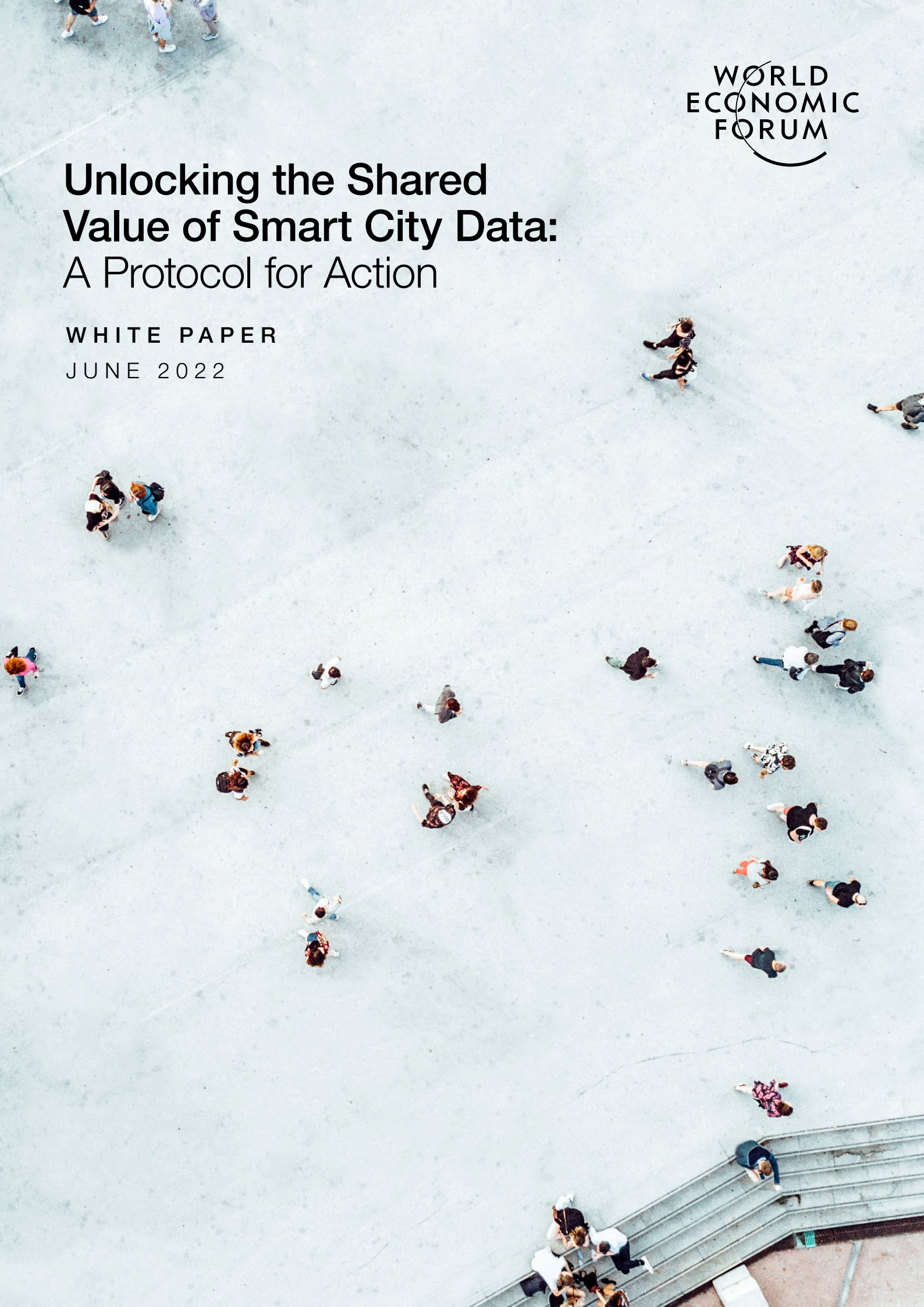
WORLD ECONOMIC FORUM

Unlocking the Shared Value of Smart City Data: A Protocol for Action

WHITE PAPER

JUNE 2022



Images: Getty Images, Unsplash

Contents

Foreword

Smart city technologies hold the potential to transform urban life.

Jeff Merritt
Head of Urban Transformation,
Member of Executive Committee,
World Economic Forum

The urban landscape of the world is rapidly changing. Currently, 55% of the world's population live in cities and this figure is projected to grow to 68% by 2050.[1] This increase in the urban population is giving rise to a plethora of economic and societal opportunities, as well as challenges.

The advent of Fourth Industrial Revolution technologies, such as the internet of things (IoT), machine learning and artificial intelligence, provides an opportunity to maximize the benefits of rapid urbanization while at the same time reducing its costs. By leveraging these technologies, cities have the potential to improve and ease the lives of residents, businesses and visitors across a wide range of issues from waste management, traffic and city operations to public health and safety.

Data is at the heart of this vision for more sustainable, resilient and equitable cities as it fuels the Fourth Industrial Revolution. More data is being generated than ever before, with the global volume of data created in the next five years predicted to be more than double that of the advent of digital storage.[2] Although an unprecedented amount of data flows across borders and devices, the regulatory environment remains fractured, with more than 130 data privacy laws in place globally.[3] Much of the data needed to tackle the world's most pressing challenges lies siloed in public and private sources, with an array of regulatory, commercial and social risks preventing the sharing of data, even for social good. As a result, only a very small amount (less than 1%) of the data is used to foster innovation, create value and drive decisions to address macro-level challenges.

It is in this context that the World Economic Forum's Centre for the Fourth Industrial Revolution and the Government of Denmark came together in partnership. Beginning in autumn 2018, a series of workshops, interviews and discussions were held both virtually and in cities across the globe, from San Francisco to Delhi and Shanghai. Through these discussions and the creation of an international working group consisting of more than a dozen World Economic Forum Fellows and Partners from Canada, China, Denmark, India, Japan and the United States, the protocol presented in this publication was born. This white paper documents the insights from this multi-year collaboration and a framework to help realize the potential of more data-driven cities.

We hope that this publication can help facilitate and encourage continued multistakeholder global collaboration on this topic, supporting the development of new innovative approaches and governance models for unlocking the shared value of smart city data.

Executive summary

A strong governance framework is critical to realizing the potential of more data-driven cities.

Every day more than 2.5 quintillion bytes of data are generated, with over 50% coming from internet of things (IoT) devices. But less than 1% of this data is fully utilized. Harnessing the full potential of IoT data has been hampered by a fractured regulatory data environment, a lack of data policies and regulations, privacy concerns, public and private data sources that remain siloed and the absence of business models based on the value of data.

Nowhere are these challenges and opportunities more visible than in cities. Smart cities hold the promise of increased efficiency, improved public services, elevated quality of life and economic growth but are predicated on access to high-quality and dependable data.

To address the challenges, realize the opportunities and manage the risks arising from the growing amounts of IoT data, a sound legal structure is required with sustainable, secure and ethical models that facilitate the dynamic and efficient sharing of IoT data, and that foster innovation while protecting privacy and building trust.

This white paper analyses the current landscape of data collection and proposes a governance framework to align the IoT data sharing value chain in three key platform dimensions, as well as six major regulatory structures rooted in global standards: data privacy, data security, interoperability, accountability and integrity of data, eligibility of platform operators, and a terms of use agreement for data user, provider and operator. To illustrate this, the paper provides examples of trusted data sharing platforms, how they are structured and the essential elements to building trust to enable the opening of public sector data, and the trading of private sector data safely. The protocol also addresses the direction and actions that can be taken to implement the protocol in local policy and regulations.

To unlock the power of IoT, data requires governance protocols and regulatory frameworks that keep pace with the times, manage the growing volumes and complexity of data, and consider the supporting and resisting factors of all key stakeholders and their different governance structures. The novel framework proposed in this publication highlights key areas for proper data management and to enable trust between data providers, platform operators and end users, protect the rights and interests of all parties, and facilitate the dynamic exchange of IoT data. Innovative regulatory and governance models are the basis for cities to manage and benefit from the growing amounts of IoT information to help guide decision making, improve the quality of public services, and become more efficient, sustainable, resilient and equitable.

1 Introduction

Leveraging Fourth Industrial Revolution technologies to improve the lives of urban populations requires effective governance to enable access, understanding, feedback and learning.

Currently, 56% of the world's population (4.4 billion) reside in urban areas. This number is predicted to increase to 68% in the next 30 years.[4] The continuous advancement of urbanization has been coupled with the increasing concentration of various resources in cities and their surrounding areas, transforming cities into hubs for economic and societal advancement.

The advent and adoption of Fourth Industrial Revolution technologies, such as IoT, artificial intelligence and big data, are pushing the frontiers of urban services and transforming the way urban residents live and work. With the increasing integration of cutting-edge technology, cities' infrastructure is becoming increasingly complex in terms of design and interconnectedness, thus creating new opportunities and challenges.

Leveraging Fourth Industrial Revolution technologies to improve the lives of urban populations requires effective governance to enable access, understanding, feedback and learning. In this context, the so-called "smart city" technologies – the use of information and communication technology (ICT) to improve citizens' welfare – need to collect, analyse and understand data from various entities, sectors and levels to make decisions. The use of IoT technology-related data can enable the conversion of urban physical information into digital signals and facilitate the interconnection between the physical and digital worlds. This creates new opportunities for cities to improve the quality of government services and achieve efficiency in using resources and many other use cases and applications.

Many challenges that come with urban growth, such as limited land, resource inefficiency and traffic congestion, could be addressed by developing smart cities and mature technologies. Therefore, smart cities' technologies have the potential to significantly contribute to the attainment of the United Nations Sustainable Development goals and "make cities and human settlements inclusive, safe, resilient and sustainable".[5]

BOX 1 | **Leveraging data and digital services to address pandemic-related vulnerabilities**

The spread of the COVID-19 pandemic led to the emergence of many global vulnerabilities in 2020. As social distancing was enforced to curb the spread of the virus, ICT became imperative for continuing to provide services for the population and relay critical information.

In Saudi Arabia, the volume of digital services provided by the government increased by 70% in 2020 compared to the previous year.[6] These services include remote training for thousands of healthcare workers, millions of online processes managed by the Ministry of Interior and applications to help spread awareness of and assist with social distancing. The Saudi Data and Artificial Intelligence Authority developed the Tawakkalna app[7] to manage all COVID-19 related issues. Users can use the app to access more than 140 services, such as booking an appointment for the COVID-19 vaccine or accessing their Health Passport, which confirms their COVID-19 vaccine dose number. The application currently has 27 million users.

In addition, Saudi Arabia's Communications and Information Technology Commission has also accelerated the approval of dozens of new delivery applications that help serve the population while they are socially distancing.[8] The requirements and guidelines developed for these applications has also helped advance the growth of the technology-based services ecosystem as a whole, providing a launch pad for future growth. With the reduction of regulatory barriers, Saudi Arabia saw an increase of 500% in delivery agents and 240% in executed orders.[9]

With an increasing amount of data being captured daily, there is a growing challenge around data silos and privacy. Every day, more than 2.5 quintillion bytes of data are generated, more than 50% of which come from IoT devices.[10] However, there has not been a widely adopted business model driven by data value. Meanwhile, cities have tended to pursue smart city initiatives in silos.[11] In addition, datasets are stored in different systems, controlled by different players and disconnected from each other without unified formats or processes.

