International Telecommunication Union

ITU-T Technical Specification

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(19 July 2019)

ITU-T Focus Group on Data Processing and Management to support IoT and Smart Cities & Communities

Technical Specification D2.1

Data processing and management framework for IoT and smart cities and communities

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The procedures for establishment of focus groups are defined in Recommendation ITU-T A.7. ITU-T Study Group 20 set up the ITU-T Focus Group on Data Processing and Management to support IoT and Smart Cities & Communities (FG-DPM) at its meeting in March 2017. ITU-T Study Group 20 is the parent group of FG-DPM.

Deliverables of focus groups can take the form of technical reports, specifications, etc., and aim to provide material for consideration by the parent group in its standardization activities. Deliverables of focus groups are not ITU-T Recommendations.

Technical Specification D2.1

Data Processing and Management Framework for IoT and Smart Cities and Communities

Summary

This Technical Specification describes the Data Processing and Management Framework organised into five dimensions that are Data lifecycle dimension, trust dimension, data commercialisation, data ecosystem and data governance. This framework covers all IoT and Smart Cities and communities applications and services.

Acknowledgements

This Technical Specification was researched and principally authored by Hakima Chaouchi (Telecom SudParis), Okan Geray (Smart Dubai), Xiaomi An (RUC), Nathalie Feingold (NPBA) and Wei Wei (RUC) under the chairmanship of Gyu Myoung Lee (Korea, Rep.of).

Additional information and materials relating to this Technical Specification can be found at: www.itu.int/go/tfgdpm. If you would like to provide any additional information, please contact Denis Andreev at tsbfgdpm@itu.int.

Keywords

Internet of Things, Smart Cities and Communities, IoT, SC&C, Data Processing, Data Management, Framework, Data lifecycle, Data Trust, Data commercialisation, DPM Value Chain, DPM ecosystem, Governance, Interoperability

Technical Specification D2.1

Data Processing and Management Framework for IoT and Smart Cities and Communities

Table of Contents

Technical Specification D2.1

Data Processing and Management Framework for IoT and Smart Cities and Communities

Introduction

Data Processing and Management framework proposes a multi-dimension representation of the data related activities. It provides the abstract view of the DPM capabilities required at each stage of the Data lifecycle considering different inherent aspects to the data such as its source (personal data, legacy data, and public data) and external aspects that are the actions to be applied to the data following the data manipulation, sharing, security and governance requirements and the commercialisation objectives. These actions require a set of DPM capabilities related to each identified dimension.

1 Scope

The Technical Specifications are expected to cover the following:

- A high level DPM framework in IoT and smart cities and communities from a capability perspective.
- The identification of the related data processing and management concepts and their relationships
- Overview of the data value chain in IoT and smart cities and communities;
- The common considerations of DPM in IoT and smart cities and communities

This document is intended to be used by:
a) Those who are engaged in Data Processing and Management activities;
b) Those who are involved in data related standardisation activities;
c) Data processing and management policy makers and regulators.

2 References

None

3 Definitions

3.1 Terms defined elsewhere

These Technical Specifications use the following terms defined elsewhere:

3.1.1 **<application>** [b-ITU-T Y.2091]: A structured set of capabilities which provide value added functionality supported by one or more services, which may be supported by an API interface.

3.1.2 **<capabilities>** [b-ITU-T X.1601]: Quality of being able to perform a given activity.

3.1.3 **<closed data>** [b-FG-DPM TS D0.1]: Data that requires access control to be divulgated.

3.1.4 **<data>** [b-ISO 16091:2018]: Information represented in a manner suitable for automatic processing.

3.1.5 **<data commercialization>** [b-FG-DPM TS D0.1]: The process of creating commercial value from data.

NOTE – It may encompass various activities, including but not limited to, monetization, valuation, pricing, licensing, distribution, marketing and sales.

3.1.6 <data exchange> [b-FG-DPM TS D0.1]**:** Accessing, transferring and archiving of data.

