International Standard

ISO/IEC 17917

Smart cities — Guidance to establishing a decision-making framework for sharing data and information services

Villes intelligentes — Recommandations pour l'établissement d'un cadre décisionnel pour le partage des données et des services d'information

First edition
2024-06

Reference number
ISO/IEC 17917:2024(en)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document is based on BSI (British Standards Institution) (as BSI PAS 183:2017) and drafted in accordance with its editorial rules. It was assigned to Joint Technical Committee ISO/IEC JTC 1, *Information technology*, and adopted under the “fast-track procedure”.

Any feedback or questions on this document should be directed to the user’s national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Contents

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see http://patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, as a fast track draft standard.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

0 Introduction

0.1 General

The term “smart city” denotes the effective integration of physical, digital and human systems in the built environment to deliver a sustainable, prosperous and inclusive future for its citizens. A basic assumption in the design of a smart city is the ability of the physical and digital systems to be interoperable. This standard gives governance guidance for decision-makers on establishing a decision-making framework for sharing city data and creating interoperable information services.

Data has the ability to transform the city and its services, providing visibility of the services available, and supporting citizen interactions with those services. Improving the design and integration of city services can serve the public better and drive innovation and efficiencies.

This standard aims to support data sharing in cities and between cities, and the establishment of data sharing agreements, particularly where data is being shared by multiple organizations to transform the delivery of city services.

Missing data or misinterpretation of data can lead to the wrong actions being taken by city decision-makers. A decision-making framework for sharing data can help ensure that they have the best overall data ecosystem on which to base decisions.

Sharing data across a city requires more than the interoperability covered by the smart city concept model (SCCM) defined in ISO/IEC 30182, which focuses by necessity on the semantics of data in a city. Full data interoperability requires a

data framework to be created across the entire spectrum of data for a city: open, closed and shared data.

This standard builds on the integrated operating model defined in ISO/IEC 37106:2021 and assumes that the governance of a smart city programme and the overall management of the city’s data assets has been understood and agreed upon by city leaders and decision-makers from the organizations delivering city services.

The value of data sharing has yet to be explored by cities, as data is predominantly currently used for a specific purpose related to the public task, additionally data

is not viewed as an essential city asset which can be used to transform a city. Data can also provide the basis for new commercial models in smart cities.

This standard defines the data framework for sharing city data to enable discussions between the specialists who build and design the physical and digital services and the decision-makers using data to transform their city.

This standard is for use by decision-makers in smart cities from the public, private and third sectors. It is also of interest to any city organization wishing to share data.

It is expected that each city will create a decision-making framework based on this standard to address its own challenges and opportunities, taking into account the priorities and needs of their city. The creation of a data ecosystem based on the interoperability and data sharing principles in this standard could create data assets that are used to improve the quality of life for citizens and create sustainable commercial models to fund innovation.

0.2 Relationship to other smart city standardization documents

0.2.1 ISO 37106

This standard has been built on the guidance in ISO 37106, *Sustainable cities and communities* – Guidance on establishing smart city operating models for sustainable communities . The particular components of a smart city framework which apply are:

a) [B2] Transforming the city's operating model with particular reference to the governance model developed and any vulnerabilities of both data and city services;

b) [B6] Establishing a common terminology and reference model; and

c) [B10] Identity and privacy management.

This standard is guidance to help with the implementation of these components of the smart city framework.

0.2.2 ISO/IEC 30182

The smart city concept model (SCCM) described in ISO/IEC 30182, *Smart city concept model – Guide to establishing a model for data interoperability* addresses the data interoperability issues that arise as a result of each sector and/or service in a city having its own model and terminology that it uses for data. This standard defines the data framework that addresses the other areas that affect interoperability, such as access rights, privacy, availability and formats. These other areas are also barriers to interoperability and portability which impact the design of the physical and digital services.

This standard addresses the barriers other than the semantics addressed in ISO/IEC 30182 , to enable data interoperability and portability, and the sharing of data and information services in a smart city.

The data framework identifies all elements which will be needed to deliver the four key types of insight when data and services are appropriately shared:

operational, critical, analytical and strategic insight. (See ISO/IEC 30182:2017, Clause 0).

0.3 Relationship to building information modelling (BIM) standards

The following documents are considered to be the foundational standardization documents to be used as part of a whole lifecycle approach to the built environment for BIM Level 2 in smart cities.

This standard assumes that the ISO 19650 series is used for all BIM Level 2 building and infrastructure assets in a smart city and that asset procuring organizations use them as part of their overall digital and smart strategies.

- ISO 19650-1, Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) – Information management using building information modelling – Part 1: Concepts and Principles;
- ISO 19650-2, Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) – Information management using building information modelling – Part 2: Delivery phase of the assets;
- ISO 19650-3, Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) – Information management using building information modelling – Part 3: Operational phase of the assets;

- ISO 19650-4, Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) – Information management using building information modelling – Part 4: Information exchange;
- ISO 19650-5, Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) – Information management using building information modelling – Part 5: Security-minded approach to information management.

1 Scope

This standard gives guidance on establishing a decision-making framework for sharing data and information services in smart cities.

It covers:

a) types of data in smart cities;

b) establishing a data sharing culture;

c) data value chain – roles and responsibilities;

d) purposes for data use;

e) assessing data states;

f) defining access rights for data; and

g) data formats/format of transportation.

This standard aims to support the sharing of data and information services within cities. For some cities there will also be a need to establish specific data sharing agreements, particularly where data is being shared by multiple organizations at once.

This standard supports a transparent approach to making decisions and creating specific data sharing agreements in order to fully realise the benefits and value of data and information services in a city.

Missing data or misinterpretation of data can lead to the wrong actions being taken by city decision-makers. A decision-making framework for sharing data can help ensure that they have the best overall data on which to base decisions.

This standard does not cover:

a) national security issues;

b) good practice for use of data by the citizen;

c) existing interoperability agreements between cities;

d) defining application programming interfaces (API) networks; or

e) any data sharing rules and regulations specific to a particular jurisdiction. It is assumed that a security-minded approach to data sharing is used by cities.

NOTE 1 Further details on the areas not covered in this standard, including information on relevant standards publications, are given in Annex A.

This standard is for use by decision-makers in smart cities from the public, private and third sectors. It is also of interest to any city organization wishing to share data.

2 Terms and definitions

For the purposes of this standard, the terms and definitions given in ISO 37100 and the following apply.

2.1 closed data

data which has been restricted for use

2.2 data

recorded information

[SOURCE: ISO 22005:2007, **3.11**]

2.3 data framework

classification of data assets as either metadata, reference or thematic data

2.4 data spectrum

differentiation of data assets on the basis of whether they are considered closed, sharable or open

2.5 data value chain

intelligent use, management and reuse of data to deliver insight

2.6 derived data

data item used in analysis and/or tables derived from one or more source data items and/or categories

[SOURCE: ISO 20252:2012, **2.2**]

2.7 metadata

data that defines and describes other data [SOURCE: ISO 24531:2013, **4.32**]

2.8 open data

data that can be freely used, modified, and shared by anyone for any purpose

2.9 reference data

data that defines the set of permissible values to be used by other data fields

2.10 thematic data

patterns of data within the data framework that are deemed important to support the provision of city services and the four levels of insight in the city

3 Data Sharing

The SCCM detailed in ISO/IEC 30182 is a basis for understanding the semantic interoperability of data in a smart city. This governance standard assumes the city has utilized the concepts in the SCCM as the basic building block of data interoperability in their city. This use of the SCCM can enable the four levels of insight – operational, critical, analytical and strategic – to be achieved from the data.

However the use of the SCCM guidance alone does not address all the barriers to data sharing in the city. Other aspects of data also need to be understood to create an appropriate framework for data sharing in a smart city.

To achieve effective data sharing, seven key data sharing areas are deemed to be additional barriers and should be considered individually and collectively by the city:

a) types of data;

b) establishing a data sharing culture;

c) data value chain;

d) purposes for data use;

e) assessing data states;

f) defining access rights for data; and

g) data formats/format of transportation.

This standard guides cities on these areas to address in order to establish a data framework to share data. The sharing models which are covered within the scope of this standard are:

a) public sector and public sector;

b) public sector and local business or community;

c) public sector and citizen;

d) business or community and business or community; and

e) business or community and citizen.

Whilst this standard does not cover data sharing between citizen and citizen, cities should make provision to allow a citizen to make requests related to the data it holds and shares. The city should also be able to respond to this request in a timely manner and keep an audit trail of these interactions.

4 Types of data

4.1 General

As local authorities transition to becoming smart cities, existing data assets form the initial data framework that is used as the initial evidence base for decision-making based on data. The city collects, processes and validates data for the essential operation of services provided to citizens. This city data estate resides largely in disconnected legacy systems which are cumbersome and costly to change and cannot be operated in new ways. Some investment is required in technology to unlock the value of data that resides in these legacy systems. This transition is not technology-led – rather technology is an enabler – it is data-led.

This is to allow the existing city data to continue adding value, alongside the new data which the city creates, for example from sensors forming part of its new transport infrastructure.

Understanding the data assets of a city is the first step in creating value from the data and maximizing the value of the data assets to the city. Concepts, of themselves, are not sufficient to derive value from city data and it is important to understand that the physical location of the data – the technology in which the data resides – is not the issue. Irrespective of the source and state of the data assets in a city, a common data framework can be created that reflects the data estate from which the city can derive value. This requires a data-centred approach – a new way of thinking about data – that develops the SCCM model articulated in ISO/IEC 30182.

Unlocking value from data requires the city to understand the value that can be created from data beyond the dataset approaches which prevail within cities, and indeed across regions, territories, and nations.

