Safeguarding Big Data Captured in Public Spaces through Standardization

Michel Girard

Key Points

→ Smart cities need high-quality data from public spaces in order to succeed. Significant growth is expected in the number of Internet of Things (IoT) devices installed in public spaces; however, there is a lack of interoperability standards to manage this data.

→ Although big data initiatives in smart cities are designed to improve citizens’ quality of life, data collected in public spaces can also erode fundamental rights and freedoms. Social acceptance is critical for these initiatives to succeed and new voluntary standards asserting that data captured in public spaces is a public good are needed.

→ Standards would frame approaches to engage citizens on acceptable levels of monitoring and surveillance in public spaces and define interoperability requirements for data collection, access and secondary uses by third parties.

Introduction

This policy brief proposes the development of voluntary standards for safeguarding big data captured in public spaces. Broad recommendations are made regarding the main features of the proposed standards for consideration by interested parties. The standards would help municipalities frame their engagement with citizens when designing smart cities projects, provide guidance on the collection and organization of data generated in public spaces and help ensure data is accessible for research and analytics while asserting appropriate intellectual property (IP) rights on data sources.

New voluntary standards are needed for managing big data captured in public spaces. They are a necessary pre-condition for the smooth functioning of smart cities. In the current landscape, the deployment of monitoring and surveillance technologies in public spaces is occurring in an ad hoc fashion. While big data from public spaces is bound to make cities smarter, there are mounting concerns regarding the impacts of increased monitoring and surveillance activities on fundamental human rights and privacy. The proposed standards would set a process for a dialogue between citizens and municipalities on the “acceptable” density of monitoring and surveillance equipment in public spaces, the nature of the data being collected, the priorities in responding to citizens’ needs and data use in general.
With the introduction of fifth-generation (5G) technologies, the number of sensors and cameras installed in public spaces is projected to grow. Data transmission costs are anticipated to go down, along with data storage costs. However, without standardization, interoperability objectives that are fundamental to smart cities will remain elusive. Data capture and creation initiatives are typically designed by an individual organization for one particular purpose. The potential for secondary use will only be unlocked when interoperability standards are implemented. Standards can help remove the interoperability barriers that prevent secondary uses of data. Voluntary standardization is seen by many as a viable approach to developing frameworks that are replicable across sectors. Other issues such as cyber security, data residency and adherence to applicable laws need to be addressed. Finally, the IP and ownership of the monitoring and surveillance equipment and the data it generates, as well as the analytics and the solutions, vary significantly from one smart city initiative to another and could benefit from standardization.

Broad adoption of standards generally results in the creation of a level playing field, interoperability, predictability and consistency while lowering implementation costs. As indicated in the CIGI paper *Big Data Analytics Need Standards to Thrive: What Standards Are and Why They Matter* (Girard 2019), thousands of voluntary standards are incorporated by reference in federal and provincial regulations and in municipal by-laws as well as in procurement documents. The proposed standards could be used by regulatory authorities when drafting and implementing new by-laws and regulations aimed at creating new smart cities projects involving the capture and use of big data in public spaces.

For the purpose of the proposed standards, four fundamental assumptions are made. They set the stage for a citizen-centric approach to big data captured in public spaces:

→ Data stemming from activities taking place in public spaces (including but not limited to streets, parks, public buildings and institutions and energy and water distribution systems) have similar attributes to other public goods and should be treated and managed as such. By asserting that data from public spaces is a public good, local authorities (and their smart cities partners) will be in a
better position to assert IP rights and to set requirements regarding its use and access.

→ Participants in smart city initiatives have the ability to generate anonymized data from public spaces.

→ Participants also have the ability to pool data from various sources, including public spaces.

→ Eligible private sector firms, policy makers and academics have an interest in accessing data captured in public spaces for the purposes of data analytics, academic research, improving policy outcomes and the development or commercialization of new products and services.