Due to factors such as insufficient mature policies and legal frameworks, the low level of public-private data sharing, or the absence of universal standards, corporate expertise and societal culture, data remains unable to live up to its full potential. According to statistics, less than 1% of IoT data has been fully utilized.[12]

Unlocking the value of data requires the building of trust. Today, both organizations and individuals have become increasingly cautious about their data rights and privacy. A specific survey of data trust in the United Kingdom, for example, shows that the "vast majority of consumers (78%) believe that businesses benefit disproportionately from data exchange".[13]

Data is as essential to the Fourth Industrial Revolution as steam was to the first and electricity was to the second. In developing future smart cities, data will be an important means of production that provides the necessary nutrients and is a critical instrument for many relevant parties to analyse and solve economic and social problems.

When data is strictly maintained by several separate entities, it is static and prevented from flowing freely as isolated islands. As will be discussed throughout this publication, to unlock productivity and stimulate the vitality of the cities of the future, it is increasingly crucial to build trusted platforms for dynamic data exchange that foster innovation while protecting privacy.

These governance protocols and policy frameworks must be robust and modern, able to manage the growing volumes and types of data being generated and shared, and take into account all key stakeholders and their different governance structures. This will create a foundation for a sustainable and mutually beneficial exchange over time.

"To unlock productivity and stimulate the vitality of the cities of the future, it is increasingly crucial to build trusted platforms for dynamic data exchange that foster innovation while protecting privacy.

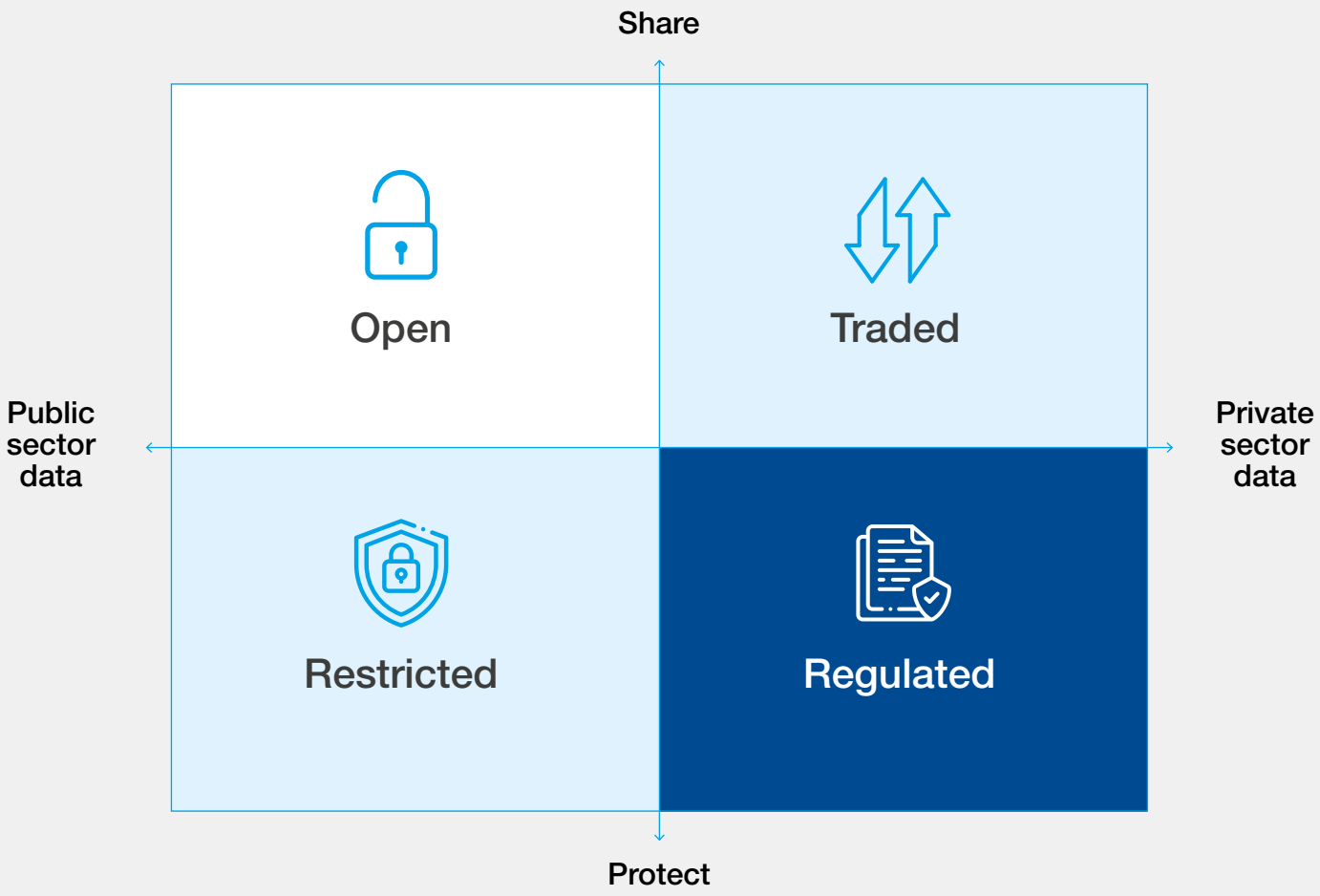
2 Data sharing in the IoT age

Data protection is vital to both public and private sectors, particularly when opening or trading data, as data sharing presents risks.

The general framework of data sharing that this protocol proposes takes into account two criteria – the source of the data and the corresponding permissible action that can be carried out with the data. Two dimensions are identified in a matrix: "public sector data" vs. "private sector data" and "share" vs. "protect". These two sets of forces divide the matrix into four quadrants: open, restricted, traded and regulated (see Figure 1).

Public sector data refers to the data generated, collected and stored by international, national, regional and local governments, and other public institutions and data created by external agencies for the government or related to government programmes and services.[14] On the other hand, private sector data refers to data generated, collected and owned by private companies or individuals, such as customer activity data, personal data, business operational data and industrial data.[15] This data is tradable among different companies and institutions.

FIGURE 1 **The framework of data sharing**



2.1 Public sector data

As a key public asset for digitization and innovation, public sector data should be shared with the public for free through open platforms. Government open data platforms have been established in many countries to provide transparency, accountability and value creation.[16] According to the 2018 United Nations E-Government Survey,[17] the total number of countries with open government data platforms reached 139 in 2018, accounting for 72% of UN member countries. Meanwhile, many other countries have also established policy frameworks related to open data. Some international organizations such as the World Bank provide free and open access to data for all users on topics such as finance, business, health, the economy and human development from countries worldwide.

2.2 Private sector data

As a privately-owned asset, private sector data is usually shared through trading among companies or individuals. Data trading is considered an economic activity that generates profits through the transfer of data ownership, data usage rights and data income rights. Data trading favours the convergence and use of various types of data and has the potential to facilitate business activity, industrial innovation, social creativity and technological progress. In some countries, platforms have been created to promote data trading. For example, the Quandl platform, created in 2013 as a professional data trading platform in Canada, offers financial, economic and alternative data worldwide, and is used by more than 400,000 people, including analysts from the world's leading hedge funds, asset managers and investment banks.[18]

Whether it is the opening of public sector data or the trading of private sector data, data sharing requires the consideration of data protection. Opening public sector data without restrictions can bring risks and challenges in safety and privacy, as well as in civil liberty. In addition, to safeguard public interests, opening data concerning national security, trade secrets, or personal privacy should be restricted. Similarly, laws and standards should regulate data trading activities that affect personal privacy, third-party intellectual property, or trade secrets. Therefore, finding the right balance between data sharing and protection is crucial to building trust between providers and users.

2.3 The unique value and challenges presented by IoT data

To understand the value and challenges associated with data generated by IoT solutions, it is crucial to identify and assess its characteristics, particularly in the smart city context.

- **Objectivity:** IoT is an important channel to capture objective data from the physical world for use and integration into digital platforms and applications. As sensors and meters capture the state of or changes in material objects, human processing of this data and decisions based on IoT data can, in turn, affect the physical world, forming a circular system ("perception-decipher-feedback-perception").
- **Real time:** Data collected by sensors is generated in real time. In some of the applications, instant judgement and feedback are very important. For example, an autonomous vehicle must perceive road obstacles in real time and avoid them.
- **Diversity:** Although the types of IoT sensors are limited, there are several in common use, such as pressure, temperature and humidity sensors. In addition, these sensors could be integrated into different application scenarios. The combinations of diversified IoT data and the know-how behind them are infinite.
- **Massive:** The frequency of data generation from IoT nodes is much higher than that from the internet. For example, most sensor nodes are in a full-time working state and the data flow is endless. By 2025, it is estimated that 41.6 billion IoT devices will be online, generating 79.4 zettabytes of data.[19]

Meanwhile, the challenges of capturing and leveraging the unique value of IoT data are also enormous:

- **Complexity:** Data transmission applies to various communication networks from a low-power, wide-area network (LPWAN), Bluetooth, Wi-Fi and GPRS to 3G, 4G and 5G. A large proportion of this data ends up in the cloud or local servers. While IoT data sources are vibrant and sensor data formats are different, there are considerable barriers to interpretation and interaction.
- **Legal principles and privacy:** The origins of IoT data can be people, machines, equipment, public facilities and the natural environment. Therefore, the ownership of IoT data and the right to use it can be ambiguous. This poses another fundamental obstacle to the sharing and application of IoT data.
- **Security:** Currently, there is no unified and verified security protocol to ensure data security from generation, transmission and storage to application.
- **Cost:** The price of sensors remains a barrier that hinders the wide deployment of IoT. Due to the massive amount of IoT data, the cost of IT infrastructure, operation and maintenance is also huge, which offsets the benefits brought by the value of this data.

2.4 A new frontier: IoT data sharing

With the rapid growth in the number of connected devices, IoT is crucial in enabling the digital transformation of various industries and applications. The Boston Consulting Group (BCG) examined 75 smart city applications that use data from a variety of sources, including IoT devices, and found that nearly 50% of these applications require data sourced from multiple industries or platforms and among a broader set of potential future applications, 40% will similarly require cross-industry data aggregation. For example, the LinkNYC programme in New York – which focused on replacing payphones with Wi-Fi-enabled kiosks[20] – required three different companies to provide the necessary data, hardware and network capabilities. Given the potential value of applications that require aggregation from multiple sources, many cities are looking to move away from IoT-enabled applications that rely on limited, siloed data and are seeking partnerships with technology providers to develop platforms and other initiatives that integrate data from multiple sources.[21]