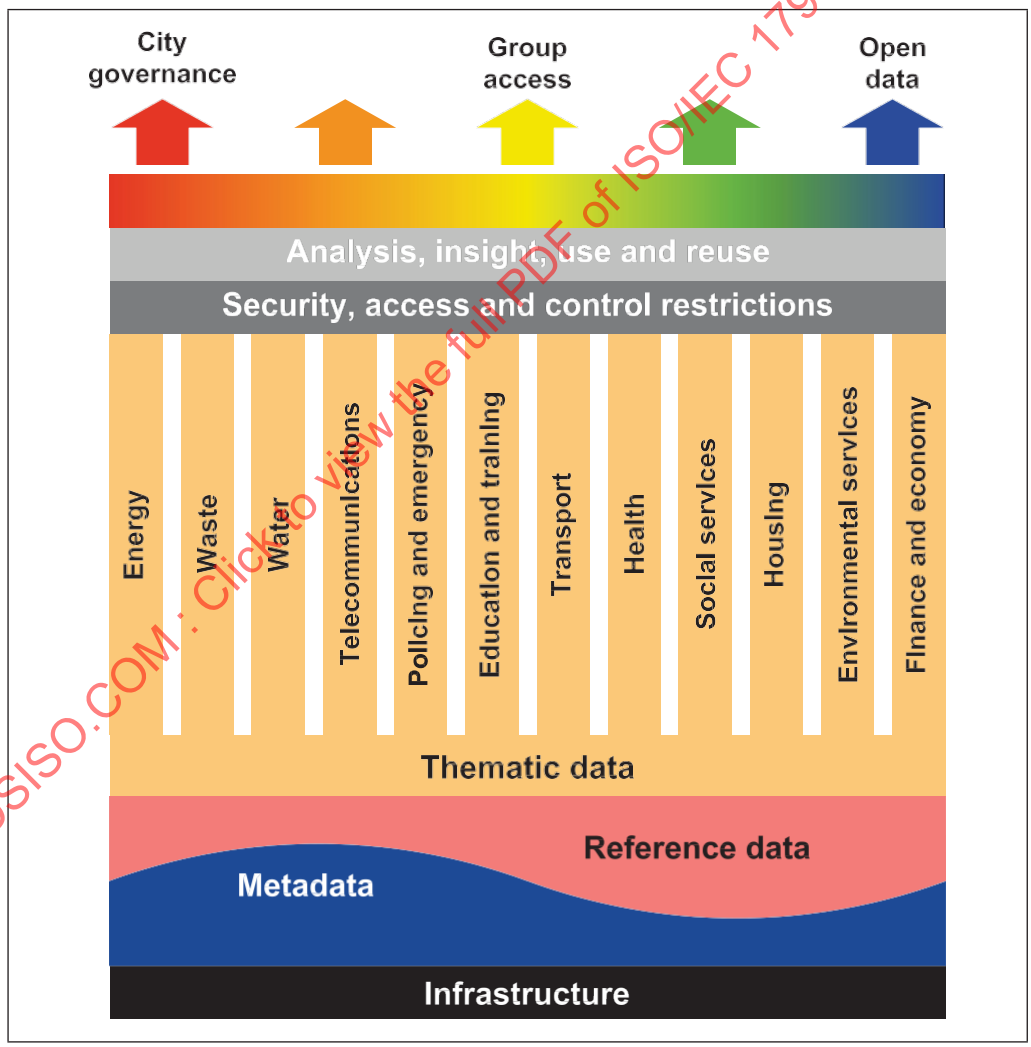
4.2 The data framework

4.2.1 General

As shown in Figure 1, the data framework for a smart city classifies the data assets as either metadata, reference data or thematic data. The data framework shows how current city data assets are transitioned from the existing siloed service provision to an interoperable data estate. The data framework also supports the collection, processing and analysis for future heterogeneous data streams that will become the norm as we transition to a connected Internet of Things landscape.

The data framework supports the active management of data across the entire data lifecycle (see Clause **7**, where the data value chain is covered in detail).

Figure 1 **Data framework**



4.2.2 Infrastructure

Infrastructure is a system of facilities, equipment and services required for the operation of a city. This encompasses both physical and data infrastructure, and all resources including technical, supporting services and people which are required to support the successful delivery of city services.

Governance of infrastructure in a smart city is complex due to the multi-agency operating model which can comprise of many organizations who are likely to have differing approaches to Information Technology (IT) governance. ISO/IEC 38500 contains the guiding principles each organization providing smart city infrastructure can use. It focuses on both the current and future use of IT including management processes and decisions related to the current and future use of IT.

4.2.3 Metadata

Metadata is used to summarize basic information about data to enable it to become more easily discoverable by both humans and computers. Metadata can facilitate understanding of the provenance of data within the data framework and support appropriate data policies, licenses and regulation. An example of metadata in a smart city data framework is the data relating to the voluntary services organizations who deliver city services on behalf of the city to citizens. Metadata exists in all cities but the availability of metadata might differ depending on the size of the city, whether the city is predominantly urban, rural or is in a coastal setting, and the maturity of the city's data framework.

Metadata has the added value of being able to be used for analysis and comparison purposes across a number of cities, or indeed at national level. When metadata is shared, consideration should be given to any metadata which has been created and the data framework updated to reflect the security, access and control rights to be considered for the metadata created.

4.2.4 Reference data

Reference data usually consists of a list of permissible values and/or textual descriptions and is used by a business process to drive value from data. This requires an agreed vocabulary approach in order for this data to support the business processes across the many services, organizations and departments in a city. In a smart city data framework there are organizations who supply a number of services to the city, or products/services which are operated as shared services by the city. An example of smart city reference data is that of a vehicle specification about a transport fleet used to determine its suitability to meet a number of service needs. This reference data is unlikely to be unique to a specific city, however because it is relevant to a specific service provision and might be based on legacy approaches, it is unlikely to be useful for cities to use for comparison or benchmarking across cities.

4.2.5 Thematic data

Thematic data in a city will initially be the datasets and legacy data that are created, processed and managed by the city in order to deliver services to citizens, such as the data related to the provision of adult care services. The metadata and reference data within the data framework with the thematic data supports the city as it moves towards the provision of citizen-centric services for adults with all data becoming part of the wider healthcare service data for the city. At this point, it becomes possible to consider attributes that exist across a number of city services, allowing a city to understand the characteristics of the thematic data and what constitutes a set of data for the city service. In this scenario, healthcare service data forms a set of data in a city which of itself has specific value, and allows a city to understand the challenges and opportunities which exist.

4.3 The data spectrum

4.3.1 General

In order to understand how a city can maximize the value of its data, it is important that the data framework classifies data for use and also differentiates the data it holds on the basis of whether it is considered closed, shareable or open. The extent to which the restrictions have been implemented can vary dependent on the security, access and control requirements. The use of data within the data spectrum is restricted to the use, reuse and the purpose for which data can be shared. ISO 31000, *Risk management - Principles and guidelines* outlines good practice on the management, assessment and analysis of risk and can be used by cities when implementing the data framework.

An appropriate risk management regime for the sharing, publishing and reuse of data should be established and implemented.

4.3.2 Closed data

Closed data is data which has been restricted for use. This data has been designated as information that is not permitted to be shared. In a city, this data includes payment details for citizens within a specific service, such as their council tax.

4.3.3 Shared data

The shared component of the data spectrum is the data which exists which cannot be considered as either open or closed data. This varies between cities and is assumed to represent the majority of the data in a city.

This standard looks in detail at:

a) the suitability of sharing data for new purposes (see Clause **9**); and

b) access rights to data (see Clause **10**).

It is important as part of the data spectrum to understand there are three top level access restrictions which apply to shared data:

1) specific access is when data is made accessible by the data owner to either named individual(s) or named organization(s);

2) group access is when data is made available to specific groups of people or organization(s) based on predetermined criteria; and

3) public access is when data is made available publicly but only under certain terms and conditions that cannot be considered open.

Publishers of city data have a duty of care when restricted data is considered for sharing to ensure that potential harm to individuals or assets is considered prior to publication. An example of shared data such as this is COMAH (control of major accidents and hazards) data.

Case study 1 in Annex A gives an example of the benefits of shared data with public access.

4.3.4 Open data

This standard uses the definition of open that is maintained by the Open Project [3]. "Open" means anyone can freely access, use, modify and share for any purpose (subject at most to requirements that preserve provenance and openness). This definition is also used to determine whether data can be classified as open data.

Case study 2 in Annex A gives an example of a city publishing open datasets.

4.3.5 Data usage

For all types of data within the data framework, it is important to consider the data usage when deciding on the value that is created from that data. Metadata and reference data determine the provenance of the data and in particular its state, for example raw, processed or archived. In some cases it is appropriate to use static data (a snapshot of data created to be used for some future purpose). Temporal data (data that varies over time) might also be used to create snapshots of scenarios in a city for future use. Versioned data (data that represents a data update cycle) might also inform scenarios and should also be considered in a city.

Consideration should also be given to the data structures used within the data framework when determining the data usage. This is covered in more detail in Clause **11**. Temporal and versioned data is important in order to understand the patterns of data which form the four key types of data insight required when sharing data in a city:

a) operational insight examines the characteristics of things such as buildings, communities and organizations, using data to evidence and improve their value for the city or deliver a service;

b) critical insight is gained from the real-time monitoring of incidents and current cases, involving all relevant organizations from across sectors, who work together to achieve the desired outcome or response or deliver a service;

c) analytical insight is the exploration of the data framework to determine patterns, correlations and predictions, allowing the development or innovation of systems or services, or the evidencing of challenges and opportunities for the city; and

d) strategic insight is the overarching approach that examines outcomes related to strategic objectives, decisions and plans.

4.4 Derived data

One aspect of smart city data sharing that is still largely untapped is the role of derived data and how it can contribute to the value of city data. Derived data has traditionally been created to support the performant processing of data attributes for processing. Derived data is when one or more measures used by a city are combined from one or multiple datasets to create new data attributes, which are then used during the exploration of data and appear in any resultant analysis.

As part of the processing that is performed to achieve the four key types of data insight, derived data is created as snapshots for particular scenarios, allowing the appropriate data to be explored for insight. Derived data is a key resource that a city uses to understand and respond to the challenges and opportunities experienced. For example, education data might be used along with transport data to gain strategic insight from the derived data about potential impacts on education provision when planning major city infrastructure initiatives.

Data creators and decision-makers should be aware that there are risks associated with the aggregation of data, by accumulation and/or association. This might result in derived data being created that relates to, or reveals sensitive information. It is important to ensure that any change or creation to the derived data is re-examined in order to ensure whether it is allowed to be shared or published.**Commercialization of data**

4.4.1 General

The open data movement has articulated that making data available free and open at the point of use adds value for the owner of that data and allows entrepreneurs to extend existing businesses or begin new companies to monetize the open data by adding value. The data within the data framework which should be considered for

commercialization extends beyond open data to consider shared data. These valuable resources created from the smart city's data framework require curation and add significant value by creating social, economic and environmental resources for the public good. The data created in a city are not just created by the city but by organizations which support the delivery of services in the city from either

the private or third sector. Putting the data into the data framework creates visibility and facilitates unforeseen exploitation. The use of this shared data as part of a commercialization model will require negotiations with all the data creators involved. To ensure the continued curation and publication of this data – particularly at city scale – it will require a new approach that provides sustainable funding. Sustainable funding would address the ongoing cost associated with the evolution of city data and its usage.