3.1.7 <data governance> [b-FG-DPM TS D0.1]: Set of activities aimed to design, implement and monitor a strategic plan for data asset management.

3.1.8 <data management> [b-ISO/IEC TR 10032:2003]: The activities of defining, creating, storing, maintaining and providing access to data and associated processes in one or more information systems.

3.1.9 <data marketplace> [b-FG-DPM TS D0.1]: An electronic marketplace whose main product is provisioning of data and/or related services around data.

3.1.10 <data processing> [b-ISO 5127:2017]: Systematic performance of operations upon data.

3.1.11 <data processing and management (DPM)> [b-FG-DPM TS D0.1]: The combination of all activities either directly performed on or indirectly influencing data.

NOTE 1 - Directly performed activities include among others [collecting/acquiring/capturing], exchanging, storing, securing, manipulating, reusing, aggregating, curating, disposing, monetizing and deleting data.

NOTE 2 - Indirectly influencing activities include among others policy and standards making, skills and innovation enhancement.

3.1.12 <data sharing> [b-FG-DPM TS D0.1]: The process of data exchange among different parties with specified conditions.

3.1.13 <ecosystem> [b-FG-DPM TS D0.1]: A set of organisations forming a distributed system with both technical and non-technical properties.

NOTE - In DPM, ecosystem refers to a data ecosystem, which is comprised of the technical and non-technical factors and mechanisms which directly or indirectly impact DPM activities in an ecosystem, based on various degrees of interoperability. Factors and mechanisms include, but are not limited to, data laws, regulations and policies, data standards, data skills, data research and development programs, data entrepreneurship, data economy financial incentives and data platforms.

3.1.14 <Internet of Things> [b-ITU-T Y.4000]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

NOTE 1 – Through the exploitation of identification, data capture, processing and communication capabilities, the IoT makes full use of things to offer services to all kinds of applications, whilst ensuring that security and privacy requirements are fulfilled.

NOTE 2 – In a broader perspective, the IoT can be perceived as a vision with technological and societal implications.

3.1.15 <interoperability> [b-ITU-T Y.101]: The ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged.

3.1.16 <minimal interoperability> [b-FG-DPM TS D0.1]: The minimal sufficient degree needed to meet a certain requirement for data sharing, use and reuse.
NOTE – This is an approach to build a set of modular mechanisms, including information models, across multiple domains, locations and events. The definition aligns with the definition of “interoperability” in [b-ITU-T Y.101].

3.1.17 <lifecycle> [b-ISO/IEC TR 29110-5-3:2018]: Evolution of a system, product, service, project or other human-made entity from conception through retirement.

3.1.18 <open data> [b-FG-DPM TS D0.1]: Any information that has been made available for anyone under a legal framework to access, alter, and share without restrictions.

NOTE - It can be from a public source, e.g. government data, or from a business, e.g. company intelligence, and can be used for both commercial and non-commercial purposes.

3.1.19 <personal data> [b-PAS 185:2017]: Data which relates to a living individual who can be identified: a) from those data; or b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

3.1.20 <processed data> [b-ISO 5127:2017]: Data which have been transformed from raw data or from an earlier data stage into a more refined stage by data cleaning, sorting, linking, verifying and similar operations.

3.1.21 <raw data> [b-PAS 185:2017]: Data that has not been processed for use.

3.1.22 <requirements> [b-ISO 8000-2:2018]: Need or expectation that is stated, generally implied or obligatory.

3.1.23 <risk> [b-ISO 31000:2018]: Effect of uncertainty on objects.

3.1.24 <safety> [b-ISO/IEC Guide 51:2014]: Freedom from risk which is not tolerable.

3.1.25 <security> [b-IEC Guide 120: 2018]: Condition that results from the establishment and maintenance of protective measures that insure a state of inviolability from hostile acts or influences.

3.1.26 <service> [b-ITU-T Y.2091]: A set of functions and facilities offered to a user by a provider.