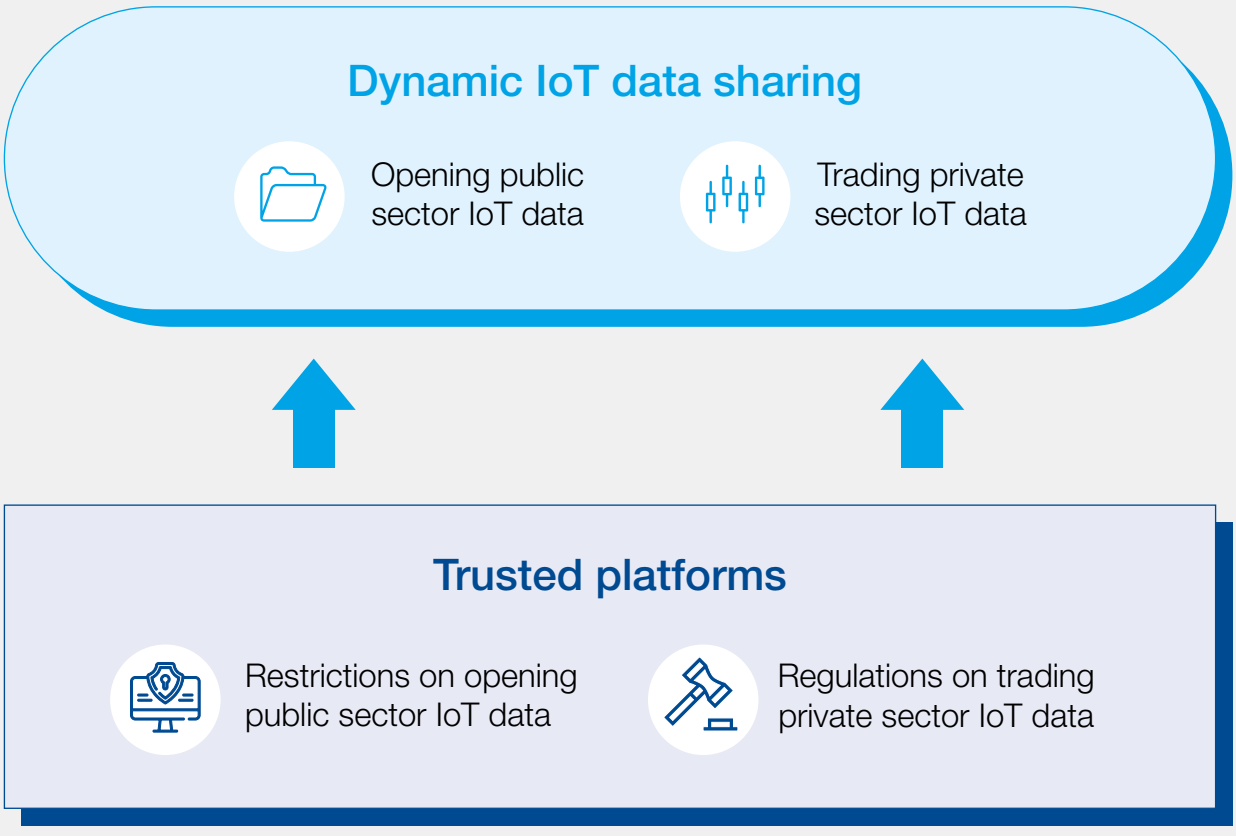
3 Defining the governance landscape for IoT data sharing

Generating, collecting, sharing and utilizing IoT data is key to creating value, as well as robust regulatory structures.

3.1 Trusted platforms

Given the characteristics of IoT data, it is necessary to develop a framework that responds to the specific context for sharing this data. Data-sharing platforms that are trusted by all stakeholders are important in enabling a dynamic exchange of IoT data. Necessary restrictions on the openness of IoT public sector data and regulations on the trade of IoT private sector data are essential to building this trust and making the public sector willing to open IoT data and making the private sector willing to trade IoT data (see Figure 2).

FIGURE 2 **Trusted platforms**

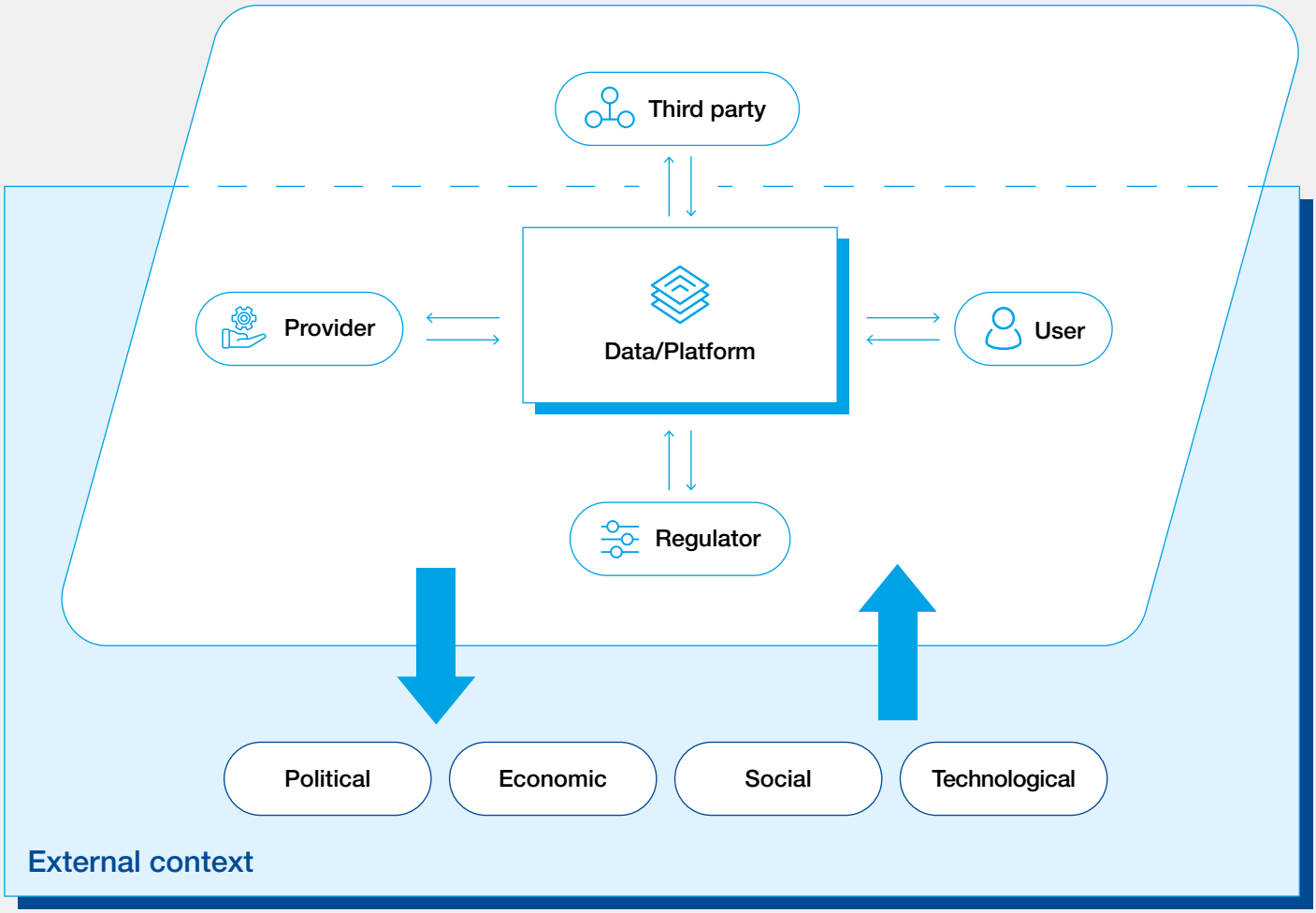


3.2 Key stakeholders and interactions in an IoT data-sharing platform

IoT data is generated, aggregated, distributed, traded and ultimately utilized to produce value. In its life cycle, the following primary stakeholders are involved (see Figure 3):

- Data providers collect and provide IoT data. They acquire IoT data from their own sensors, through customers on specific projects, or through cooperation with governments.
- Data users use their technological and analytical capabilities to apply IoT data in business, operations and research. They pay attention to the quality, velocity and variety of IoT data, and demand that data should be real-time, complete, authentic and accurate. In the value chain, first-hand IoT data users provide processed data to other second-hand users.
- Platforms are used as carriers to open or trade IoT data. Platforms offer places that are recognized and trusted by both data users and data providers.
- Regulators play a key role in fostering, standardizing and overseeing the collection, sharing and circulation of IoT data. Regulators formulate laws, regulations and policies related to IoT data flow, security and privacy protection.
- Third-party agencies include industry associations and research institutions that set standards, promote interoperability, enforce self-discipline and conduct research and evaluation to facilitate IoT data sharing.
- The external context includes the political, economic, social and technological environment that will influence the relationships among key stakeholders. Among these, the advancement of technology will continuously influence the interaction among key stakeholders and the long-term effects.

FIGURE 3 **Main stakeholders**



3.3 Dimensions and characteristics of trusted platforms with dynamic data sharing

A trusted platform with dynamic data exchange can be characterized by a combination of features from the following three classifications:

1. Variance in data types

The types of data managed by the platform are numerous and include raw data, calibrated data and processed data of different grades and sources, as well as insights generated from data processing. The sources also vary widely from public data collected by government agents and proprietary data collected by device owners, network operators, service providers and other business enterprises to data contributed by individuals and organizations. For example, the Basic Data Programme, developed under the Danish government's eGovernment strategy 2011-2015, covers personal, business, real property, address, geographic and other types of data collected by the government which can be accessed for free.[22]

WAGRI, an agricultural data collaboration platform, combines various kinds of public data, from soil and weather information to market conditions, along with commercial data from private companies, so that WAGRI participants involved in agricultural activities can have complete information throughout the agricultural production, circulation and consumption process.[23]

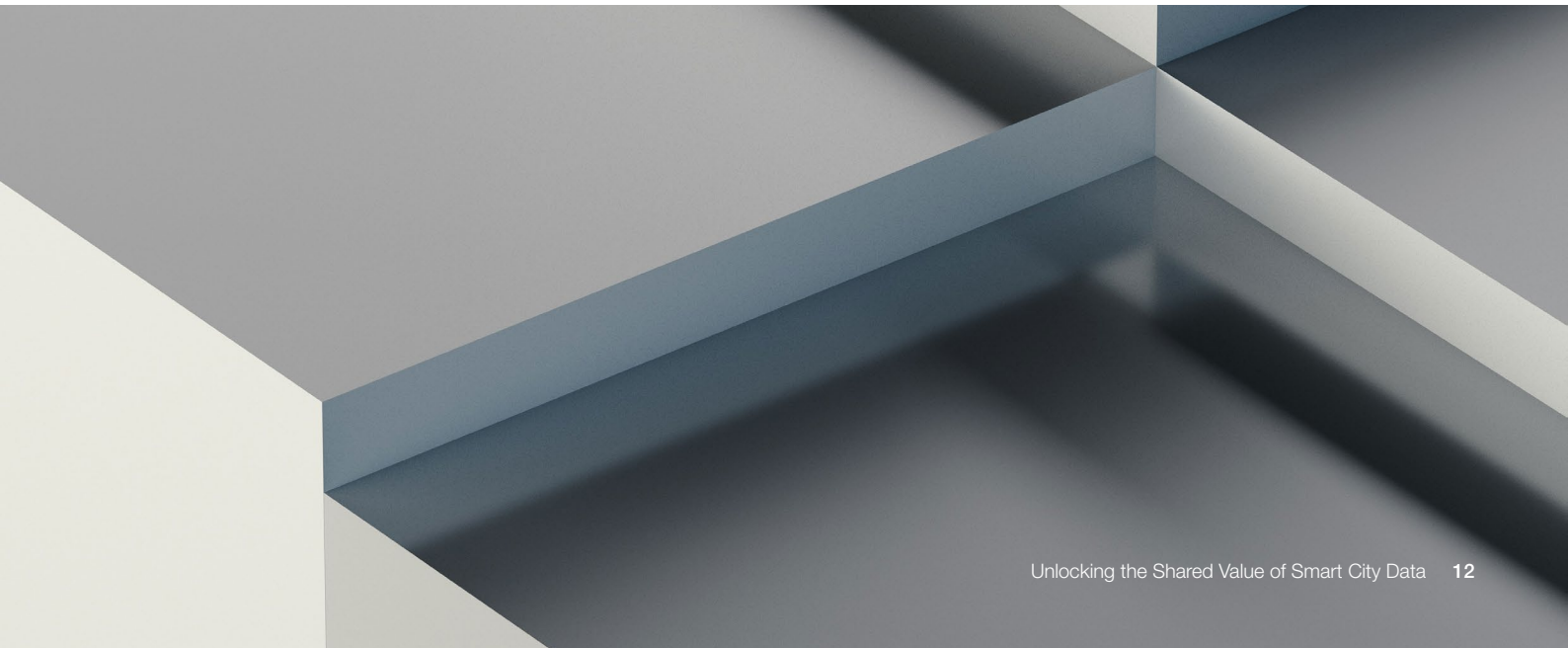
In Colombia, the Centre for the Fourth Industrial Revolution is leading the Moonshot project,[24] which deals with data as a strategic asset to generate social and economic value in society. Through this effort, the centre plays a key role in supporting the development and implementation of Colombia's National Data Infrastructure Plan as one of the digital enablers to help combat the effects of the pandemic and reactivate and strengthen the digital economy. The Moonshot project has been developed in partnership with the World Economic Forum's Data for Common Purpose Initiative.[25]

Since sensors typically generate IoT data in real time, a dynamic data platform must address the complexity and large scale of real-time data transfer. Depending on the platform design, the available data could be real-time only, while other platforms may provide a combination of historical and real-time data. The retention period of historical data depends on the individual design of each platform to address the needs arising from data analysis, legal record-keeping requirements and other business or administrative operations.