4.4.2 Demand-led model

The initial supply-driven approach to making open data available, as a result of the work of the Open Data User Group (ODUG) has been superseded not just in the UK but globally. The new approach is a demand-led approach to the

curation and publication of open data. This demand-led approach is based on the publication of data for which there is an audience who can articulate the value of the release of the data.

Case study 3 in Annex A gives examples of some of the benefits of a demand-led model.

In a smart city context, a new approach is needed to data curation and publication if the initial building block created by the release of open data from cities is to be extended to include not just open data, but shared data.

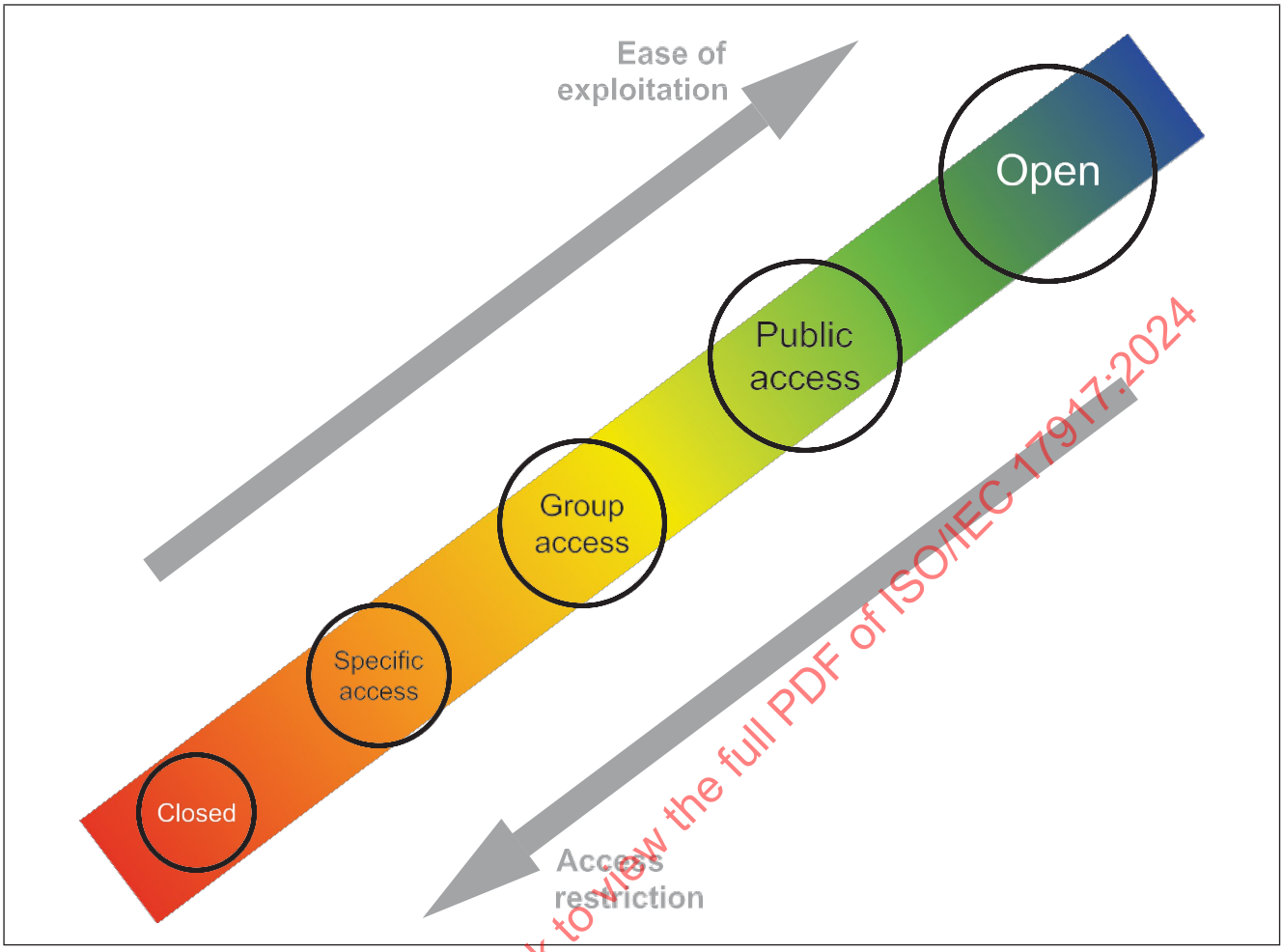
There is an opportunity for cities to evolve commercial models for data that they curate and publish. This commercial opportunity exists for the combination of open and shared data and business models should be used to add value beyond the four levels of insight (see Clause **4**) for the city and city services. The city should examine business models that build on the open data already released and include any or all of the shared access data that the city holds, where access is currently restricted.

As evidenced with the open data agenda, a supply-led approach to these new models is unlikely to lead to a commercial funding model that contributes to or wholly funds the curation and publication of data shared commercially by a city.

A demand-led approach should be used for city shared data; this can also highlight the commercial opportunities to collect, process, curate and publish new shared city data.

As depicted in Figure 2, the commercialization opportunities diminish in direct proportion to the closing of data. In a city where the curation and publication approach results in the majority of data that could be shared remaining closed, this limits the commercial opportunities. By contrast, a city which explores the opportunities which exist with both open and shared data will be able to evolve new commercialization models to fund the data framework.

Figure 2 **Commercial opportunities for the data spectrum**



4.4.3 Commercial models

One of the keys to unlocking the demand-led open data agenda was the need to understand who consumed the data that had already been released, and to provide a mechanism for any citizen to request more data based on the presentation of a use case. Many countries are now using the demand-led approach to open data pioneered by ODUG. This approach has the additional benefit for cities in that it prioritizes requests for data with the data framework.

The creation of an evidence base that understands the demand, showcases what data exists and makes that data discoverable, helps a city determine the commercial value of shared data. This can allow commercial models based on incentivizing the safe and secure sharing of data to be developed.

Case study 4 in Annex A showcases the exploration of a commercial model. Not all of the data in the data framework is necessarily chargeable: the city as the authoritative source of this data should consider which data might be released as searchable and open, as this evidence of demand can highlight further commercial opportunities.

4.4.4 Research-based model

A research-based approach to commercial opportunities for city data should also be explored. This requires an investment to explore the data that a city holds using statistical, mathematical and data science techniques, however this investment could

enable incentives for sharing to be understood and possible commercial models that could be applied to emerge. Case study 5 gives an example of the potential benefits of a research-based approach to urban planning.

The outcomes of data sharing can be quantified using a pre-determined scale. Using a scale around outcomes can support a city's ability to monitor and report on data sharing outcomes, to encourage appropriate changes to the approach. This mechanism could be as simple as a 1-5 scale of effectiveness / impact. A 5, for example, would indicate that the particular outcome of data-sharing provides a significant benefit that was otherwise not able to be realised. Adding elements which track effort and/or risk might also be useful.

Whichever approach is taken to understand the business model that maximizes the opportunity to commercialize shared data, it is important that the outcome of data sharing is measurable.

5 Establishing a data sharing culture

5.1 General

The result of the city transitioning to managing data as an asset in its own right and in collaboration with other significant data owners in the city is a fundamental change to the way in which legacy data has been viewed by the city authorities.

Guidance note [B2] from ISO 37106 - Transforming the city's operating model should be used as a starting point for understanding how to manage data as an asset.

This recommends specific data leadership within the city and coordination of all the owners of data within the data framework. Coordination of activities to

establish a data sharing culture should have appropriate focus on understanding the demand for the data and the most appropriate mechanisms for ensuring that data is discoverable.

When sharing or publishing data, particular care should be taken to address safety and security considerations. In particular, steps should be taken to identify and protect information that could impact on the safety and security of individuals, services, sensitive assets and systems, and the benefits which city assets and systems exist to deliver.

5.2 Identifying the benefits of data sharing

As part of the transition to a data sharing culture, the city should articulate the benefits of this transition to all organizations with a role in the new city operating model. The benefits will initially accrue from having a single view of city data via the data framework.

There are many technical options to achieve this transition and it is not necessary to combine data in a single technical implementation, provided all the data is accessible for discovery. Indeed, as cities transition from using structured data - which is the predominant city data prior to the adoption of the IoT - to curating IoT and other streaming data - which is either unstructured or semi-structured - NoSQL technologies and cloud services built for this data will need to be used.

From a technical viewpoint, what is important is that technologies used are loosely coupled and interoperable, to prevent the technology being a barrier to unlocking the benefits of data sharing.

Irrespective of the technologies employed, the value of the data framework is derived from harnessing shared data estates rather than the current data silos built as part of individual city services. It is important for cities to understand the security and

resilience advantages of a shared data estate rather than a single data repository for the data framework.

The benefits of this approach can break down any barriers between data providers and prevent the creation of barriers in the future. Consideration should be given to changes which take place in organizations over time and a mitigation strategy developed to include contingency measures for the continuity of service, including key personnel. The data framework can allow cities to explore duplications or enhancements for existing services and discover new opportunities to provide new services based on insight from the data. A single view of the city data can highlight improvements and efficiencies and help cities understand how best to improve service delivery.

5.3 Knowledge creation approach

If cities merely transition the technology they use to create a data sharing culture, this can enable some of the value to be derived from the data framework and some data capacity to be created with this incremental change. Data sharing in smart cities has the potential to create not just incremental change, but a step change for the city that creates not just data capacity but data capability: this requires a knowledge creation approach.

A knowledge creation approach is focused on ensuring that data is federated beyond data specialists, and is in the hands of the majority of the domain specialists, responsible for management and oversight of city services. Domain specialists and data users need a method of discovering and accessing the data within the data framework on demand, as part of their business as usual process. The provision of a feedback loop should be developed to monitor and prevent errors in the data, creating a dynamic data resource for all. Any updates or revisions to the data within the data framework should be transparent in order to ensure that the provenance of the data is maintained. A data framework which is designed and implemented in this way can reduce duplication and provide a better user experience for all within the city. Additionally this can allow cities to benchmark themselves against other cities, and understand similarities and differences on a regional or national basis.