Scope and Intended Users

The standards would focus on data captured in public spaces, including facilities owned by municipalities and municipal infrastructure. They could cover data generated from the delivery of municipal services, such as public transport and garbage collection. In addition, activities in public spaces undertaken by private sector organizations that require municipal permits, such as restaurants, taxis and ride-sharing services, could also be covered if the standards are incorporated in relevant regulations, by-laws and procurement documents. The standards would cover three distinct stages: citizen engagement; data collection and organization; and data access, sharing and retention.

The potential number of intended users could be large, starting with the authorities responsible for public spaces, facilities, lands and infrastructure. Monitoring and surveillance devices used in public spaces would include all types of sensors and cameras designed or primarily intended to collect, retain, process or share audio, electronic, visual, location, thermal, olfactory or similar information.

Citizens have a significant stake in public spaces. The fundamental freedoms of association, expression, religion or peaceful assembly are often expressed in public spaces. Data captured in public spaces can improve our quality of life, enhance our health and help protect our environment through enhanced monitoring and management of traffic and transportation systems, power plants, water supply networks, waste management and information systems. Appropriate monitoring and surveillance of public spaces can also result in reduced crime and better schools, libraries, hospitals and community services.

On the other hand, without agreed-to standards framing their use, monitoring and surveillance devices could end up eroding citizens’ privacy or other fundamental rights. For example, by using existing technology, inexpensive software and publicly available data, it is now feasible to recognize an individual in a public space through facial recognition software, track their movement through sensors and keep a detailed record of that person’s actions and conversations in public spaces (Gladstone 2018). If they fall into the wrong hands, monitoring and surveillance devices and the data they generate could ultimately pose an existential threat to our individual freedoms and to the personal safety of specific categories or groups. According to Bianca Wylie (2019), citizens should not only know how and why data is collected in public spaces, they should have a meaningful role in the design, planning and execution of data collection initiatives taking place in public spaces, because both public spaces and the data they generate belong to them. These core issues need to be addressed before secondary uses for data generated in smart cities can be explored (ibid.).
Private sector firms would have a definite interest in the development of standards required for big data initiatives in smart cities to work. Canada’s digital industry is highly dependent on standardization in order to compete against tech giants. Although our digital industry sector is already bigger than primary industries such as mining, forestry and oil and gas (Evans 2019), a recently published digital industries report from Canada’s Economic Strategy Tables notes that it is highly fragmented. Almost all of Canada’s 40,000 information and communications technology firms are small and micro-sized (98.6 percent) and 85 percent of them have fewer than 10 employees (Canada’s Economic Strategy Tables 2018, 5). At the moment, there are no individual players in Canada, public or private, with the scale and might to impose big data standards on public spaces. Collaboration between governments, municipalities, not-for-profit organizations, academics and private sector firms is essential to design and implement a standardized approach to smart cities, including how big data captured in public spaces is managed.

In addition, it has been well documented that access to data is a major issue for smaller innovative artificial intelligence (AI) companies (Vincent 2016). Although there is more data generated within firms and AI solutions are being designed for specific applications (Bean 2018), third-party data sharing is not occurring on a systematic basis. Currently, the vast majority of data owners, custodians and controllers, including government departments and agencies, are not willing or able to share or sell data to third parties. Private sector firms generally see proprietary data as giving them a market advantage over competing firms and have few incentives to share with third parties. Most firms are looking at monetizing data through analytics they can sell to an established customer base.

The good news for innovative Canadian AI firms looking for data is that once standards are introduced, most players in the newly created ecosystems end up winning. New standards quickly bring about technological and product certainty, which lowers the risk for data custodians upstream and AI firms downstream. Standards also allow smaller players to enter the market and introduce price competition.