With IoT data coming from various sensors and sources, data formats can vary. The multiplicity of data formats poses challenges for processing and synthesizing data from different sources. As a result, some operators standardize data formats and structures across their platforms. Reaching a consensus on standards and formats can facilitate efficient data sharing and use.

In the face of growing privacy concerns, personal data rights advocates proposed the creation of a "personal data warehouse" to let individuals control the flow of their data. Instead of sharing raw data, data analytics can be done inside a personal data store and provided to apps or services, limiting the possibility of revealing data to third parties with whom the person has no prior relationship.[26] One instance is the Hub of All Things project, a personal data platform that "confers intellectual property rights of personal data to individuals through their ownership of a dedicated database, wrapped with containerized microservices" and allows them to own a personal data server (called the HAT Microserver) and all the data within it.[27]

"Since sensors typically generate IoT data in real time, a dynamic data platform must address the complexity and large scale of real-time data transfer.



BOX 2 | **India's urban data exchange (IUDX) platform**

Making quality and usable datasets affordable and easily accessible can enable the start-up ecosystem and industries to deliver innovative solutions for smart cities. With this motivation in mind, the India Smart Cities Mission launched the India Urban Data Exchange (IUDX) platform.[28] This open-source software platform aims to support the secure and authenticated exchange of various types of data between platforms, third-party applications, data producers and data consumers in a simple and uniform manner. The IUDX platform allows data providers and data users to share, request and access datasets related to cities, urban governance and urban service delivery, all in one platform. Examples of use cases include smart water and energy management and predictive stormwater management.[29]

To protect privacy and security in data exchange, the platform can classify data into different categories for differentiated uses and management. Data classification can help ensure that any collection, transaction and use of data complies with regulations and rules regarding data privacy and security. The platform offers a five-level data use classification: "public" – for public consumption and use; "internal use" – accessible only by municipal corporation employees for managing operations and the delivery of public services; "sensitive" – regulated by any city/state/central law or regulation, e.g. on issues such as privacy; "protected" – for sensitive data, such as the identity of individuals, and in the event of any breach or loss of such data the municipal corporation must disclose/notify in a timely manner; and "restricted" – data that poses a risk or threat to life or loss of public assets or critical infrastructure. This provides full control for the data owners regarding data access.[30]

2. Platform operations

The platform may grant differentiated access to different groups of users depending on the nature of the data and the operation of the platform. Some open data platforms are accessible to all users, while others are accessible only to certain certified user groups within the scope of approved uses. Even in the case of government-run data platforms, not all public data collected by authorities is publicly available.

The degree of openness is often correlated with how the data is classified and data relating to privacy, public safety and other sensitive data is often restricted within authorized government agencies.

In addition to user classification, the business model of a data trading platform also determines its access and permission model. EverySense, the first IoT data exchange market launched in 2015 in Japan, started with a trial membership system, with an initial registration fee of 50,000 yen and a monthly payment of 50,000 yen.[31]

OpenDataSoft, a platform launched in 2011 initially as a project to build France's national open data portal, offers customized data-sharing solutions on a single platform that enables organizations to publish, manage, analyse, visualize and share data in various formats. Both public sector and private companies can use its services to design and operate their own open data platforms for private data sharing within companies, or as an IoT and smart city data solution.[32]

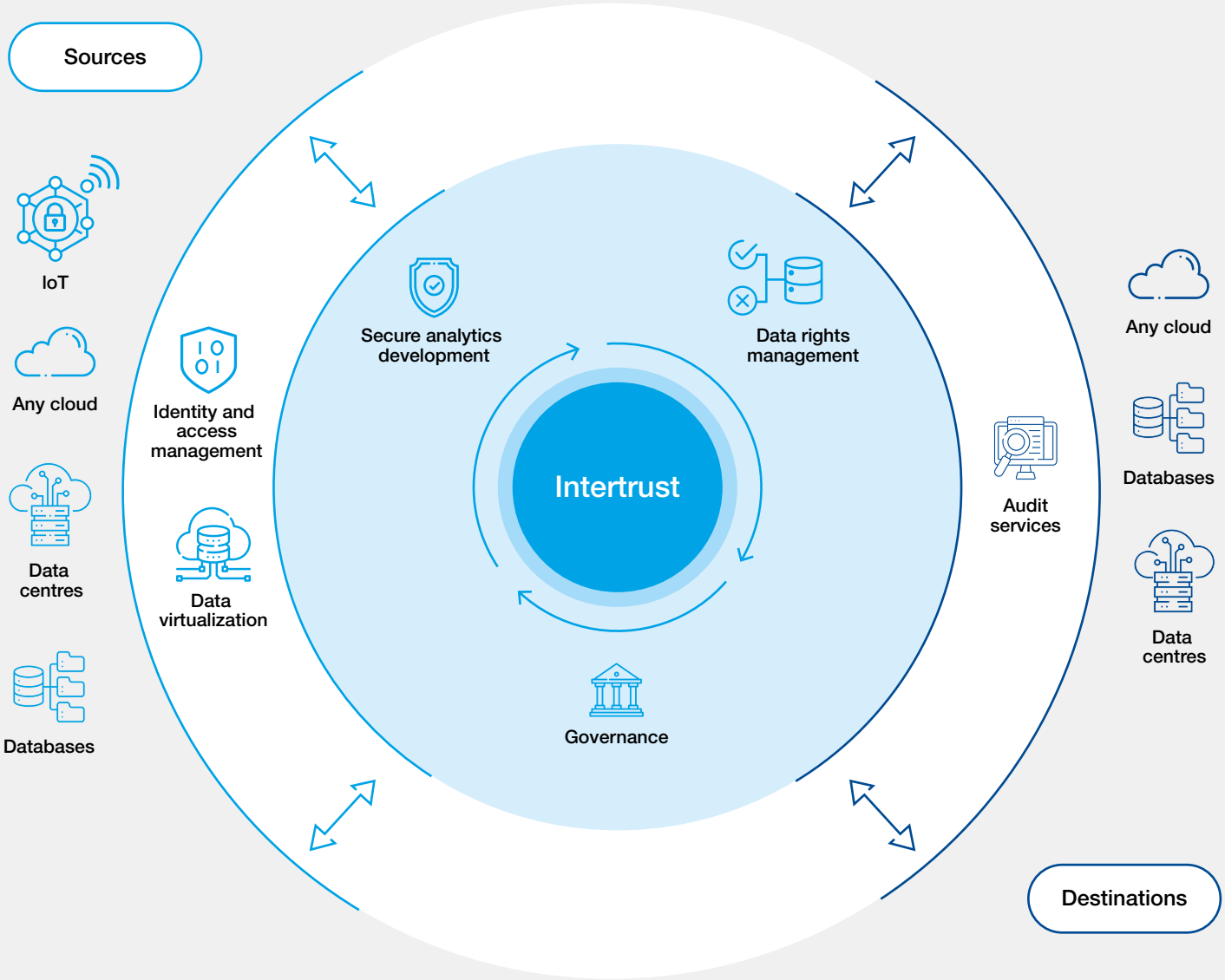
Data monetization is a typical data marketplace strategy. It should be noted that data monetization is not limited to data trading but also includes the type of activities that involve profits generated from data analytics and other information-based products and services. One big challenge is to unlock the power of proprietary data often concentrated in the hands of a few conglomerate operating system providers. An example initiative to tackle this challenge is the Decentralize Data for Dubai programme, which involves the use of crypto tokens to manage a digital infrastructure that is decentralized, private and secure.[33] With distributed ledgers and a decentralized data exchange framework, users, companies and government agencies can manage, buy, sell and rent data. This strategy has the potential to allow all parties to contribute and use the data, and reduce the competitive advantage of hoarding private data.

To improve its services and build trust between data providers and users, the platform may incorporate a feedback mechanism and allow users to rate datasets. A feedback mechanism can also help the platform to adapt its features and services over time to reach a broader set of users. New York City's Open Data platform offers interactive features for users to rate data, contact data providers, submit and read comments, and interact with government agencies related to specific datasets.[34]

In this context, there is an increasing need to find a secure, fast and cost-effective method for sharing large amounts of sensitive and valuable data. The Intertrust Modulus platform has developed a model that combines data virtualization, data governance and a secure data service that enables trusted data sharing between companies and partners. Instead of the traditional approach of moving data around to perform computations, the platform offers a secure system that essentially enables end-to-end data sharing, permissions and provenance between entities[35] (see Figure 4).

"There is an increasing need to find a secure, fast and cost-effective method for sharing large amounts of sensitive and valuable data.

FIGURE 4 | **The Intertrust Platform's approach to enabling interoperability across diverse datasets and devices**



The COVID-19 pandemic had a visible impact on all administrative sectors and generated the need to react quickly and create applications on a massive scale in a short period of time, which promoted the acceleration of the digitization of public and private services.

3. Platform governance

How a platform is governed depends on the roles of different actors and parties in setting and enforcing the rules, providing financial support and executing operations. Apart from a lack of coordination, these initiatives are often missing an overarching strategy, which indicates potential fragmentation in city agencies and the issue of legacy technologies. Thus, to fully realize the potential of smart cities technologies, city governments need to overcome organizational and technological siloes.[36]

To address this, national and local governments have created open data platforms and directories to allow greater access to and visibility of available data. In a review by BCG, approximately 40% of the smart cities analysed have open data platforms with integrated back-end databases with some 15% prioritizing the building of back-end connectivity.[37] Two examples of these initiatives are Portland Urban Data Lake[38] and Smart Seoul Data.[39]

To fully realize the potential of smart cities technologies, city governments need to overcome organizational and technological siloes.