5.4 Promoting trust and participation

In addition to the need for a city to articulate the benefits of data sharing, it is important to determine how to build a trust model that can be shared with citizens and used to develop a citizen participation model for citizen data that is shared by the city. A digital ethics code should be developed in consultation

with citizens to provide publicly available guidance related to those citizens who will be providing their data to the data framework. One approach would be the development of a shared data charter explaining why and how data will be used for each of the top level access restrictions that apply to shared data, in order to promote trust and participation across the entire data spectrum

Case study 6 in Annex A gives three examples of existing open data charters in cities, illustrating the basis for the creation of shared data charters. A shared data charter should identify the security measures to be used by data owners to underpin the trust model for the city.

5.5 Anonymization of data

5.5.1 General

As anonymization is a key technique for smart city data sharing, an overview of the guidance needed has been included in this standard (see **5.5.2**).

The term “anonymization of data” refers to data that does not itself identify any individual and that is unlikely to allow any individual to be identified through its combination with other data.

Anonymization of asset data should also be considered by cities. Examples of asset-related data sets where anonymization would be required include those containing:

a) information on alarms connected to response centres with automatic police response;

b) some COMAH data; and

c) energy consumption at building or occupier level.

Different jurisdictions have specific approaches to the anonymization of personally identifiable data, local laws and registrations will inform the approach which needs to be taken. Although it is not always possible for anonymized data to be completely risk-free, rather it is possible to mitigate the risk of re-identification until it is remote. If the risk of re-identification is likely, the information should be regarded as personal data. The term “re-identification” describes the process of turning anonymized data back into personal data through the use of data-matching or similar techniques.

5.5.2 Anonymization – key points

The key aspects of anonymization of data in cities are:

a) data protection law does not apply to data rendered anonymous in such a way that the data subject is no longer identifiable;

b) fewer legal restrictions apply to anonymized data;

c) where the anonymization of personal data is possible, it can help service society’s information needs in a privacy-friendly way;

d) can help all organizations that need to anonymize personal data, for whatever purpose; and

e) can help cities to identify the issues that need to be considered to ensure the anonymization of personal data is effective.

5.5.3 Governance and management of anonymization

It can be easier to seek to use anonymized data in new and different ways because the local laws do not apply, for example the UK data protection laws on purpose limitation rules do not apply. Anonymization allows organizations to make information public while still complying with their data protection obligations. The disclosure of anonymized data is not a disclosure of personal data - even where the data controller holds the key to allow re-identification to take place.

The city should work with organizations processing city data to put in place an effective governance structure in relation to their anonymization processes. This can support dealing with complaints to regulatory authorities about the processing of personal data including its anonymization, or if a regulator carries out an audit. A thorough risk analysis should be carried out on the likelihood and potential consequences of re-identification at the initial stage of producing and disclosing anonymized data.

The risk of re-identification differs according to the way the anonymized information is disclosed, shared or published. Publication to the world at large is more risky than limited access; limited access allows the disclosure of richer data but relies on robust governance and contractual arrangements. In cases where the consequences of re-identification of anonymized data could be significant - for example, because it would leave an individual open to damage, distress or financial loss - the city should work with organizations processing city data to:

a) seek the data subject's consent for the disclosure of the data, explaining its possible consequences;

b) adopt a more rigorous form of risk analysis and anonymization;

c) in some scenarios, such as those where the consequences of the risk of re-identification could be significant, the data should only be disclosed within a properly constituted closed community, and with specific safeguards in place;

d) provide information on personal data collected and the purpose for which it is used; or

e) ensure that spatial data related to objects cannot be used to reveal the identity or patterns of life of individuals or groups of individuals.

5.5.4 Mitigating re-identification risks

Re-identification risks can occur across the lifecycle of de-identified data in a smart city. These risks can occur as a result of any organization which who is involved in the creation of de-identified shared or closed data in the smart city. ISO/IEC 27559 provides a technology agnostic framework for identifying and mitigating risks which can be used to support compliance and regulatory requirements and relevant privacy principles.

5.6 Spatial data

Spatial data is any data with a direct or indirect reference to a specific location or geographical area. Spatial information includes postcodes, GPS data, trajectories (the path of a moving object or person) and map references, and these sometimes constitute personal data. The approach cities should take to handling spatial information is dependent on available related information, and the size of the dataset they are dealing with. To avoid disclosure of personal data, and to reduce re-identification risk, for some types of spatial information, the removal or blurring of certain elements should be considered, for example, using partial postcodes rather than full postcodes. Small numbers in small geographical areas present increased risk but this does not mean that it is necessary to remove small numbers automatically. For example, removing numbers relating to five or ten individuals or fewer might be a reasonable approach for minimizing risk of identification in a proactive disclosure scenario. People's trajectories are particularly sensitive and will require special treatment.

5.7 Using a shared language for data governance

In a city it is important that the language used for the sharing of data is common across the city and city organizations in order to ensure that appropriate governance is in place. Ensuring appropriate governance requires leadership and comprehensive coordination from the city. This governance should include consideration of any issues around data quality, to enable the correct actions to be taken when sharing data with the appropriate access across the data spectrum.

In some jurisdictions regulation will require that the city has the appropriate opt-out process and an auditable trail of evidence in place to prove that information has been removed once a request for removal has been processed.

5.8 Data sharing in practice

Although different jurisdictions have different regulatory regimes smart cities globally operate a multi agency model and in practice there are two key scenarios for the sharing of personal data: systematic data sharing arrangements as well as ad hoc or one-off requests (see 5.8.1).

"Data sharing"in a smart city setting means "the disclosure of data from one or more organizations to a third party organization or organizations, or the sharing of data between different parts of an organization".

Data sharing can take the form of:

- a reciprocal exchange of data;
- one or more organizations providing data to a third party or parties;
- several organizations pooling information and making it available to each other;
- several organizations pooling information and making it available to a third party or parties;
- exceptional, one-off disclosures of data in unexpected or emergency situations; or
- different parts of the same organization making data available to each other.

5.8.1 Two types of data sharing

Some data sharing doesn't involve personal data; for example where only statistics that cannot identify anyone are being shared. In smart cities there are two key types of data sharing:

a) systematic, routine data sharing where the same data sets are shared between the same organizations for an established purpose; and

b) exceptional, one-off decisions to share data for any of a range of purposes.

Different approaches apply to these two types of data sharing. Some of the good practice recommendations that are relevant to systematic data sharing are not applicable to one-off decisions about sharing. Systematic data sharing generally involves routine sharing of data sets between organizations for an agreed purpose. It could also involve a group of organizations making an arrangement to pool their data for specific purposes.

Adopting the good data sharing can help organizations to collect and share personal data in a way that complies with regulation, is fair, transparent, and in line with the rights and expectations of the people whose data is being shared.

Much data sharing in cities currently takes place in a pre-planned and routine way. As such, it is governed by established rules and procedures. However, as cities realise the value from the data framework, they might decide, or be asked, to share data in situations that are not covered by any routine agreement. In some cases, this might involve a decision about sharing being made in conditions of real urgency, for example in an emergency situation.

6 Data protection reform

6.1 General

At the time of writing, new data protection legislation has or is being introduced across the world. In many cases terminology differs in each jurisdiction. It is important to establish the terms which should be used in the specific jurisdiction which apply to the smart city to ensure that the implementation of this standard has been adapted to the specific data protection setting.

This standard provides some overarching guidance (see 6.2 to 6.10) based on best practice advice recommendations to prepare for data protection compliance which is applicable to the smart city implementing this standard. This will not contain all the guidance necessary in the applicable data sharing jurisdiction. Further reading on regulatory setting is listed in the Bibliography.

Emerging data protection regulatory regimes are a window of opportunity for cities to create the data framework for the open and shared data across the data spectrum. The data framework is an ideal governance mechanism for building data protection compliance. Adoption of this standard to initiate data sharing for a smart city also helps support the business case for the transition of existing technologies that are not considered adequate to ensure regulatory data sharing compliance.

6.2 Awareness

It is essential to start planning an approach to compliance as early as possible, and to gain buy-in from key people in the organization. It might be necessary, for example, to put new procedures in place to deal specifically transparency and individuals' rights provisions. In a large or complex business, this could have significant budgetary, IT, personnel, governance and communications implications. Implementation of regulatory compliance requires detailed planning for all the organizations which contribute and consume data from the data framework. It is recommended that should be led by the city to ensure all affected organizations are included.

6.3 Information that is held

It is a requirement of many regulatory jurisdictions to assess and document what personal data is held, where it came from and who it is shared with. It might be necessary to organize an information audit, across the organizations within the city, or for particular city services. This is particularly important for a smart city to recognize privacy rights for a networked world. For example, if inaccurate personal data is held and shared with another organization, the other organization is required to be informed about the inaccuracy so it can correct its own records. This is not possible without knowing what personal data is held, where it came from and who it is shared with. This information should be documented. Doing this could also help organizations to comply with the accountability requirements of the regulatory setting which if applicable requires organizations to be able to show how they comply with the data protection principles, for example by having effective policies and procedures in place.Communicating privacy information

Currently, when personal data is collected, there is a requirement to give people certain information, such as the identity of the organization collecting the information and their intended use of the information. This is usually done through a privacy notice. In some jurisdication there is additional information that is required to be shared, for example the data retention periods. Individuals have a right to complain to a regulator if they think there is a problem with the way their data is being handled. It is best practice that the information to be provided is in concise, easy to understand and clear language.

6.4 Subject access requests

The rules for dealing with subject access requests vary across jurisdictions in many cases it will not be possible to charge for complying with a request and there is usually a time limit when a reply must be sent. In some jurisdictions they may also be different grounds for refusing to comply with subject access request, for example manifestly unfounded or excessive requests may be able to be charged for or refused. If refusing a request, policies and procedures will be required to be in place to demonstrate why the request meets these criteria.

Organizations are often required to provide some additional information to people making requests, such as data retention periods and the right to have inaccurate data corrected. If the organization handles a large number of access requests, the impact of the changes could be considerable, so the logistical implications of having to deal with

requests more quickly and provide additional information will need thought through carefully.

A new operating model based on data sharing for a city could ultimately save a city and city organizations a great deal of administrative cost if they can provide people with the ability to access their information easily online. It is recommended that organizations consider conducting a cost/benefit analysis of providing online access.

6.5 Consent

The city should work with organizations processing city data to review how they are seeking, obtaining and recording consent and whether they need to make any changes. Usually consent has to be freely given, specific, informed and unambiguous.