Citizen Engagement

As indicated above, public spaces and the data they generate are assumed to be public goods for the purposes of the proposed standards. As such, the standards would codify an approach and mechanisms to engage citizens on acceptable levels of monitoring and surveillance in public spaces as well as on the modalities of data collection, sharing, access, retention and secondary uses. Although new citizen engagement models can be designed from scratch, the standards could refer to the “Model Surveillance & Community Safety Ordinance” published by the American Civil Liberties Union (ACLU). It codifies a model ordinance text that can be adapted into a voluntary standards format. The model proposes the development of a surveillance use policy by relevant authorities, consultation with citizens on its modalities and formal adoption of the policy once approved through an ordinance (or a by-law). It provides detailed guidance on the features of surveillance use policies, compliance and oversight. Variants of the model surveillance ordinance are being adopted and implemented in cities across the United States (Conley et al. 2016).

However, the standards would have to go further than the model surveillance ordinance in order to establish a comprehensive framework for the creation and management of big data from public spaces. Citizens can be engaged to determine what are the most pressing issues that need to be addressed through smart cities. They have a stake in how data should be categorized, anonymized, accessed and used. They may also be interested in the benefits associated with data analytics from the commercialization of new algorithms to new academic research insights. The standards would provide guidance to engage citizens on these aspects.

Recent pilot projects aimed at collecting information through monitoring and surveillance tools show that engagement and social acceptability are key for success. A recent example includes an Ivanhoé Cambridge initiative at a shopping centre in Quebec where video cameras were installed to collect demographic data about clients. Storefront signage was installed and a comprehensive communications campaign was conducted about the objectives of the pilot project and the features of data capture (such as data anonymization and short data retention timelines).
Nevertheless, the Quebec Information Access Commission received more than 10 complaints from concerned consumers and launched an investigation on the technology (Bussière 2019).

Data Collection and Organization

Municipalities aiming to become smart cities face significant challenges related to data collection and organization across among a multiplicity of departments and agencies. The standard PAS 181:2014 Smart city framework — Guide to establishing strategies for smart cities and communities from the British Standards Institution (BSI) provides guidance on approaches to develop and implement a smart cities framework and break down individual silos (BSI 2014).

Detailed guidance regarding data collection and organization for smart cities already exists. Ideally, the proposed standards would align with the smart city concept model (SCCM) outlined in ISO/IEC 30182 — Guidance for establishing a model for data interoperability.¹ The standard, published in 2017, provides guidance on a model that can provide the basis of interoperability between component systems of a smart city, by aligning ontologies in use across different sectors. It proposes concepts (such as organization, place, community, item, metric, service and resource) and outlines an approach to map out relationships and associations between concepts (for example, an organization has a resource, an event occurs at a place) in order to create the necessary connections between needs and available data. The SCCM can be used to catalogue data holdings from different organizations; promote identifiers and categorizations as reference information; agree on data standards for specialisms; understand data sets from other sectors; and construct a local data ecosystem where data can be contributed and consumed by different organizations and people in a city.²

The identification and labelling of monitoring and surveillance data sources in public spaces, including standardized signage and QR codes, may improve social acceptability. Both identification and labelling would certainly benefit from standardization. Different schemes have been developed over the years to facilitate the identification of security cameras in public spaces. For example, long-established programs such as Surveillance Watch propose specific signage and information guidelines for surveillance cameras located in public and commercial spaces.³

Recently, the Toronto Sidewalk Labs project unveiled a signage prototype for monitoring and surveillance devices installed in public spaces. Through a project entitled Designing for Digital Transparency in the Public Realm, Sidewalk Labs is looking for partners that want to advance the use and adoption of these signage standards in the public realm. This initiative showcases the wide range of monitoring and surveillance devices currently in place in urban infrastructures and the need for standardizing how citizens can relate to them.⁴ For example, a Sidewalk Labs prototype for data sources in buildings includes the following types of monitoring and surveillance equipment:

- infrared sensors (to detect motion and measure occupancy);
- security cameras (to monitor security);
- Numina sensors (to assess mobility patterns);
- temperature sensors (to monitor indoor temperatures);
- infrared depth sensors (to monitor the occupancy of a space);
- smoke detectors (to monitor signs of smoke);

---

¹ ISO is the International Organization for Standardization and IEC is the International Electrotechnical Commission. They are standards development organizations that develop and maintain international standards.