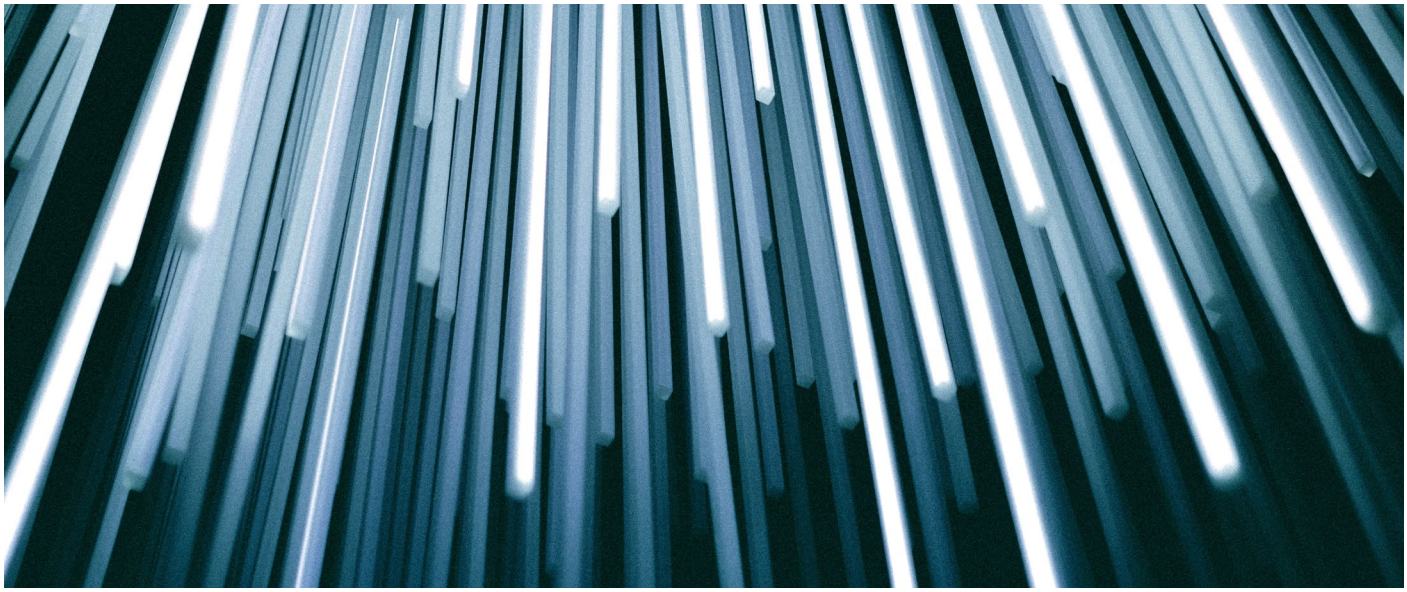
Open public data platforms typically operate with full government funding, while others are funded partially by the government or entirely by private funding. A platform's funding may also evolve through the different stages of development. For example, a data platform initially funded by the government may attract investment from companies once an ecosystem has been built to support business activities.

With growing concerns about digital sovereignty, an increasing number of governments are taking steps to invest in digital and data infrastructure to ensure self-determination and control stored and processed data. The GAIA-X project, initiated by the German government in 2019 in response to the dominance of non-European cloud providers and data privacy and security concerns stemming from this dominance, aims to join forces with other European countries to create a federated data infrastructure.[40] This data structure combines existing and newly created edge and cloud solutions, distributes and coordinates data processing, storage and analysis between local edge and central cloud services, and has the potential to address the challenges that come with the huge amounts of real-time IoT data.[41]

In terms of rulemaking and enforcement, governments or government-affiliated agents can directly regulate and enforce a platform, which is typical for public data platforms. It can also be operated by third-party enterprises using government-defined rules or run and owned by private entities which set and enforce their own rules of operations in compliance with laws and regulations. Some data exchange technology companies, such as Dawex, which specializes in data marketplaces for the monetization of datasets, application programming interfaces (APIs), raw data, refined data and insights, allow data suppliers to configure the terms and conditions (such as data format, history, transaction-based or subscription-based) for the monetization of data.[42] Another instance is the Shanghai Data Exchange,[43] a mixed-ownership enterprise founded in 2016, which provides on its platform a set of guiding principles and standards for data exchange. These guidelines cover aspects from personal data protection, legal rights and liabilities of data owners, operators and users, principles for data processing and a list of data types forbidden from circulation. It also prescribes a set of guiding standards for de-identification, data tag assignment, data updating frequency and statistics cycle, data retention period, data pricing methods and others.[44]

From the standpoint of facilitating the proper utilization of personal data, the government of Japan held a series of internal consultations on the concept of "information banks", a system intended to promote the sharing and utilization of data and to implement such schemes through public-private collaboration. In addition, an arbitrated certification system was developed to ensure that personnel meet certain minimum criteria for reliable information management.[45] Since 2018, the Japan Information Technology Federation has accepted applications for certification as an information bank based on the Guidelines for Certification Schemes Relating to Information Trust Functions version 1.0.[46] And in 2019, the information bank certification was granted to Sumitomo Mitsui Trust Bank and FeliCa Pocket Marketing, which accredits them as trusted organizations that can use personal data for business purposes while always safeguarding individual privacy.[47]

Some cities are also sharing data with private companies to foster innovation by hosting Startup in Residence programmes, which offer startups various resources, such as a physical space and access to the data needed to develop solutions for the public sector.[48] One example is the City of Boston's Beta Blocks programme, which aims to create a platform for civic experiments, bringing together the city government, technology companies and local communities to reimagine how technologies can help solve neighbourhood problems and assess the impact of these technologies on the city.[49]



3.4 Developing a regulatory protocol framework rooted in global standards

A robust regulatory structure that considers data privacy, security and interoperability is crucial for properly managing data and nurturing stakeholders' trust in platforms. This is important to protect the rights and interests of all parties and facilitate dynamic sharing and best uses of data for both public and private interests.

In this regard, this publication proposes a model framework composed of six components:

1. Data privacy

The ubiquity of IoT sensors deployed in public and private spaces raises privacy issues. To build trust in IoT devices and their key actors, policy and legal frameworks should be in place to ensure that data collection, sharing and use do not infringe upon privacy.

One fundamental step to protect privacy is to classify data according to its type and with varying degrees of sensitivity. The definition and scope of "personal data" (also known as "personal information" or "personally identifying information") can vary depending on the jurisdiction. The General Data Protection Regulation (GDPR) defines "personal data" as "any information relating to an identified or identifiable natural person; an identifiable natural person can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person". The California Consumer Privacy Act (CCPA) defines "personal information" as "information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household".

One key difference is that "personal data" under the GDPR covers publicly available data, mainly referring to information lawfully made available from government records, whereas "personal information" under the CCPA does not cover publicly available information. The GDPR also provides "special categories of data", referring to the sensitive data that reveals "racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership and the processing of genetic data, biometric data to uniquely identify a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation". The GDPR prohibits the processing of such sensitive data except under specific circumstances as prescribed by law. The CCPA does not separately define "sensitive data" or "special categories of personal data". It provides a definition of "biometric data" but does not mandate special rules for collecting or sharing biometric data.[50]

In China, the Information Security Technology – Personal Information Security Specification defines "personal information" as any information that is recorded, electronically or otherwise, that can be used alone or in combination with other information to identify a natural person or reflect the activity of a natural person.[51] "Personal sensitive information" is personal information that, once leaked, illegally provided or abused, could endanger personal and property safety or easily lead to the damaging of personal reputation and mental and physical health or discriminatory treatment.[52]

Best practices for data privacy may include:

- Consent for data collection, storage and use.
- Minimal data collection and processing for specified, explicit and legitimate purposes.
- Measures of security such as access control, privileged access management, credential choices and management, multifactor and granular authentication, encryption and web app security.
- Removal of personal information through measures such as de-identification.
- A clear declaration on the scope of authorized uses and life cycle of the shared data storage limitation, as well as whether and how the data will be shared with third parties.
- Awareness of the risks and related policies regarding collecting data from special groups of the population, such as children.

2. Data security

The influx of real-time IoT data and additional entry points into the network increases potential vulnerabilities and risks in any IT infrastructure. Risks include unauthorized access to data, infiltration of IoT devices and the use of compromised IoT devices to attack other devices on the network. There is a need to ensure a high level of data security through policies, standards, frameworks, certification and other data security practices.

"To build trust in IoT devices and their key actors, policy and legal frameworks should be in place to ensure that data collection, sharing and use do not infringe upon privacy.

The influx of real-time IoT data and additional entry points into the network increases potential vulnerabilities and risks in any IT infrastructure.

In 2018, the National Institute of Standards and Technology published the *Interagency Report on the Status of International Cybersecurity Standardization for the Internet of Things (IoT)*, which provides a comprehensive assessment of IoT cybersecurity risks, the landscape of standards for IoT cybersecurity, as well as other valuable resources for strengthening IoT security.[53]

IT services management frameworks such as the Information Technology Infrastructure Library, which provides guidance on the strategy, design, operation and improvement of IT services, can provide useful resources for maintaining IoT security. In general, some data security best practices recommend that all parties involved in data exchange, including the provider, the administrator of a data platform and the data user, keep the IoT infrastructure and/or devices updated and patched on schedule.

An internationally recognized standard for implementing, managing and maintaining information security is recommended to platform operators. The International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 27000 family of standards provides the requirements for an information security management system, including a dozen standards for a systematic approach to managing sensitive information. In the specific case of IoT devices, it may be particularly useful to consult the ISA/IEC 62443 (formerly ISA-99), aimed at their application in industrial automation and control systems, covering the early design and implementation phases, systems integration, as well as day-to-day use, management and maintenance.

Some of the best practices for data security may include:[54]

- Defining in-transit data protection requirements, such as encryption standards.
- Implementing in-transit encryption and key security.
- Using a certificate management service and defining a revocation process to revoke certificates if they are compromised.
- Finding a detection tool or mechanism to detect any attempt to move data outside the defined boundaries as soon as possible.
- Implementing secure protocols such as Transport Layer Security or Internet Protocol Security to reduce the increase in data manipulation or loss.

BOX 3 | **Data privacy and security certification**

Certification is a conformity assessment activity performed by neutral third parties to assess and attest compliance with specified requirements generally derived from technical standards or legislation. The 2017 report *Recommendations on European Data Protection Certification*, published by the European Union Cybersecurity Agency, identifies and analyses the challenges and opportunities of data protection certification mechanisms.[55]

The ISO does not perform certifications. However, it recommends that, when choosing a certification body, it is important to check if the body a) uses the relevant Committee on Conformity Assessment of ISO standard; and b) is accredited, although accreditation is not mandatory.[56]

The International Accreditation Forum manages accreditation bodies' adherence to international standards in areas such as management systems, products, services, personnel and other conformity assessment programmes. The International Laboratory Accreditation Cooperation is the international organization of accreditation bodies in laboratory and inspection accreditation.

Below is a selected list of data protection certification bodies:

Everbridge

- Globally applicable certifications: ISO/IEC 27001:2013, Statement on Standards for Attestation Engagements No. 18's (SSAE No. 18) Service Organization Control 3 (Soc 3)
- United States government certifications: the Support Anti-Terrorism by Fostering Effective Technologies Act (SAFETY Act) of 2002, the Federal Risk and Authorization Management Program (FedRAMP)
- EU privacy and security compliance: General Data Protection Regulation (EU GDPR), C5 (Cloud Computing Compliance Controls Catalogue) developed by the German Federal Office for Information Security (BSI) (BSI C5), Privacy Shield
- United Kingdom government certifications: G-Cloud, Information Commissioner's Office (ICO).

Underwriters Laboratories
Data Centre Certification Programme: Helps mitigate risk for data centre owners and operators by sharing best practices. The certification has six critical components that can be attributable to data centre outages: concurrent maintainability, reliability, security, sustainability, commissioning and safety.

PRISM International
PRISM Privacy+ Certification: Verifies compliance of records and information management services with all known data protection laws and meets customers' regulatory due diligence obligations.

Europrivacy Certification
Delivers gap analysis and GDPR compliance certification.