Consent is usually a positive indication of agreement to personal data being processed it cannot be inferred from silence, pre-ticked boxes or inactivity. If an organization relies on individuals' consent to process their data, they will be required to make sure it meets the regulatory compliance in their jurisdication. If not, the organization will be required to alter its consent mechanisms or find an alternative to consent.

Consent is verifiable and individuals generally have stronger rights where consent is relied upon to process their data. It is recommended that the mechanisms used for recording consent are reviewed to ensure an effective audit trail is created.

6.6 Privacy regulation related to children's data

In some jurisdictions regulation has been updated relating to the sharing of data from children. Smart cities should consider whether it is an optimal time to consider putting systems in place to verify individuals' ages and to gather parental or guardian consent for the processing of children's data as they adopt the data framework.

This processing of children's data could have significant implications for children's services in a city where their personal data is collected as part of the service provision. When collecting and sharing children's data it is recommended that any documentation is in language that children will understand.**Data breaches**

Some organizations are already required by regulators to notify the regulatory authority (and possibly some other bodies) when they suffer a personal data breach. This will be new to many organizations. Not all breaches are usually required to be notified to the regulatory authorities, only ones where the individual is likely to suffer some form of damage, such as through identity theft or a confidentiality breach.

In some cases, an organization will be required to notify the individuals whose data has been subject to the breach directly, for example where the breach might leave them open to financial loss. Smart cities are advised to develop policies and procedures for managing data breaches, whether at a central or local level. A failure to report a breach when required to do so could result in a fine, as well as a fine for the breach itself.

In conclusion, it is recommended that a smart city and the other organizations sharing data from the data framework makes sure they have the right procedures in place to detect, report and investigate a personal data breach.

7 Data value chain

7.1 General

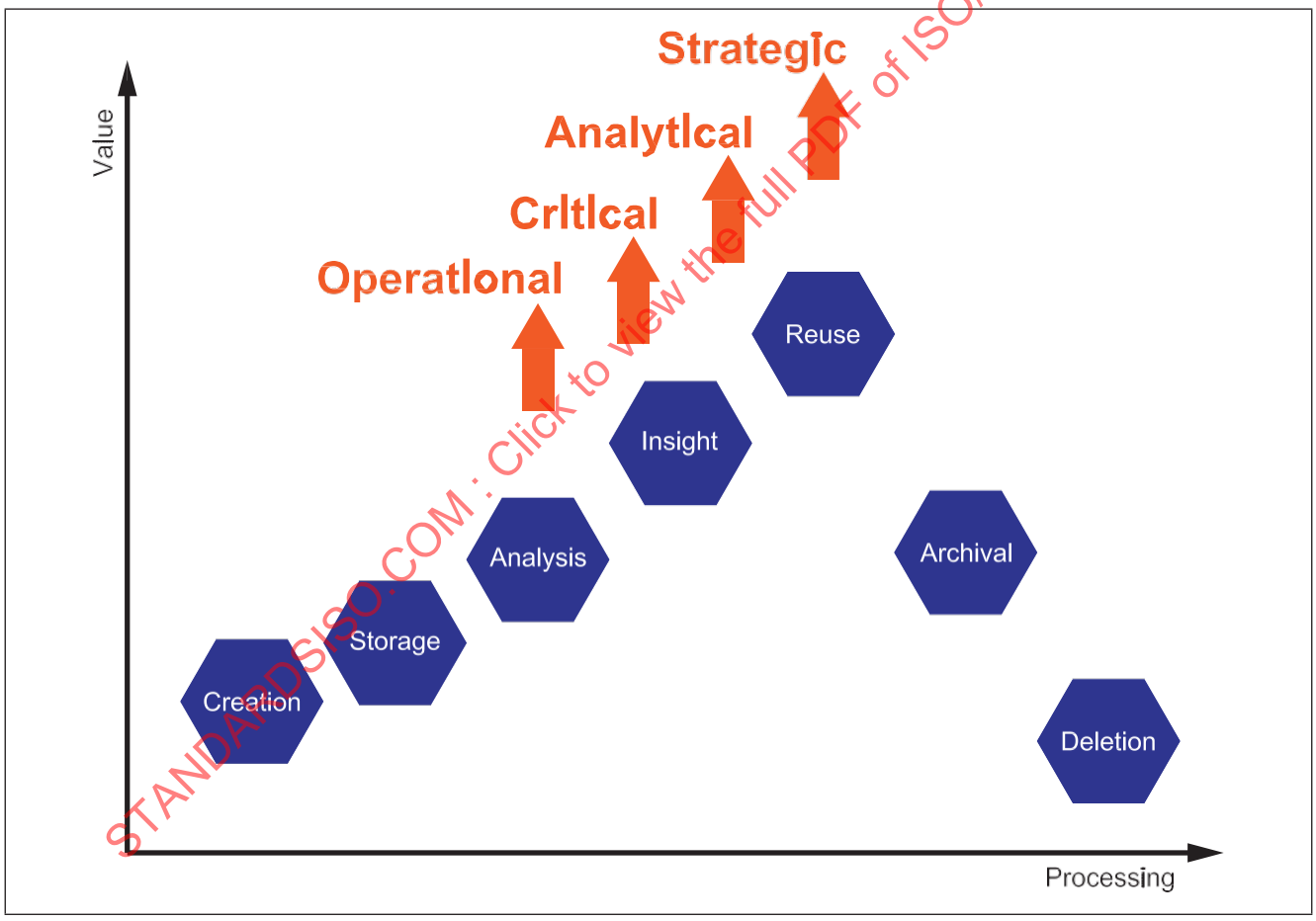
The data value chain is a mechanism through which cities can view the flow of data from the point that it is collected throughout its entire lifecycle. Each city has its own flow of data and needs from data, but the four key levels of insight within a city are

created by analytical processes which turn the data framework into different kinds of actionable intelligence.

The city should ensure that any risk of an adverse impact on an individual, organization or asset is considered prior to the sharing of the data from the data value chain. The data value chain depends on a blended technology ecosystem in a city which acts as disruptive force to understand traditional, static practices and supplant them with innovative, purpose-built solutions based on data analytics. Essentially this transforms data into an additional city asset.

The data value chain as shown in Figure 3 has the power to disrupt cities with new ways of thinking and doing. The data value chain also has the ability to unify disparate city services by putting data – knowledge – into the hands of decision- makers – not just data specialists. A data value chain should be used across each department or organization in the city, informing daily workflow. This approach can support decision-making with data from the simple deployment of resources, the strategic placement of those resources, to the ultimate value those resources provide in return.

Figure 3 **Data value chain**



7.2 Data roles

7.2.1 General

Although individual cities have their own data value chain, there are five key roles to be fulfilled in order to maximize the impact of the data framework in a city.

The roles that exist across the data value chain include:

a) data creator (see **7.2.2**);

b) data owner (see **7.2.3**);

c) data custodian (see **7.2.4**);

d) primary publisher (see **7.2.5**);

e) secondary publisher (see **7.2.6**);

f) users (see **7.2.7**); and

g) other cities (see **7.2.8**).

7.2.2 Data creator

The data creator role defines those organizations who both collect and/or transform data for the city or its services. This role can be passive where the organization is responsible for the creation of data for a city, as part of the provision of a city service, for example the creation of the city data relating to the location of lampposts in the city. Additionally, this role can be a reactive role where operational insight data is collected and is transformed to provide the city with critical insight, for example a transport operator in a city who supplies data collected from cameras in the event of an critical incident. For derived or aggregated data, the data creator is the provider of the process which transforms the data created by others.

7.2.3 Data owner

The data owner is the designated curator for the data related to a city service on behalf of the city. The responsibilities of this role include the authority to change the data where appropriate and maintain the transparency for the provenance of the data within the data framework on behalf of the city.

7.2.4 Data custodian

The data custodian role differs to the data owner role as this organization does not own the data, it merely is the custodian of the data for a specific purpose or task related to the provision of a service within the city. See case study 7 in Annex A for an example of the difference in the roles of data owner and data custodian in city service provision.

7.2.5 Primary publisher

The primary publisher role relates to the organization that performs the publication role for all data across the data spectrum. All sources of data can be viewed by the organization who performs this publisher role, all data however might not be published. Publication of the data depends on which part of the data spectrum the data belongs to and the access restrictions which apply.

7.2.6 Secondary publisher

In a smart city an additional publication role exists. The publication of some of the data on the data spectrum is facilitated by the primary publisher. As a result, for some of the published data an organization creates additional value from the city data which has been published. This secondary organization should be encouraged to publish the new value data which has been created, performing the role of secondary publisher. The secondary publisher should monitor the quality of the data in the data framework, feeding back to the city on any variance detected as part of the data publication process.

Any access restrictions to the data to be published as part of this secondary publication role are determined by the primary publisher. A feedback loop should be incorporated which supports the primary publisher delegating authority to the secondary publisher to oversee the publication of the data itself.

7.2.7 Users

There are a number of organizations who can have differing roles in the data value chain but are also considered to be the users of city data. Although this varies between cities, the key user groups that are common to all cities are:

a) city organizations who support the operation of city services, for example emergency services, community health services and contractors;

b) third sector organizations providing or supporting city services;

c) business users, for example corporates and SMEs;

d) citizens; and

e) academic organizations.

Other aspects of the data value chain that exist in a city but which are outside the scope of this standard are:

f) other cities (see **7.2.8**); and

g) citizen roles (see **7.2.9**).

7.2.8 Other cities

This other cities role relates to cities where data sharing agreements are in place, particularly where the other city is considered a near neighbour, which might or might not be geographically co-located, but where city services are shared between the cities. Although outside the scope of this standard, any impact of other cities on the data value chain should be considered in these situations.

7.2.9 Citizen roles and responsibilities to other citizens

Although outside the scope of this standard, the data value chain might be impacted by citizen roles and therefore they should be considered.

It is important that city organizations understand these roles in order to fully utilize the data framework, in the context of the statutory duties that exist for the protection of personal data when sharing data across the city.