³ See https://surveillancerights.ca/.

⁴ See www.sidewalklabs.com/dtpr/.
thermostats (to control temperature);
→ light switches (motion sensors to turn lights on and off);
→ faucet switches (motion sensors to turn faucets on and off); and
→ door locks (keyless systems to provide access).

The signage provides information about the monitoring equipment in use and whether it collects identifiable information. QR codes are clearly visible to allow individuals to acquire additional information on the devices and their intended uses.

Similar inventories of monitoring and surveillance equipment and signage conventions would be needed for other types of public spaces, from roads to public transit to facilities providing municipal services.

Guidance is also needed regarding the categorization of data sources in order to assess data quality. Proper data characterization and categorization are essential for effective data analytics. On that front, the ISO published in 2015 the ISO 8000 standard Information and Data Quality: Concepts and Measuring. Additional standardization work will be required over time in order to establish a robust data taxonomy for smart cities based on the ISO 8000 framework.5

Finally, guidance on data security, from its origin through IoT devices to storage, needs to be addressed (Pandey et al. 2019). Fortunately, normative documents and best practices covering these issues already exist or are being developed, including the recently published UL 2900-2-3 cyber security standard. The standard has been recognized by the US Food and Drug Administration for medical devices and can be used for the certification of IoT devices (Bean 2018; UL 2017a; UL 2017b).

Data Access, Sharing and Retention

Data access, sharing and retention issues likely require the most effort from a standardization perspective. Although software, hardware platforms and protocols already exist for data access, sharing and pooling, choices have to be made regarding the principles, criteria and required specifications for big data captured in public spaces to be made available. At the core will be the challenge to choose suitable IoT platforms to allow for data sharing between various IoT streams. The choice of platforms is also growing. In 2018, there were more than 450 different IoT platforms available in the global marketplace, but the number could soon reach close to 1,000 different available platforms (McClelland 2018). A recent article from the IoT For All blog states that data sharing is still a major challenge and that it is difficult for developers and start-ups to innovate in the smart cities space without access to rich data sets. Open IoT specifications and architecture could be a solution if they prove to be conducive to interoperability (Margossian 2018).

Guidance on appropriate mechanisms for the anonymization of data sources is another topic that is key for smart cities to succeed. Significant work has been done over the past 25 years on techniques for data anonymization, which could be used as input in the proposed standards.6 On this front, Canada stands to win if voluntary standardization can provide the right framework for the use of anonymized data. In the United States, for instance, privacy case law has been described as an impediment to the use of anonymized data sets for research (Rubinstein and Hartzog 2015).

There are a number of possible approaches and processes to make data generated from public spaces available to private sector firms and academics. Ideally, the standards would provide guidance on essential requirements for data sharing and access. One promising avenue, called a public data sharing framework, was designed in Australia through a collaborative effort between governments and the digital industry sector and coordinated through the Australia Computer

---


6 See www.caida.org/projects/predict/anonymization/.
Safeguarding Big Data Captured in Public Spaces through Standardization

Society (ACS). The framework, which aims at safely sharing data relating to individuals, storing it securely and ensuring it is only accessed and used by approved users with pre-determined credentials, is based on an updated version of the Five Safes framework. It may provide much needed guidance for future standards focused on managing data access, sharing and retention (Oppermann 2018).

Data trusts are another approach to pool and manage access rights to data generated from smart cities initiatives. Like legal trusts, data trusts appoint a steward (trustee) to manage an asset for a purpose on behalf of a beneficiary or beneficiaries who own the asset. According to Sean McDonald (2019), the concept has gained in popularity after being proposed by the United Kingdom as a vehicle to pool data in support of its AI industry growth strategy. Recently, Sidewalk Labs proposed using a civic data trust to govern the data collected as part of its flagship smart city project in the Quayside area of Toronto. New forms of data access governance may need additional legislative support but offer interesting models to consider.