EuroPriSe
Provides certification for IT products, IT-based services, websites and data processing commissioned in compliance with European data protection law.

Bureau Veritas
Provides data protection certification with GDPR requirements.

3. Interoperability

Interoperability, broadly defined as the ability to share and use data across systems, platforms, locations and jurisdictions, is crucial to overcoming data silos and unlocking the value of shared data. A key step is establishing and adopting standards for data formats and structures.

According to the International Telecommunication Union (ITU), interoperability includes at least three features: a) context information management (cross-domain information metamodel); b) shared data models (domain-specific information models); and c) ecosystem transaction management (conditions for exchange). The ITU also proposed a data process management framework to categorize five types of interoperability by different contexts and dimensions, including the data life cycle, data trust, commercialization, the ecosystem and governance. Data life-cycle dimension interoperability is concerned with technical issues such as communication standards, data format, structure and content standards. The other four are all concerned with technical and non-technical issues, including social, economic, legal and ethical factors.[57]

Standardization efforts are underway to establish the equivalent of international telecommunication standards such as those for the Signalling System 7 (SS7), container manifests (the documents accompanying shipping containers listing their contents prepared by the party filling the container) and Air Traffic Control for the IoT. Yet achieving this for the IoT and machine-generated data involves many more variables. Early cases of out-of-band signalling for machine-generated data can be found in vertical industry sectors. These highlight the need for a universal format to describe sensors, data provenance, regulatory and compliance aspects,

“ Interoperability, [...], is crucial to overcoming data silos and unlocking the value of shared data. A key step is establishing and adopting standards for data formats and structures.

and monetization in a control layer. In software terms, this is defined and included in metadata. There are known standards for metadata, such as ISO/IEC 11179, Digital Object Identifier (DOI) and Resource Description Framework (RDF), to name just a few.

The ITU also set up a framework to support data interoperability in IoT environments which contains three dimensions of data interoperability: syntactical, semantic and object abstraction. The ITU Focus Group on Data Processing and Management to Support IoT and Smart Cities and Communities has provided extensive guidelines for IoT data processing and technical management specifications.[58]

Other organizations and institutions are also actively seeking to establish different standards addressing specific issues in data sharing. The ISO published *Cutting tool data representation and exchange – Part 72: Creation of documents for the standardized data exchange* (ISO Technical Specifications or ISO/TS 13399-72:2016), which defines the necessary text elements of a drawing frame and determines a standardized data exchange format.[59]

Precise interoperability mechanisms can be referred to as Open and Agile Smart Cities Minimal Interoperability Mechanisms (MIMs).[60] The interoperability points of MIMs are threefold: the context information API, which allows access to real-time contextual information, the shared data model with guidelines and catalogues, and the marketplace API, which exposes functionalities such as catalogue management and ordering management.

Other examples of organizations that have implemented different standards to facilitate data exchange are the Institute of Electrical and Electronics Engineers, who launched an initiative to create an autonomous distributed sharing model for a data trading system.[61] Likewise, the Fix Trading Community set up a Market Model Typology Initiative (MMTI) committed to achieving a practical and common solution for standards on post-trade data across all asset classes subject to the Markets in Financial Instruments Directive (MiFID II), a legislative framework instituted by the EU.[62]

4. Accountability and integrity of data

Data accountability is crucial and can be addressed by validating and declaring the data provider/source, evaluating for potential bias and securing the data source and flow's transparency and traceability. This is already the case for static data but will be even more important when it comes to IoT data due to the vast variety of data sources, the real-time nature of data streams, which allows little time for human overseeing or interventions, as well as the scope and scale of potential use cases. Given these characteristics, it is critical that the user is provided with as much contextual information as possible:

- **Validated data provider and source**
  An essential step to ensure the accountability and quality of data, which may include but is not limited to verifying the data source's SQL statement, checking that all required fields are mapped and checking for valid matching rules. Some organizations, such as the Electronic Commerce Code Management Association, will certify and register a company as a Quality Data Provider and compliant with ISO 8000-110:2009, as long as it has demonstrated that it can submit, generate and respond to data queries.

- **Transparency and traceability of data source and flow**
  Traceability – the ability to trace the results back to the original source data and track each transformation or linkage of data – underpins data integrity and the validity of a study based upon a dataset. Data traceability mainly concerns where the data is from and from whom, who is collecting the data and how the data source was manipulated and modified to produce its result. The platform can provide and document data stewardship from its raw form to each of its transformations. When moving and merging data, it is also crucial to ensure that data from different sources and repositories conforms to business rules and is not corrupted due to inconsistencies in type or context. Maintaining the transparency and traceability of data also has key security advantages.

- **Evaluation for potential bias arising from the datasets**
  As biased data fed into an algorithm can lead to biased decision-making, it is important to develop tools and methods to evaluate and measure bias in data. A risk assessment process is particularly necessary for decision-making in public policy and security. Groups and think tanks such as the Future of Privacy Foundation have developed methods to evaluate potential bias in a city's open data and recommend measures to mitigate the bias for transportation safety.[63] The Urban Institute, a United States-based non-profit research organization, is also developing a prototype bias assessment tool to detect bias in geospatial point data to ensure fairness in decision-making related to public matters.[64]

A system of data quality assessment is crucial to guarantee data accountability. The Data Management Association UK has identified six core dimensions to describe the quality of data, which include accuracy, completeness, uniqueness, consistency, timeliness and validity.[65] Regular data quality audits and analyses help ensure the accuracy of data.

5. Eligibility of platform operators

To build trust in the data exchange platform, it is important to ensure that the platform operators have the legal right to purchase, collect, store and exchange the data, and that there are regulatory structures for the certification and periodic review of data brokers through legislation, Charter requirements and other means.

The US Federal Trade Commission recommends that the data broker industry adopts several best practices: "First, they should implement privacy-by-design, which includes considering privacy issues at every stage of product development. Second, the Commission encourages data brokers to implement better measures to refrain from collecting information from children and teens, particularly in marketing products. Finally, the Commission recommends that data brokers take reasonable precautions to ensure that downstream users of their data do not use it for eligibility determinations or unlawful discriminatory purposes."[66] The Commission also calls for laws that enable consumers to learn about the existence and activities of data brokers.

Data brokers' regulations and policy frameworks can vary across regions and under different jurisdictions. Vermont's data broker privacy law (H.764), effective 1 January 2019, is the first of its kind in the US to regulate the businesses that collect, aggregate and sell data about consumers with whom they have no relationship. The law forbids data brokers from acquiring personally identifiable information (PII) with intermediaries by fraudulent means, as well as using the PII to harass, stalk, commit fraud, or engage in unlawful discrimination. It also requires data brokers to register annually with the Vermont Secretary of State, pay a registration fee and provide detailed information about their activities. Similarly, California Assembly Bill No. 1202 requires data brokers who collect and sell personal information about consumers with whom they have no direct relationship to register with the California Attorney General beginning on 1 January 2020, pay a registration fee and provide information including the name of the data broker and its primary physical, email and website addresses.

6. Terms of use agreement for data users, providers and operators

An agreement of terms of use provided by the data-sharing platform will help ensure the practices of data collection, exchange and use comply with laws and regulations and other privacy and security requirements, and define the liability and rights of all parties involved in data exchange and trade activities.

A typical use agreement may provide users with an overview of the platform, including its structure and how the platform works, methods and restrictions for registration if registration is required, as well as agreed methods for the platform to communicate with its users (e.g. consent to receiving notices through registered email addresses). The agreement should define the obligations and rights of the data provider and the platform operator, respectively. This agreement can include which party should be responsible for the legitimacy of data sources and content, whether there is any authorization of rights (e.g. rights to copy, modify, distribute and process data) to the operator along with the transfer of data to the platform and the rights of the data provider to terminate the licence and delete data, as well as other service usage restrictions.

Between the platform operator and the data user, besides basic information regarding the use of the platform, the agreement can define the terms and restrictions of data use, terms of confidentiality, whether there are any warranties about the suitability, completeness, timeliness, reliability, legality, or accuracy of the datasets, as well as the definition of liability and related indemnity resulting from data use.

If the platform is collecting information from its users, it should also obtain their consent and provide information on what data will be collected, how the data will be used and whether and how it will be shared with third-party websites, advertisers and services. The agreement can also define the proprietary rights of materials on the platform, such as software, images, text and graphics, as well as the intellectual property rights of data and intellectual products derived from data use and analysis.

In the Terms of Use agreement for open data platforms such as the Shenzhen Government Open Data Platform,[67] users are authorized to have indiscriminate access to the platform data and the right to use and share this data free of charge. However, users are not authorized to transfer the data acquired from the platform elsewhere. They are obliged to acknowledge the source whenever using the data from the platform and adhere to lawful usage. The platform reserves the right to interrupt the service if the user violates the corresponding regulations and agreements. The platform claims not to guarantee the accuracy and timeliness of data, and not to be liable for any direct or indirect losses resulting from the use of data or any derived analysis. The intellectual property rights of images, text, graphics, codes and other content on the website are protected by law. These terms are typical for open public data platforms, which usually clearly define liability and the user's obligations.

BOX 4 Data governance initiative in China

Shanghai
The Shanghai government proposes data brokers establish a standardized, transparent, secure, controllable and traceable data transaction service environment, formulate transaction service procedures and internal management systems, and take effective measures to protect data security, personal privacy, personal information, trade secrets and confidential business information.[68]

Tianjin
The Tianjin government requires data transaction service agencies to review the authenticity and legality of the identities of both data providers and users, and requests that data providers explain the data sources and keep records of these transactions and reviews.

Shenzhen
The Shenzhen government proposes that the government organizes the formulation of local standards such as those standards for data processing activities, data products and services, data quality, data security, data value evaluation and data governance evaluation. The government will support data-related industry organizations to formulate group standards and industry norms; provide information, technology, training and other services; guide and urge market entities to regulate their data behaviour; and promote the healthy development of the industry. At the same time, market entities are encouraged to formulate data-related corporate standards and participate in the formulation of relevant local and group standards.

To address risk and vulnerability in the data infrastructure, the government of Shenzhen city, China established security requirements for data processors. These include anonymizing personal data and other important data collected by the state and storing it separately from data that can be used to re-establish the identification of specific people. Data processors should also carry out domain-level and hierarchical data storage management and select storage carriers with security and protection features that meet the required security levels. The Shenzhen Government Open Data Platform also need to monitor and give early warnings for abnormal situations such as data leakage, damage, loss and tampering.[69]

Shandong
The Shandong government proposes strengthening the protection of personal information by the following means:

1. Strengthening the supervision and compliance of data sources; regulating the collection of user data by enterprises and institutions; improving the publicity of the rules of collection, processing and the use of personal information and the security of the evaluation systems; improving the channels for complaints and the notification of personal information security incidents; and stopping the illegal collection of personal information.

2. Strictly controlling the phenomenon of excessive collection of personal information by applications, standardizing the type, scope and sensitivity of information collection, and strengthening the application software review.

3. Strengthening the security protection of data activities, data desensitization, data leakage prevention and data encryption before use.[70]

A path forward

As discussed above, to unlock the power of IoT data it is necessary to have in place governance protocols and policy frameworks that foster sustainable, trusted, secure and ethical models of IoT data capture and sharing. Thus, the current white paper proposes a novel framework that highlights six areas for the proper management of data and to enable trust between data providers, platform operators and end-users, protect the rights and interests of all parties and facilitate dynamic IoT data sharing and best uses of data for both public and private interests.

This concluding section provides more detailed instructions for stakeholders and the direction and actions needed to implement the protocol in local policy and regulations.