7.3 Data sharing responsibilities

7.3.1 General

There are three key roles which smart cities should use for the sharing of personal data:

a) data controller (see **7.3.3**);

b) data processor (see **7.3.4**); and

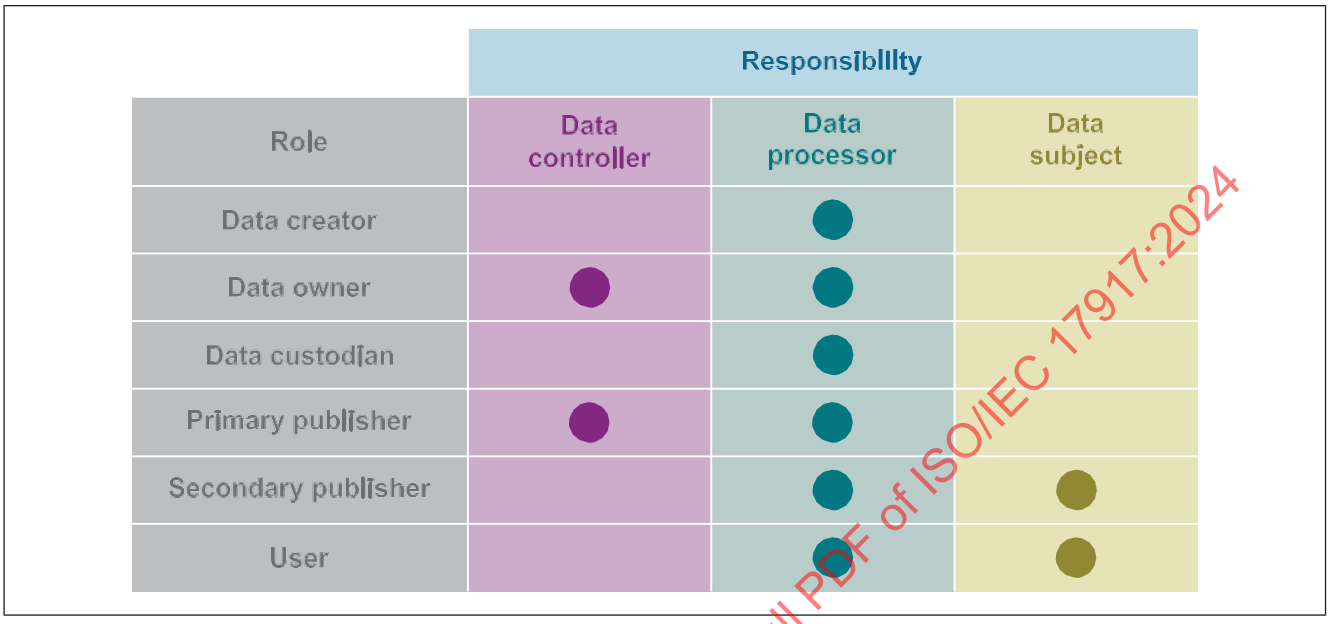
c) data subject (see **7.3.5**).

7.3.2 Data roles and responsibility matrix

It is important that a city has a clear view of these roles and responsibilities before new business models are explored around the commercialization of data.

As a result of the legislation and regulation in the specific jurisdiction which surrounds the sharing of personal data, the roles and responsibility matrix shown in Figure 4 should form the basis for testing whether business models can be considered appropriate.

Figure 4 **Roles and responsibilities matrix**



7.3.3 Data controller

In relation to personal data, a data controller is a person who (either alone or jointly or in common with other persons) determines the purposes for which and the manner in which any personal data are, or are to be, processed.

7.3.4 Data processor

In relation to personal data, a data processor is any person (other than an employee of the data controller) who processes the data on behalf of the data controller.

7.3.5 Data subject

A data subject is an individual who is the subject of personal data.

7.4 Data controller vs data processor

It is important for cities involved in the processing of personal data to be able to determine whether they are acting as a data controller or as a data processor in respect of the processing. This is particularly important in situations such as a data breach where it will be necessary to determine which organization has data protection responsibility.

There is a distinction between a data controller and a data processor in order to recognise that not all organizations involved in the processing of personal data in smart cities have the same degree of responsibility. Irrespective of the regulatory jurisdiction it is usually the data controller that exercises control over the processing and carries data protection responsibility for it.

8 Purposes for data use

8.1 General

Based on the data that is in the data framework and the data spectrum that has been agreed upon, it is important for a city to understand the purpose of the data sharing that is intended. Equally important is the need to consult stakeholders on these data sharing intentions, utilizing the learning from the open data agenda and engaging with stakeholders or stakeholder groups both offline and online. The maturity of a city's approach to data sharing can support the development of an initial engagement with stakeholders to empowering communities with data, through to active participation in meeting the challenges and opportunities for the city.

In order for smart cities to evolve towards becoming sustainable and resilient cities, it is important to use the data framework as an evidence base not just for understanding the value that data can bring to current city services, but for exploring future innovations based on data, while remaining mindful of the need to protect the safety and security of individuals, services, sensitive assets and systems, and the benefits that city assets and services exist to deliver.

The following key areas (see 8.2 to 8.5) should be considered as the basis for understanding the purpose of data in a smart city, which can be used as a basis for continuous consultation with stakeholders.

8.2 Primary purpose

Smart city data sharing is predominantly done across the data spectrum for the benefit of citizens. A city initially looks to engage stakeholders using data to communicate about city services, challenges and opportunities. This allows an understanding to be developed on the real needs that the city is addressing and how this might be focused and refined. Ultimately a citizen and stakeholder participation approach to data sharing in a city allows the move towards a self-service model for data that is shared.

The efficient functioning of a city and key services such as the transport infrastructure and effectiveness of city utilities are also key reasons to share data in a city.

As a city matures its approach towards data sharing this can facilitate the growth agenda for local businesses or sectors, especially as commercial models are developed for data sharing. Throughout this data journey it is important to develop this approach to data alongside the political agenda and around the use and need for data: an evidence-led approach to change.

8.3 Secondary purposes

Alongside the primary purposes for data sharing, a number of secondary purposes accrue. When the data framework is used to explore these new purposes, the most stringent requirements for the data used should apply.

Data sharing can be used to establish a future innovation approach, as the evidence base created from data sharing can be extended to include new initiatives and opportunities.

Data sharing can also highlight where opportunities for new business streams exist and where sector growth might be encouraged. A data sharing strategy can become the basis of a business case to seek new funding. There is a need to understand sustainable funding models for smart cities which enable smarter integration of city services. These new funding models should be delivered via partnership style agreements. Data sharing is a key innovation which can enable a mixed economy funding model to facilitate the integration of existing services with new delivery partners across

different sectors. For example, an initiative which addresses the growth ambitions for a sector or community could involve the city, private and third sector organizations who share data to drive the innovation.

8.4 Reuse and distribution

8.4.1 General

When data sharing from the data framework relates to the sharing of asset or service data this should take account of any sensitivities related to the data which is being shared.

When assessing the reuse and distribution of data from the data framework, it is important to ensure that this is adequately understood for personal data in the two different data sharing scenarios that are usually present for smart cities.**Systematic data sharing scenario**

When entering into an agreement to share personal data on an ongoing basis the following data sharing checklist - systematic data sharing should be considered.

Is the sharing justified?

Key points to consider include the following.

a) What is the sharing meant to achieve?

b) Have you assessed the potential benefits and risks to individuals and/or society of sharing or not sharing?

c) Is the sharing appropriate and proportionate to the issue you are addressing?

d) Could the objective be achieved without sharing personal data?

Do you have the power to share?

Key points to consider include the following.

a) The type of organization you work for.

b) Any relevant functions or powers of your organization.

c) The nature of the information you have been asked to share (for example was it given in confidence?).

d) Any legal obligation to share information (for example a statutory requirement or a court order).

If deciding to share, it is good practice to have a data sharing agreement in place. A smart city data sharing agreement should cover:

a) what information needs to be shared;

b) the organizations involved;

c) what to communicate to people about the data sharing and how to communicate that information;

d) measures to ensure adequate security is in place to protect the data;

e) arrangements for providing individuals with access to their personal data if they request it;

f) agreed common retention periods for the data;

g) processes to ensure secure deletion takes place.

8.4.2 One-off requests scenario

In certain circumstances, a city might be asked to share personal data relating to an individual in one-off circumstances and the following Data sharing checklist – one off requests, should be considered.

Is the sharing justified?

Key points to consider include the following.

a) Do you think you should share the information?

b) Have you assessed the potential benefits and risks to individuals and/or society of sharing or not sharing?

c) Do you have concerns that an individual is at risk of serious harm?

d) Do you need to consider an exemption in regulatory laws?

Do you have the power to share?

Key points to consider include the following.

a) The type of organization you work for.

b) Any relevant functions or powers of your organization.

c) The nature of the information you have been asked to share (for example was it given in confidence?).

d) Any legal obligation to share information (for example a statutory requirement or a court order).

If you decide to share

Key points to consider include the following.

a) What information do you need to share?

   1) Only share what is necessary.

   2) Distinguish fact from opinion.

b) How should the information be shared?

   1) Information should be shared securely.

   2) Ensure you are giving information to the right person.

c) Consider whether it is appropriate/safe to inform the individual that you have shared their information.

Record your decision

Record your data-sharing decision and your reasoning, and whether or not you shared the information.

If you share information you should record:

a) what information was shared and for what purpose;

b) who it was shared with;

c) when it was shared;

d) your justification for sharing;

e) whether the information was shared with or without consent.

8.5 Access and control

When establishing the purpose of data sharing in a city and the basis on which it is to be used, it is important to decide how data is to be accessed and by whom. The access requirements should be based on the purpose for which the data has been shared.

It is particularly important that security, access and control restrictions are considered if the data being shared has been repurposed from the original reason that the data was collected. When the data framework is used to explore these new purposes, the most stringent requirements for the data used should apply.Although the city might have defined these principles for open data this should be re-examined for shared data, ensuring that the roles and responsibilities for shared data are equally well understood. For example, footfall data collected by a city to understand priorities related to retail development would need different security access and control mechanisms in place if the data was collected manually by survey, than if video camera data was used. Video camera data would need to anonymized before this data could be shared.

Depending on the data collection and processing services for shared data, the format of this data is likely to vary depending on the city services that collect and use this data. Provided the data to be shared is in machine-readable format – as a minimum in comma separated values (csv) format – it can be repurposed as part of a data sharing initiative.

8.6 Technology considerations

Smart Cities operate a multi-agency model which may be complicated by the technology deployed and the considerations needed for specific technologies. Usually there are contractual arrangements in place which specify the permitted processing of shared data by a technology vendor or operator as to assume the governance role of a data processor of smart city data. Another example of the technology considerations needed is the use of cloud data storage which uses security boundaries to customize the access, security and control privileges for Cloud Service Customers. ISO/IEC 23751:2022 gives details of appropriate data sharing agreements where data is intended to be processed using one or more cloud services, or other distributed platforms.

