.).

B.4.19.7 Government Approvals

No information found.

B.4.19.8 Public Interest and End Goals

The Health Data Hub will be a trusted third party for health data which will:

- Provide a one-stop shop to facilitate access to data.
- Ensure quality of the data that will be shared.
- Play an essential role in promoting innovation with the development of an environment in which innovation can flourish.
- Allow the pooling of technological and human resources within a secure technological platform.

B.4.19.9 Equity and Human Rights

No information found.

B.4.19.10 Data Collection, Categorization and Storage

The Hub will use several data sources: from the NHDS (National Health Data System) database to more scattered data such as data from (general practitioners or specialists) or laboratory data. Other targeted data include data from the DMP (Shared Medical Record) or provided by the patients themselves.

B.4.19.11 Consent

The secondary re-use of health data requires, in principle, individual information to be given to the data subjects on the purposes of this new processing operation, where possible, to enable them to exercise a right to erase their data. However, the GDPR allows this information requirement to be overridden if it is impossible or would require a disproportionate effort.

It also allows an exemption from the right of erasure (provided for by the GDPR for other types of processing) for reasons of public interest, or for research purposes, insofar as this right is likely to make impossible or seriously compromise the achievement of the processing objectives.

For these activities, the authorization of France's national data protection agency, *la Commission nationale de l'informatique et des libertés* (CNIL),[44] is required. The CNIL's Methodology for Privacy Risk Management 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;
- Measures are used to treat risks in a proportionate manner.

B.4.19.12 Privacy

The data is hashed – an algorithm generates data of a fixed length from the original string of characters.

B.4.19.13 Cybersecurity and Hardware Management

The technological platform and associated processes must comply with the requirements related to the hosting of health data and comply with the SNDS security reference framework (decree of 22 March 2017 on the security reference framework applicable to the National Health Data System).

B.4.19.14 Data Residency and Data Sovereignty

Data sovereignty is an objective, but there is no evidence of its implementation as of yet.

B.4.19.15 Ethical Use of Data and Technology

No information found.

B.4.19.16 Transparency and Individual Control

In the context of the GDPR, individuals can be informed by directly contacting the Data Protection Officer

The re-use of health data is possible without consent provided that the processing is in the public interest and that appropriate safeguards for the rights and freedoms of the data subjects are in place.

[44] Cuggia and Combes, "The French Health Data Hub and the German Medical Informatics Initiatives," 430.

B.4.19.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

The GDPR provides that health data is not the property of the producer.

B.4.19.18 Open Data

The technology platform will use open source solutions. A version management tool such as Github will be used to share algorithms.

B.4.20 Japanese Information Banks

B.4.20.1 Profile

- Location: Japan
- Organization type: Private sector
- Level of documentation: Low-level documentation
- Maturity level: Nascent

B.4.20.2 Background

In Japan, information banks have a similar objective as data trusts to protect data but use a different mechanism. With information banks, individuals would be able to deposit their information with a trusted third-party, decide how the information is shared with third parties and 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.[45]

B.4.20.3 Regulatory Framework

An appropriate regulatory framework is still under development. As of 2018, the Japanese government has established a study group for the guidelines of certification schemes concerning functions of information trusts.[46]

B.4.20.4 Democratic Accountability and Transparency

No information found. As information banks are to be hosted by financial institutions, it is unclear whether there would be accountability mechanisms in place.

B.4.20.5 Supporting Policies, Principles, Frameworks and Risk Mitigation Strategies

No information found.

B.4.20.6 Public Consultation

No information found.

[45] D.A. Consortium, “Pilot Testing Begins on an ‘Information Bank,’ a New System for Storing Personal Data | Online Advertising DAC,” DAC, accessed October 1, 2019, http://www.dac.co.jp/english/index.php/press/2018/20180910_ib.

[46] Japan. Ministry of Economy, Trade and Industry, “Release of the Guidelines of Certification Schemes Concerning Functions of Information Trust Ver. 1.0,” June 26, 2018, https://www.meti.go.jp/english/press/2018/0626_002.html.

B.4.20.7 Government Approvals

No information found.

B.4.20.8 Public Interest and End Goals

No information found.

B.4.20.9 Equity and Human Rights

No information found.

B.4.20.10 Data Collection, Categorization and Storage

No information found.

B.4.20.11 Consent

In a proof of concept by Mitsubishi UFJ Trust, individuals chose which data to provide via an app.

B.4.20.12 Privacy

No information found.

B.4.20.13 Cybersecurity and Hardware Management

No information found.

B.4.20.14 Data Residency and Data Sovereignty

No information found.

B.4.20.15 Ethical Use of Data and Technology

No information found.

B.4.20.16 Transparency and Individual Control

Users can voluntarily disclose their medical history habits or recent activities but may specify what purposes it can be used for.

B.4.20.17 Intellectual Property Rights, Ownership and Equitable Distribution of Value

The Mitsubishi pilot users earned 500 to 1000 yen per month from each company with whom they chose to share data.

B.4.20.18 Open Data

No information found.