Suggestions to key stakeholders

City regulator and administration

- Technical: Improve the infrastructure of IoT devices and their application context in cities and build up the IoT network to collect the large amount of IoT data.
- Organization: Set up special departments to manage the issues of IoT data sharing; identify job functions and key performance indicators.
- Standard: Establish relevant standards to facilitate the utility of IoT data.
- Policy: Formulate policies and regulations, management regulations and supervision framework at the regional level and summarize practices and experience for the higher-level government to promote the formation of national policies and regulations.
- Mechanism: Improve the integration and utility of the data both in society and business, and support the establishment of public-private data platforms.
- Mechanism: Support third-party industries, organizations and associations and encourage them to facilitate unlocking the shared value of data.
- Segment selection: Suggest conducting pilot experiments from the perspective of urban smart infrastructure and application – e.g. public transportation, urban energy production and use, smart buildings/communities, environmental monitoring and protection.

Platform

- Emphasize approaches to solve the problem of information asymmetry.
- Build up the mechanism of trust for the supply and demand sides.

Data provider/user

- Compliance and integrity: Create and maintain the sustainable development of a data-sharing mechanism by respecting and following current regulations.

Third-party/industry association

- Coordinate the relationship and contradiction between the demand and supply sides of industry data and promote the self-discipline of industry data sharing.
- Actively engage in the policy discussion and policy-making process.

Suggestions for implementing the protocol

The protocol should be adjusted to local measures and conditions to achieve the maximum public and business interest value. Therefore, by starting from a small scale, from the perspective of pilot cities and industries, experience can be accumulated and then scaled up to form a national IoT data governance regulation and mechanism. It is also necessary to learn from other countries and industries. The following three steps are proposed to carry out the implementation:

Step 1: Pilot

- Select pilot cities/regions and focus on a maximum of two industrial applications with IoT data and data-sharing demands.
- Bring together key stakeholders (e.g. government, businesses, academic experts, industry groups, citizen groups) to form a project team and formulate the specifications of the data platform.
- Enact provisional regulations and conduct policy testing.

Step 2: Expansion

- Expand from one city or urban area to multiple cities and form a cross-city data-sharing platform.
- Expand the application from one industry to multiple industries and form a cross-industry data sharing mechanism.
- Summarize the experience and conduct an exchange of experiences and references on the international platform.

Step 3: Conclusion

- Form a national IoT data-sharing platform and mechanism based on the pilot experience and feedback results to release the value of IoT data within the country.
- Form international agreements to promote global data sharing.

This protocol is expected to create the minimum necessary conditions for a dynamic, secure and trusted data exchange, encouraging the public sector to open IoT data and the private sector to trade IoT data for the benefit of cities and their inhabitants.

Appendix: Global benchmarking of platforms for open data and data sharing

The following list of references is intended to provide an overview of the main open/data-sharing platforms, which can serve as a reference for stakeholders in the design of future platforms and policy development. It should not be considered an exhaustive list.

Name	Country	Name	Country	Website
Australian open government data	Australia	http://data.gov.au/	Open	2013
Adex	Germany	https://www.theadex.com/	Trade	2013
BDEX	US	https://www.bdex.com/	Trade	2014
BODIK	Japan	https://www.bodik.jp/about/	Trade	2013
Brazilian Open Data Portal	Brazil	http://dados.gov.br/	Open	2011
Data Republic	Australia	https://www.datarepublic.com	Open	2016
DataScouts	Belgium	http://www.datascouts.eu	Trade	2014
Dex	Singapore	https://www.dex.sg/	Trade	2016
EverySense	Japan	https://every-sense.com/	Trade	2014
Factual	US	https://www.factual.com/	Trade	2008
Government data Singapore	Singapore	https://data.gov.sg/	Open	2011
Government data UK	UK	https://data.gov.uk	Open	2010
Information Technology Federation of Japan	Japan	https://www.itrenmei.jp/summary/	Trade	2016
Japan Data Exchange	Japan	https://j-dex.co.jp/en/index.html	Trade	2016
Japan Exchange Group (JPX)	Japan	https://www.jpx.co.jp/english/	Trade	2013
Open Data Denmark	Denmark	http://www.opendata.dk/	Open	N/A
Open Government Canada	Canada	http://open.canada.ca/en	Open	2011
Open Government Data (OGD) Platform India	India	https://data.gov.in/	Open	2012
Open Government Data Japan	Japan	https://www.data.go.jp/?lang=english	Open	2014
Open platform for French public data	France	https://www.data.gouv.fr/en/	Open	2011
Qlik	US	https://www.qlik.com	Trade	1993
Quandl	Canada	http://www.quandl.com/	Trade	2011
RapidAPI	US	https://rapidapi.com	Trade	2014
Spaziodati	Italy	http://www.spaziodati.eu/	Trade	2012
The Beijing municipal government data resource network	China	http://tjj.beijing.gov.cn/EnglishSite/	Open	2012
The Danish Data Distributor	Denmark	www.datafordeler.dk	Open	N/A
The National Agriculture and Food Research Organization	Japan	https://wagri.net/	Trade	2017
The Open Data Platform of Zhejiang	China	http://data.zjzwfw.gov.cn/	Open	2015
The Open Data Platform of the Guizhou provincial government	China	http://data.guizhou.gov.cn/home	Open	2016
The Open Data Platform of Shanghai	China	https://data.sh.gov.cn	Open	2012
WingArc 1st	Japan	http://www.wingarc.com/service/dg/	Trade	2004

Contributors

World Economic Forum

Jeff Merritt
Head of Urban Transformation; Member of the Executive Committee

Punit Shukla
Lead, Artificial Intelligence and Machine Learning (2018-2020)

Yamin Xu
Lead of Greater China, Urban Transformation Platform

Fudan University

Lei Zheng
Professor, Head of Lab for Digital & Mobile Governance

Xinping Liu
Deputy Head of Lab for Digital & Mobile Governance

Xinyu Qi
Lab for Digital and Mobile Governance

Government of Denmark

Anders Raahauge
Head of Division for Digital Preservation, Danish National Archives; World Economic Forum Fellow (2018-2019)

McGill University

Xiao Liu
Assistant Professor; World Economic Forum Fellow (2018-2019)

Acknowledgements

Working group members

Sanae Imai
Vice-President, Industry Transformation; World Economic Forum Fellow (2019-2020)

Tadashi Kaji
Senior Chief Researcher, Hitachi; World Economic Forum Fellow (2019-2023)

Hitomi Sanao
Associate Director, Corporate Strategy, Eisai; World Economic Forum Fellow (2019-2020)

World Economic Forum

Arushi Goel
Project Specialist, Data Policy and Blockchain, Centre for the Fourth Industrial Revolution India

Saiful Salihudin
Platform Curator, Urban Transformation

Karla Yee Amezaga
Platform Curator, Data Policy

Project community experts group:

Gengzhen Yan
President, Yida China Holdings; China Representative, SIGFOX

Eric-Yongan Mou
Group Senior Vice-President, Isstech

Yongshuo Xu
Senior Digital Expert, SAP

Yu Zhang
Module Lead, Intel China

Raymond Yu Zhao
Vice-President and General Manager of Sales and Marketing, AIM Stek

We would also like to thank the following for their input:

Nasser Alrayes
Deputy Chief Executive Officer for Smart Cities, Saudi Data and Artificial Intelligence Authority, Saudi Arabia

Aamena Alshamsi
Data Science Lead, Smart Dubai

Will Barkis
Principal, Smart Cities, Orange

Fulvio Bartolucci
General Manager and Co-Founder of Adenergy, Aden

Thais Blumenthal de Moraes
Go to Market Strategy Lead, Waze

Andrew Collinge
Strategic Adviser, Smart Dubai

Lin Fan
Chief Executive Officer, Beijing Taoyou Tianxia Tech

Anyu Fang
Urban Designer, AECOM

Søren Fauerholm Christensen
Head of Division, Danish Agency for Data Supply and Efficiency

Bo Fristed
Head of ITK, City of Aarhus, Denmark

Feng Gao
Director, Open Data China

Eduardo Gómez Restrepo
Data Leader, Centre for the Fourth Industrial Revolution Colombia

Rodolphe Heliot
Vice-President Business Incubation, Schneider Electric

Marie Jordan
Senior Director, Global Standards Management, Visa

Hiromitsu Kato
General Manager, Hitachi

Marianne Knudsen
Land Surveyor, Danish Agency for Data Supply and Efficiency

Xin Lai
Chief Engineer, Thunder Chain, Xunlei

Hyejin Lee
Programme Director, SUGAR Network for Global Design Innovation, Tongji University

Fanyu Lin
Chief Executive Officer, Fluxus

Hongzhi Liu
Senior Vice-President, AECOM

Jinsong Liu
Deputy Director-General, Department of International Economics, Ministry of Foreign Affairs of the People's Republic of China

George Livingston
Principal, Business Group, Orange Silicon Valley

Nakagawa Masaya
Assistant Director, Planning Division, Planning and Coordination Bureau, Kobe City

Per Kolbeck Nielsen
Special Advisor, Danish Agency for Data Supply and Efficiency

Mariam Nouh
Managing Director, Centre for the Fourth Industrial Revolution, Saudi Arabia

Fabien Pfaender
Acting French Dean, Sino-European School of Technology of Shanghai

Hu Qin
Senior Director of Research, Beijing Representative Office, Environmental Defence Fund

Philippe Rapin
Founder and Chief Executive Officer, Urban Radar

Denice Ross
Senior Fellow, National Conference on Citizenship, Fellow at the Beeck

Fumihiko Shimizu
General Manager, Technology Platform Department

Martin Skjold Grøntved
Adviser, Business Development, Danish Agency for Data Supply and Efficiency

Jo Skotte Johansson
Head of Section, Danish Agency for Data Supply and Efficiency

Kim Søvsø
Head of Department ITK City Lab, City of Aarhus, Denmark

Kostas Terzidis
Professor, Tongji University

Mette Vestergaard Dam
Tech and Cyber Adviser, Denmark's Tech Ambassador's Office

Sean Wihera
Vice-President, Business Development and Partnerships, Clarity Movement