8.7 Other considerations

Although the agenda around data sharing in smart cities is only just being explored, it can enable a city to understand how a partnership model can be developed between the city and the service providers who deliver the city services. The development of this partnership model is one which is more mature in other domains and it allows cities to learn from similar use cases from other sectors. Case study 8 in Appendix A gives an example of how a partnership model can have benefits for the citizen.

9 Assessing data states

9.1 General

To fully embrace the commercial opportunities that are available to smart cities based on the sharing of data new business models will need to be developed to reflect data sharing strategies. The data framework should provide an understanding of the data states for the data it holds.

Traditionally, there has been a publication-averse culture that exists as a result of many factors such as: a lack of data skills or knowledge, unwillingness to expose issues related to data quality or completeness, lack of knowledge or understanding, or the legislation or regulation. However, such aspects need not restrict the publication or

reuse of this data, providing the barriers are acknowledged in the data framework. One way of removing these publication barriers is to acknowledge the data states of the data in a city's data framework. The following key areas (see 9.2 to 9.6) should be included in the data framework to provide support to the development of new commercial models for shared data in a city.

9.2 Completeness

Understanding any level of completeness in the context of the purpose for which the data is intended to be used can allow the suitability of the data for that purpose to be understood by all. The data owner is responsible for ensuring that a measure of completeness (i.e. the data coverage and its continuity) is included in the data framework.

It is not always possible for a city to consider the data being shared with a measure of completeness, therefore as a minimum a completeness measure in the extent of the data's coverage and its continuity should be created by the data owner for the shared data in the data framework.

9.3 Accuracy

In addition to a measure of completeness, the data framework should contain a measure of accuracy for the data that is to be shared. The data owner should consider this in terms of the reliability of the data in the context of a specification, or for a particular data standard. This should be a precision measure, a valid range of measurements, tolerances or sampling rates which are declared as part of the data framework. The data owner should also record any conversion processes that are required, for example to convert analogue data to digital data.

Whilst the measure of accuracy of the data being shared is determined by the data owner, ultimately the data user evaluates and determines the accuracy of the data which has been shared. The city should embed a change control mechanism as part of a feedback loop for the data framework to allow the accuracy measure and the associated metadata to be updated based on feedback from the data user.

9.4 Availability

Availability encompasses the publication of data across the data spectrum and the availability of the technology which is used to implement the data framework.

The availability of data – including closed data – across the data spectrum should be declared as part of the data framework. The details related to closed data should be confined to an acknowledgement that it is part of the data framework and is not available as part of any data sharing arrangements. Despite this, it is important that the closed data that a city holds is included in the data framework. This can enable a trust and participation model to be developed with citizens and other city organizations.

The success of commercial opportunities for data sharing in smart cities relies not just on the availability of data in the data framework, but on the reliability of the refresh timetable for this data. It is important that the city creates a regular refresh timetable which data users can rely on. The refresh timetable should be in enough detail to clarify the refresh periods for all data in the data framework. It is particularly important for data which relies on updates from a number of organizations in the city that refresh timetables include any periods when data is not available as part of the data refresh cycle.

ODUG evidenced nine barriers to the usage and adoption of open data one of which was the reliability of the data refresh cycle (Data users need to be confident that the refresh cycle is adhered to by the city, preferring a longer reliable refresh cycle for data available from the data framework, rather than more frequent but less reliable data refreshes.

The availability of data from the data framework is governed by the security, access and control restrictions which apply to all data across the data spectrum. The city should use the restrictions which apply as a focus for the development of commercial models. Although APIs are outside the scope of this standard (see 11.5 and Annex A) the successful development of commercial models is likely to depend on the provision of these interfaces in order to develop the four levels of insight for a city and make this data accessible via the data framework.

9.5 Timeliness

A measure which allows a data user to determine the temporal nature of the data they wish to use, should include a measure of timeliness. Timeliness is the time lag between data being updated in the real world and an update occurring in the data framework. For example, some data such as that from sensors as part of the city transport infrastructure might occur in real-time, whereas reference data related to school provision might only be updated annually. This measure is a key part of the data framework which can facilitate all four levels of insight in a city.

It is also important to consider the frequency with which a measure is updated. This is particularly important where a measure can change rapidly or can be subject to periodicity issues.

In addition to including this measure as part of the data framework, the city should consider whether this also needs to be formalized as part of the data sharing agreements if it provides a part of the commercialization of their data framework.

9.6 Provenance

For all the data states a city includes as part of its data framework, provenance is a fundamental element of the ability of a city to monetize its data. The city should use the data value chain as a key mechanism to understand the provenance of the data it holds. An understanding of the origin or history of the data across the data value chain can directly determine the commercial value of data in a city. For sensor data, specific reference data, including location and measurement assurance data (used to determine whether the sensor is operating within tolerances) should be published as part of the data framework. A data user relies on the provenance of the data they use, reuse and publish to determine the value of a city's data.

10 Defining access rights for data

10.1 General

This standard considers the access restrictions that apply to all data across the data spectrum from legislative, regulatory and standards viewpoints. In stressing the importance of these considerations, it is also important to consider the governance structure that is needed to support access rights to city data.

The utilization of the data framework to support the four levels of insight in a city creates unique aggregations of data relevant to the challenges and opportunities experienced by the city. This should also be considered when

defining access rights to any of the data that the city holds as part of its data framework. The city should consider the requirements of users, stakeholders, organizations, machines and devices.

10.2 Individual requests for data

Many cities already support individual requests for data under existing freedom of information legislation, however it is recommended that this support should be extended to include the data within the data framework. These individual requests for data are expected to be based on thematic data and the custodianship of these requests might need to be federated across the city, based on the operating model that is in place.

The provision of data to support these individual requests should be based on a feedback system to allow the individual requester or organization to question the use of this data and report any inconsistencies or errors they find.

In order to reduce the burden of managing these requests for data, there should be an automated mechanism in place for this purpose. This mechanism should be able to be used by both citizens and organizations. The mechanism should also provide an audit trail of the request that can be used to review decisions taken. Additionally this mechanism should be the basis for any further regulatory or legislative investigations that might be required; these investigations can then be based on the interactions that have taken place as a result of the request for data.

10.3 Use of identity management methods

10.3.1 General

The use of identity management methods is an emerging topic in the adoption of services used by citizens. Smart cities will need to decide which method best supports citizens to share data and their access to city services. In sectors such as utilities, for essential services such as gas, electricity, water, where data is not subject to legal obligations access is controlled internally by the individual companies in a manner that best supports their own processes. Even as these services move online, the operation of these services is likely to be based around the needs of the organization and not the citizen. However, smart cities should take a different view, with the needs of the citizen at the centre of the transformation. Smart cities take a different view of the interaction between the citizen and the city, with the needs of the citizen at the centre of the transformation to be a smart city. In this scenario, a single credential allows access to the central data resource and the use of a choice of identity management access allows the citizen to choose the identity management service to use.

In the majority of government scenarios, the identity management approach mirrors the organization-centric approach of the private sector, with some very complex data sharing arrangements in place. These data sharing arrangements and gateways have evolved over time and have increased in both number and complexity. For those citizens and organizations that are the subject of the data being shared, there is no direct access to these services by individuals. This will need to change for smart cities in order for them to deliver citizen-centric services.

10.4 Authorization and authentication of data

10.4.1 General

Authorization and authentication of data is a complex area and there are a number of factors to consider, some which apply across the whole of the data framework, and some which are specific to the access restrictions related to shared data. Additionally, it is important to consider any IP, security or commercial sensitivities which are related not just to personal data but device and asset data.

It is important to consider not just the legislative and regulatory requirements that relate to shared data but also to consider the access rights which apply.

10.4.2 Authorization of data

In some cases the authorization of data for a city service will need to make use of a gateway to gain permission to share the data. Where these gateways are used there will be restrictions on how and with whom the data can be subsequently shared. For example the use of the Driver Vehicle Licensing Authority (DVLA) authorization service in order to share personal data related to UK Citizens, vehicles and ownership.

10.4.3 Authentication of data

Authentication services are specific to a domain, would predominantly be used to authenticate specific city services, and use thematic data from the data

framework. In many cases for government-owned data, this involves adhering to the existing data sharing legislation and regulation, often through legal gateways. Case study 9 in Annex A gives an example of the use of authentication and authorization of data to deliver a city service.

These authentication and authorization services determine the approval or denial of the request, based not just on the attribute exchange required but on the intended purpose for the request. In some cases there might be levels of restrictions related to the data, for example it might be possible to request to share data from the index of a register but not the remainder of the data held within the data framework

From a citizen viewpoint, the sharing of data might need to be facilitated with an authorized agent, for example a legal guardian or someone who holds power of attorney on behalf of another citizen. This should be accommodated within the authorization and authentication data sharing mechanisms available within the city.

10.5 Audit of access requests

The city should have in place an accessible audit facility for all authorization and authentication requests for its city services as part of the data framework. There may be a need for this audit facility to be available to all organizations providing the city services to meet legislative and regulatory obligations. It should also facilitate analysis of requests, to better understand the use of the data framework for the sharing of city data.

It is important to understand the reasons for request, frequency, by whom and the reasons for multiple requests for the same or similar data. This analysis insight on access requests is unlikely to need the provision of additional capability – it is an extension of the analytics provision for the four key levels of insight. This approach to the analysis of the access to city data can help ensure that this facility can be provided at minimum extra cost and also minimize the burden of the provision on the city.

10.6 Commercial use of access data

The access data related to a city's data framework can be used by the city to understand

the development of the data ecosystem and highlight specific challenges and opportunities for the development of a city's commercial data offering. It can also yield secondary benefits allowing the city to understand patterns and profiles of usage for the data framework, for example the most used thematic data, or the most infrequently used reference data. This could allow a city to develop a commercial license offering around the data framework based on, for example, demand patterns or frequency.

Much of the place-based information in a smart city relates to the environment, for example large parts of planning, transport, environmental health and conservation data. Reuse of environmental data might also be limited by third party data and intellectual property rights.

There is however a balance to be struck, at the discretion of the city, between the sharing of data and information services for the public good and the commercial use of data. This is especially true where citizens are required to provide data

in order to receive a city service or meet legal requirements and the effect the commercial use of data could have on public trust.

11 Data structure

11.1 General

There are a number of considerations related to the structure of the data to be shared from the data framework. While the format of the data to be shared is dependent on the technologies used to support the data framework, there are a number of data structures that a data framework uses.