Next Steps

The development of voluntary standards to safeguard big data captured in public spaces will not be a panacea but is required for smart cities to succeed. Around the world, a number of IoT labs and open cities labs have been launched. AI start-ups and academics collaborate with citizens and municipal officials to generate new insights by accessing big data generated from public spaces. In Canada, the newly created Open City Network is intended to address domestic cloud, data, residency, identity, authentication, access rights for various categories of users and application programming interface issues. It aims at creating an operational and technical platform, which will provide the core infrastructure for innovators to quickly build, deploy and scale open city innovation which will require standards.

The CIO Strategy Council, based in Canada, has recently discussed the merits of developing voluntary standards to support big data initiatives in smart cities. Other standardization bodies may be interested in coordinating the required standards development work. By taking a careful, step-by-step approach, interested parties can contribute to the laudable goals pursued by smart cities while keeping our public spaces whole and safeguarding our fundamental rights.

---

7 For more information on the framework, see Australian Computer Society (2017).

8 Noteworthy IoT smart cities labs include Busan, South Korea (www.startiot.or.kr/main/view); Helsinki, Finland (https://fiksukalasatama.fi/en/smart-city/internet-of-things-iot-trials-in-smart-kalasatama/); Amsterdam, Netherlands (http://iotlivinglab.com/) and other living labs that are part of the European Community of Living Labs (https://enoll.org/).

9 See https://theopencity.org/.
Works Cited


About the Global Economy Program

Addressing limitations in the ways nations tackle shared economic challenges, the Global Economy Program at CIGI strives to inform and guide policy debates through world-leading research and sustained stakeholder engagement.

With experts from academia, national agencies, international institutions and the private sector, the Global Economy Program supports research in the following areas: management of severe sovereign debt crises; central banking and international financial regulation; China’s role in the global economy; governance and policies of the Bretton Woods institutions; the Group of Twenty; global, plurilateral and regional trade agreements; and financing sustainable development. Each year, the Global Economy Program hosts, co-hosts and participates in many events worldwide, working with trusted international partners, which allows the program to disseminate policy recommendations to an international audience of policy makers.

Through its research, collaboration and publications, the Global Economy Program informs decision makers, fosters dialogue and debate on policy-relevant ideas and strengthens multilateral responses to the most pressing international governance issues.

About CIGI

We are the Centre for International Governance Innovation: an independent, non-partisan think tank with an objective and uniquely global perspective. Our research, opinions and public voice make a difference in today’s world by bringing clarity and innovative thinking to global policy making. By working across disciplines and in partnership with the best peers and experts, we are the benchmark for influential research and trusted analysis.

Our research programs focus on governance of the global economy, global security and politics, and international law in collaboration with a range of strategic partners and have received support from the Government of Canada, the Government of Ontario, as well as founder Jim Balsillie.

À propos du CIGI

Au Centre pour l’innovation dans la gouvernance internationale (CIGI), nous formons un groupe de réflexion indépendant et non partisan doté d’un point de vue objectif et unique de portée mondiale. Nos recherches, nos avis et nos interventions publiques ont des effets réels sur le monde d’aujourd’hui car ils apportent de la clarté et une réflexion novatrice pour l’élaboration des politiques à l’échelle internationale. En raison des travaux accomplis en collaboration et en partenariat avec des pairs et des spécialistes interdisciplinaires des plus compétents, nous sommes devenus une référence grâce à l’influence de nos recherches et à la fiabilité de nos analyses.

Nos programmes de recherche ont trait à la gouvernance dans les domaines suivants: l’économie mondiale, la sécurité et les politiques internationales, et le droit international. Nous comptons sur la collaboration de nombreux partenaires stratégiques et avons reçu le soutien des gouvernements du Canada et de l’Ontario ainsi que du fondateur du CIGI, Jim Balsillie.