Erez Zaionce
Director, Centre for the Fourth Industrial Revolution Colombia

Yunpo Zhang
Executive Director, Paper Research Institute, The Paper

Pengyuan Zhou
Principal, Kearney China

Qi Zhou
Deputy Director, Institute of Zoology, Chinese Academy of Sciences

Endnotes

1. "68% of the world population projected to live in urban areas by 2050, says UN", *United Nations Department of Economic and Social Affairs*, May 2018, https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html#:~:text=Today%2C%2055%25%20of%20the%20world's,increase%20to%2068%25%20by%202050.
2. "Data Creation and Replication Will Grow at a Faster Rate than Installed Storage Capacity, According to the IDC Global DataSphere and StorageSphere Forecasts", *IDC*, March 2021, https://www.idc.com/getdoc.jsp?containerId=prUS47560321.
3. "Data Protection and Privacy Legislation Worldwide", *United Nations Conference on Trade and Development,* n.d., https://unctad.org/page/data-protection-and-privacy-legislation-worldwide.
4. United Nations Habitat, *World Cities Report 2020: The Value of Sustainable Urbanization,* 2020, https://unhabitat.org/World%20Cities%20Report%202020.
5. "11-Sustainable Cities and Communities", *The Global Goals*, n.d., https://www.globalgoals.org/11-sustainable-cities-and-communities.
6. "How Saudi Arabia is deploying ICTs against COVID-19 – and beyond", *ITU*, July 2020, https://www.itu.int/hub/2020/07/how-saudi-arabia-is-deploying-icts-against-covid-19-and-beyond.
7. "About Tawakkalna", *Tawakkalna*, n.d., https://ta.sdaia.gov.sa/en/index.
8. Ibid.
9. Ibid.
10. Marr, Bernard, "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read", *Forbes*, 21 May 2018, https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#692ca36860ba.
11. Russo, Massimo and Tian Feng, "The Risks and Rewards of Data Sharing for Smart Cities", *BCG*, 10 August 2020, https://www.bcg.com/publications/2020/smart-cities-need-to-understand-the-risks-and-rewards-of-data-sharing-part-3.
12. McKinsey Global Institute: McKinsey & Company, *The Internet of Things: Mapping the Value Beyond the Hype*, June 2015, https://www.mckinsey.com/~/media/McKinsey/Industries/Technology Media and Telecommunications/High Tech/Our Insights/The Internet of Things The value of digitizing the physical world/The-Internet-of-things-Mapping-the-value-beyond-the-hype.pdf.
13. "Data privacy: What the consumer really thinks", *DMA Group*, February 2018, https://dma.org.uk/uploads/misc/5a857c4fdf846-data-privacy---what-the-consumer-really-thinks-final_5a857c4fdf799.pdf.
14. Open Data Charter, "International Open Data Charter: Principles", n.d., https://opendatacharter.net/principles.
15. Demchenko, Yuri, Wouter Los and Cees de Laat, "Data as Economic Goods: Definitions, Properties, Challenges, Enabling Technologies for Future Data Markets", *ITU Journal: ICT Discoveries,* 2018, No. 2, https://www.itu.int/en/journal/002/Documents/ITU2018-12.pdf.
16. "Starting an Open Data Initiative", *World Bank*, n.d., http://opendatatoolkit.worldbank.org/en/starting.html -:~:text=Transparency.,public budget expenditures and impacts.
17. United Nations Department of Economic and Social Affairs, *United Nations E-Government Survey 2018*, 2018, https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2018-Survey/E-Government Survey 2018_FINAL for web.pdf.
18. "The world's most powerful data lives on Quandl", *Quandl*, n.d., https://demo.quandl.com/.
19. "About IDC", *International Data Corporation*, n.d., https://www.idc.com/about.
20. "Main page", *LinkNYC*, n.d., https://www.link.nyc/.
21. Russo, Massimo and Tian Feng, "The Risks and Rewards of Data Sharing for Smart Cities", BCG.
22. The Danish Government/Local Government Denmark, *The eGovernment Strategy 2011-2015: Good Basic Data for Everyone- A Driver for Growth and Efficiency*, October 2012, https://en.digst.dk/media/18773/good-basic-data-for-everyone-a-driver-for-growth-and-efficiency.pdf.
23. "About WAGRI", *WAGRI*, n.d., https://wagri.net/en-us/aboutwagri.
24. "Moonshot, Mercado de datos para el bien común", *Centre for the Fourth Industrial Revolution Colombia*, n.d., https://c4ir.co/?page_id=2549.
25. "Data for Common Purpose Initiative (DCPI)", *World Economic Forum,* n.d., https://www.weforum.org/projects/data-for-common-purpose-initiative-dcpi.
26. Hevens, John C., "Heartificial Intelligence: Embracing Our Humanity to Maximize Machines", *Penguin Random House*, 2016, pp. 118-126.
27. The Hub of All Things, *Hub of All Things*, n.d., https://www.hubofallthings.com/main/what-is-the-hat.

28. Ministry of Housing and Urban Affairs-Government of India, "India Urban Data Exchange (IUDX)", https://smartcities.gov.in/India_Urban_Data_Exchange.
29. World Economic Forum, *Towards a Data Economy: An enabling framework*, August 2021, https://www3.weforum.org/docs/WEF_Towards_a_Data_Economy_2021.pdf.
30. "India Urban Data Exchange (IUDX)", Ministry of Housing and Urban Affairs-Government of India, n.d., https://smartcities.gov.in/India_Urban_Data_Exchange.
31. "Everysense Pro", Everysense, n.d., https://every-sense.com/.
32. "Create the best data experiences", *Opendatasoft,* n.d., https://www.opendatasoft.com/.
33. Outlier Ventures, *Convergence in Smart Cities: Building the Digital Infrastructure for the Fourth Industrial Revolution,* 2019, https://www.smartdubai.ae/docs/default-source/default-document-library/convergenceinsmartcities_en.pdf.
34. "Open Data for All New Yorkers", *NYC Open Data*, n.d., https://opendata.cityofnewyork.us/.
35. "Intertrust Platform, For a new era of commerce", *Intertrust*, n.d., https://www.intertrust.com/platform/.
36. Russo, Massimo and Tian Feng, "The Risks and Rewards of Data Sharing for Smart Cities", BCG.
37. Ibid.
38. "Portland Urban Data Lake (PUDL)", Portland Bureau of Transportation (PBOT), n.d., https://www.portlandoregon.gov/transportation/article/681572.
39. "Seoul to set up "S-Data," an integrated storage for all public data", *Seoul Metropolitan Government*, 7 November 2019, http://english.seoul.go.kr/seoul-to-set-up-s-data-an-integrated-storage-for-all-public-data/.
40. A federated data infrastructure is one in which each node forms an autonomous unit, provides interfaces for simple and secure data exchange, enables the use of third-party applications and functions, and allows easy data integration thanks to standards. As part of this infrastructure, certification of service providers, nodes and services is recommended, covering areas such as IT security, service levels, the degree of data sovereignty achieved and the conditions of the contractual framework. More information here: Federal Ministry for Economic Affairs and Climate Action, *Project GAIA-X: A Federated Data Infrastructure as the Cradle of a Vibrant European Ecosystem*, 2019, https://www.bmwi.de/Redaktion/EN/Publikationen/Digitale-Welt/project-gaia-x.html.
41. Ibid.
42. "Data Exchange, unleash the value of your data", *Dawex*, n.d., https://www.dawex.com/en/.
43. "Main page", *Shanghai Data Exchange*, n.d., https://www.chinadep.com/.
44. Ibid.
45. "Establishment of the Study Group for Ideal Approaches to a Certification Scheme Concerning Functions of Information Trust", *Ministry of Economy*, Trade and Industry (METI), n.d., https://www.meti.go.jp/english/press/2017/1106_004.html.
46. "Release of the Guidelines of Certification Schemes Concerning Functions of Information Trust ver. 1.0", Ministry of Economy, Trade and Industry (METI), n.d., https://www.meti.go.jp/english/press/2018/0626_002.html.
47. "Japan grants certification for first time to 'information banks'", *The Japan Times*, 9 July 2019, https://www.japantimes.co.jp/news/2019/07/09/business/japan-grants-certification-first-time-information-banks/
48. Russo, Massimo and Tian Feng, "The Risks and Rewards of Data Sharing for Smart Cities", BCG.
49. "Collaboratively Exploring the Public Value of Technology in Boston", *Beta Blocks*, n.d., https://betablocks.city/.
50. For a comparison of the GDPR and CCPA, see: "Comparing Privacy Laws: GDPR v. CCPA", *Data Guidance and Future of Privacy Forum*, n.d., https://fpf.org/wp-content/uploads/2018/11/GDPR_CCPA_Comparison-Guide.pdf.
51. The Information Security Technology – Personal Information Security Specification (GB/T 35273-2020) (PI Specification) was proposed by the National Information Security Standardization Technical Committee (TC260) on 6 March 2020, as an amendment to and replacement for the November 2017 version (GB/T 35273-2017). The PI Specification took effect on 1 October 2020. For more information, please see: "Personal Information Security Specification" English Version Announced", *National Information Security Standardization Technical Committee*, 20 September 2020, https://www.tc260.org.cn/front/postDetail.html?id=20200918200432.
52. "Information security technology – Personal information (PI) security specification", *National Standard of the People's Republic of China*, 6 March 2020, https://www.tc260.org.cn/upload/2020-09-18/1600432872689070371.pdf.
53. Interagency International Cybersecurity Standardization Working Group, *Interagency Report on the Status of International Cybersecurity Standardization for the Internet of Things (IoT)*, November 2018, https://doi.org/10.6028/NIST.IR.8200.
54. The AWS Well-Architected Framework recommends some best practices for cloud-based architecture and protecting data in transit. See: "SEC 10: How do you anticipate, respond to, and recover from incidents?", *AWS*, n.d., https://wa.aws.amazon.com/wat.question.SEC_10.en.html.
55. European Union Agency for Cybersecurity, *Recommendations on European Data Protection Certification*, 27 November 2017, https://www.enisa.europa.eu/publications/recommendations-on-european-data-protection-certification.
56. ISO, "Certification", n.d., https://www.iso.org/certification.html.

57. "Technical Specification D3.3 – Framework to support data interoperability in IoT environments", *International Telecommunication Union (ITU)*, n.d., https://www.itu.int/pub/T-FG-DPM-2019-3.3.

58. "Focus Group on Data Processing and Management to support IoT and Smart Cities and Communities", *ITU*, n.d., https://www.itu.int/en/ITU-T/focusgroups/dpm/Pages/default.aspx.

59. "Cutting tool data representation and exchange – Part 72: Creation of documents for the standardized data exchange – Definition of properties for drawing header and their XML-data exchange", *ISO*, July 2016, https://www.iso.org/standard/66794.html.

60. "Annex 1: Minimal Interoperability Mechanisms (MIMs)", *Open and Agile Smart Cities*, 16 January 2019, https://oascities.org/wp-content/uploads/2019/06/OASC-MIMs.pdf.

61. "Data Trading System Initiative", *IEEE Standards Association*, n.d., https://standards.ieee.org/industry-connections/datatradingsystem.html.

62. "Market Model Typology", *Fix Trading Community*, n.d., https://www.fixtrading.org/mmt/.

63. Safe-D Safety Through Disruption, *Sources and Mitigation of Bias in Big Data for Transportation Safety*, November 2018, https://safed.vtti.vt.edu/wp-content/uploads/2020/07/02-026_Final-Research-Report_Final.pdf.

64. Narayanan, Ajjit and Graham MacDonald, "Toward an Open Data Bias Assessment Tool", *Urban Institute*, 5 March 2019, https://www.urban.org/research/publication/toward-open-data-bias-assessment-tool.

65. "Meet the data quality dimensions", *United Kingdom Government*, 24 June 2021, https://www.gov.uk/government/news/meet-the-data-quality-dimensions.

66. Federal Trade Commission, *Data Brokers. A Call for Transparency and Accountability*, May 2014, https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

67. See the user agreement at: "Shenzhen Government Data Open Platform Terms of Service", *Shenzhen Government Service Data Administration*, n.d., https://opendata.sz.gov.cn/maintenance/forward/toTermOfService.

68. "Shanghai Data Regulations", *Shanghai Municipal People's Government,* 29 November 2021, https://www.shanghai.gov.cn/nw12344/20211129/a1a38c3dfe8b4f8f8fcba5e79fbe9251.html.

69. "The full text of the "Shenzhen Special Economic Zone Data Regulations" has been announced!", *SZ News*, 7 July 2021, http://www.sznews.com/zhuanti/content/2021-07/07/content_24368291.htm.

70. "Digital Province Construction Plan" (text in Chinese), Shandong Province, n.d., http://gxt.shandong.gov.cn/module/download/downfile.jsp?classid=0&filename=89c86dd3cc7b4f07a550497344079b99.pdf.

WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744
contact@weforum.org
www.weforum.org