11.2 Data structures

11.2.1 General

A city uses both encrypted and unencrypted formats across the data framework and these formats might need to be accessed and/or rendered in different ways. Although most cities are familiar with the needs of structured data, this is in reality the minority of the data in the data framework for smart cities. Semi-structured and unstructured data forms the majority of the data in the data framework, especially as the city embraces the internet of things.

11.2.2 Structured data

Structured data refers to any data that resides in fixed fields within a record or a file. Structured data typically resides in relational database management systems (RDBMs) and spreadsheets. Within the data framework, the structure of the data should be declared. This can help ensure that all users of the data are able to understand how to maximize usage of not just the thematic data, but also the reference data and metadata. Declaring the format of the data held can enable the city and city organizations to make best use of all the data in the data framework.

11.2.3 Semi-structured data

Semi-structured data does not conform to the formal structure associated with traditional RDBMs but nonetheless as the name suggests contains tags or other markers which allow hierarchies of records and fields to be understood. Semi-structured data is often referred to as self-describing data.

Semi-structured data will become increasingly useful for understanding the four levels of insight in a city. The data framework will need to declare the data it holds in order

that data from streams such as X and email can be utilized to build insights across a city, or for particular city services data streams. Typical formats used for semi-structured data include XML, JSON or GeoJSON.

11.2.4 Unstructured data

Unstructured data refers to information that either does not have a pre-defined data model or is not organized in a pre-defined manner. Unstructured data is a key data resource within the data framework for a city to develop and understand the four levels of insight.

Unstructured data contains no semantic tagging and is typically text, audio or video data. There is usually no reference data when data is unstructured and therefore the storage of the related metadata is that which allows this data to be utilized. Unstructured data is not usually repeatable and the analysis of this data usually creates aggregated data that is stored for future analytics use.

11.3 Semantic data model

The data framework for a smart city is assumed to be based on the SCCM described in ISO/IEC 30182. There are 22 prime concepts in the SCCM: 5 group concepts and 26 relationships that define the semantic interoperability of data in a city. The SCCM allows the data from any of the specific thematic domains in a city to be mapped to the SCCM to facilitate the sharing of meaning and data at a city level.

11.4 Analogue data

Analogue data refers to the data within a city that has not yet been digitized for use in the data framework. There is often valuable insight from analogue data, and the city should consider creating a digital data resource from this data. Examples of analogue data in a city are historic records, planning applications and complaint letters which might be in the original hard copy or even microfiche.Case study 10 in Annex A gives an example of the use of analogue data within the data framework. The metadata within the data framework should reflect that this has been sourced from analogue data.

11.5 Real-time data

Real-time data is information that is delivered immediately after collection. Dealing with real-time data means that the immediacy of data creates unique challenges. APIs are commonly used to process real time data. APIs are a set of routines, protocols and tools for building software and applications. Although API standards are currently not defined for smart cities, they are a key component to enable the development of the data framework, as they allow the different technology components to be made interoperable.

The development of APIs for the data framework of a smart city is outside the scope of this standard but might be a future area of smart city standardization work. In order to utilize real-time data as part of the data framework, APIs will be needed and used for at least the core city services, and to enable the four levels of insight in a smart city.

Annex A (informative) Case Studies

A.1 Shared data with public access

Project title	Smart Education for UK Cities
Project profile	**Background** The Office for Standards in Education, Children's Services and Skills (Ofsted) regulates and inspects to achieve excellence in the care of children and young people, and in education and skills for learners of all ages. It regulates and inspects childcare and children's social care, and inspects the Children and Family Court Advisory and Support Service (Cafcass), schools, colleges, initial teacher training, further education and skills, adult and community learning, and education and training in prisons and other secure establishments. It assesses council children's services, and inspects services for looked after children, safeguarding and child protection. Data View shows Ofsted inspection outcomes over five years and is updated with interim data at various points throughout the year. The tool can be used to show whether providers in an area are getting better over time and to see whether they are doing as well as those in similar areas in other parts of England. Data presented in Data View is released as official statistics. The data is provisional and shows the most recent inspections for the majority of education providers. **Three views of UK education data are provided** FlyingBinary have developed the Data View platform to publish UK education data from the internally-held Ofsted data. The platform is used internally within Ofsted and externally by providers, local enterprise partnerships, parents and journalists. FlyingBinary have used the principles of data journalism to build this platform to support a wide audience of data and non-data specialists alike, to give "one version of the truth" about education performance. **Compare regional performance over time** This Data View dashboard shows regions alongside all England results. When a region is selected, by clicking on its name, change over time is shown for that region in bar charts.

		Explore national, regional and local data Data on this Data View dashboard is shown in four panels, starting with the England level. Clicking on a region in the “Region” panel returns a list of local authority areas in the ”Local authority area” panel below. Clicking on a local authority area returns a list of individual providers and their inspection grade at the selected point in time, in the bottom panel. From there, the user can link to individual provider information on the Ofsted website. The user can also use the drop down box to change to a constituency level view. **Compare local authority areas** This Data View dashboard compares results for a selected local authority area with regional and national results. When a region and local authority area are selected, Data View shows performance over time for the selected local authority area. The user can click on a date to see a chart that ranks the selected area with up to 10 of its closest statistical rather than geographical neighbours. This allows cities to understand performance for their education provision ranked against other cities. Data View can be accessed at: https://public.tableau.com/profile/ofsted#!/vizhome/Dataview/Viewregionalperformanceovertime/ Further information is available at: https://www.gov.uk/government/publications/exploring-ofsted-inspection-data-with-data-view/
Place		England, United Kingdom
Time		2017 - Present
Reference		n/a
Relevance to this document	4.3.3	Shared Data
	Clause 4	Types of data

A.2 Answering the demand for data

Project title	Demand-led data sharing
Project profile	The Open Data User Group (ODUG) created a local authority (LA) incentive scheme. The scheme was allocated funding (£721,360) from the Cabinet Office's Release of Data Fund (RoDF) to support the first ever bulk release of demand-led open data, focusing on the priorities of the UK open data community. The open data request roadmap developed by ODUG is the evidence base for demand-led open data. This roadmap was used by the open data community with over 17,000 viewers and over 1,000 data requests raised to request that new data be released as open data, with local government data related to the smart cities agenda a key focus for open data release in many cases. When the ODUG LA incentive scheme closed, there were 209 datasets committed from 90 local authorities. This data is across all three agreed data themes: public conveniences, planning and licensed premises. This was the first bulk release of demand-led open data in the UK. The data also carries a quality mark via the Open Data Institute (ODI) open data certificates, at the RAW level. One of the key feedback points ODUG received from the scheme participants was that the use of standards, schemas and certificates of data quality allowed local authorities to release their data to an agreed data standard, in a way which was repeatable. ODUG was keen to understand whether releasing data on a thematic basis, with some but not all of local government involved, would be enough to create applications or services based on the newly released data. The scheme data is now loaded onto data.gov.uk the home of open data for the UK government. One example of the change created was related to the data for public conveniences. This data has been used to help create "The Great British Toilet map" – not just useful for tourists but being used by business travelers when they visit cities in the UK and to make UK cities more accessible for all visitors. Further information related to the demand-led open data work of ODUG can be found at: https://odug.org.uk/open-data-request-roadmap/
Place	Open Data User Group, UK Government
Time	2012 - 2015

Reference		n/a
Relevance to this document	4.6.2	Demand-led model
	Clause 4	Types of data

A.3 Research based model

Project title	Natural urban design
Project profile	University College London (UCL) have conducted some in-depth research looking into how urban planning techniques could benefit from new models for data sharing in a city [3]. Cities appear to display similar features and mechanisms even though they might be quite different in terms of their geographies or other urban indicators. This phenomenon seems to hold despite the effect of different planning policies in different cities. In questioning the nature of this behaviour, UCL examined how cities reveal their different characters through a wide range of different data sources: some historical, some narrative, some quantitative and some current. This research showed the surprising outcome that by combining these different data streams it is possible to understand better how and why the city has been, how it is and – more importantly – how it could evolve in the future. To reveal how the dependencies between different data streams build up in time, UCL modelled and visualized the spatiotemporal relationships between accessibility and form-function variables over time in Manhattan (1880–2010). The resulting dependency network of urban transformations explained some aspects of temporality in the relationships between the network structure of streets, street widths, building heights, land values and retail land uses, but the analysis also raised new questions about how the city changed in the way that it had. These questions could only arise because of the analysis of data from different sources that had been shared in the course of this research.

		Urban development in places as different as Manhattan and Barcelona is, by and large, subject to planning interventions and zoning constraints, hence the need to acknowledge the role of planning policy in shaping the landscape of these urban regions. But understanding planning policy and its impacts also needs the use of social, quantitative and technical data and knowledge of the culture and political situations that occurred over the years and the interrelationships between all of these in space and time. Thus the possibility to share such data could yield new insights in to the city's history but also could indicate potential pathways for the future, so the availability of different data streams and the willingness to combine them is crucially important. An explanatory theory about how cities display this convergent behaviour would be very helpful for urban planners and designers when considering possible interventions, but its development would be impossible without the sharing of multifaceted spatiotemporal data as used in this research. The findings still need to be supported by generalizations to other cities, to rule out the role of specific planning policies and historical circumstances that might have controlled growth patterns in these particular cases. Access to the published research can be found at: http://journals.sagepub.com/doi/full/10.1177/0265813516650200
Place		Manhattan
Time		1880 - 2010
Reference		n/a
Relevance to this document	4.6.4	Research based model
	Clause 4	Types of data

A.4 Answering the demand for data

Project title		Data Charters
Project profile		A number of UK cities have prepared charters for the sharing of data to promote trust and participation in decision making related to city services. Smart cities can create a similar approach to a charter for open and/or shared data. Three examples data charters are: — Open Data Manifesto, Leeds: Data City — Camden's Data Charter — London Data Charter A number of UK cities are now using this approach to open data charters to develop a charter for shared data and are creating shared data agreements amongst the organizations delivering services on behalf of a city. Further information is available from the following URL links. https://datamillnorth.org/dataset/information-and-guidance-documents https://www.camden.gov.uk/data-charter https://www.londonfirst.co.uk/what-we-do/competitiveness/london-data-charter
Place		Leeds, Camden and London
Time		2016 - 2021
Reference		n/a
Relevance to this document	5.4	Promoting trust and participation
	Clause 5	Establishing a data sharing culture