3.1.27 <shared data> [b-PAS 185:2017]: Data where the data owner has the legal authority to share it with one or more organizations, subject to a data or information sharing agreement which specifies that access is granted subject to specific restrictions and, where between different legal entities, is legally enforceable

3.1.28 <Smart Cities and Communities> [b-FG-DPM TS D0.1]]: Effective integration of physical, digital and human systems in the built environment to deliver a sustainable, prosperous and inclusive future for its citizens.

NOTE – This definition aligns with the definition of "Smart City" in [b-ISO/IEC 30182:2017] and with the recommendation from the IEC/ISO/ITU Smart City terminology coordination Task team [b- IEC/ISO/ITU Daft White Paper:2019].

3.1.29 <stakeholders> [b-ISO/Guide 73:2009]: Person or organization that can affect, be affected by, or perceive themselves to be affected by a decision or activity.

3.1.30 <trust> [b-ITU-T X.1252]: The reliability and truth of information or the ability and disposition of an entity to act appropriately, within a specified context.

3.2 Terms defined in this document

These Technical Specifications defines the following terms:

3.2.1 ***IoT & Smart cities Data category:*** Refers to the data classification based on the corresponding security, privacy, quality and governance rules. Examples are personal data which require privacy respect and open data which is technically and legally open to access.

3.2.2 ***IoT & Smart cities Data types:*** Corresponds to the reception regularity of data at the databases and applications, and the way this data is retrieved either directly by the application or

spontaneously by the sensors. Data types in IoT and Smart cities applications count passive, active and dynamic data.

3.2.3 ***IoT Passive data:*** is the data that is not sent on a regular basis, it is obtained by the application requesting explicitly the sensor to send its data.

3.2.4 ***IoT active data:*** is the data streamed on a regular basis by the sensors. It is also named Time Series data.

3.2.5 ***IoT dynamic data:*** refers to data generated by sophisticated sensors that are able to send and receive information to the applications directly. These types of sensors allows a full range of capabilities, including the ability to change the data that's produced, change the format of the data, change the frequency, and even deal with security issues and provide automated software updates to dynamically deal with issues.

3.2.6 ***IoT Time series data:*** is any data which is time stamped. Time-series data workload differs from the workload of other sorts of data as time-series data is primarily inserted but rarely updated. Time-series databases require treating this data on a high priority basis. Recording a huge amount of time-series data allows to build different applications that allow to analyze the past, monitor the present or predict the future of a given IoT and SC application.

3.2.7 ***DPM data lifecycle dimension interoperability:*** Refers to a set of technical and organizational arrangements that allow both Communication networks interoperability and data interoperability, and data sharing business interoperability

3.2.8 ***DPM Data Lifecycle:*** Concerns all the steps of Data since its creation to it use and disposal.

4 Abbreviations and acronyms

These Technical Specifications use the following abbreviations and acronyms:

ITU International Telecommunication Union

IoT Internet of Things

SC&C Smart Cities and Communities

STI Science and Technology Innovation

5 Conventions

None.

6 DPM Framework concepts and high-level considerations

Cities and communities and their Challenges

Cities are dense and large human settlements in which a large majority of natural and man-made resources are consumed. More than half the world's population currently lives in urban areas. These urban areas consume 75 percent of natural resources, produce 60 to 80 percent of greenhouse gas emissions, and generate more than 50 percent of all waste. Cities contribute an estimated 80 percent of gross domestic product (GDP) globally. Therefore, our social, economic and environmental challenges are mostly manifested in urban environments. Additionally, urbanization is further progressing at a hitherto unforeseen rate with 68% of the world population expected to reside in cities by 2050.

In 2015, United Nations (UN) has adopted 17 global sustainable development goals (SDGs) for the year 2030 covering social, economic and environmental development issues. Specifically SDG11, Sustainable Cities and Communities, intend to make cities and human settlements inclusive, safe, resilient and sustainable, among others.

The unprecedented growth of cities and communities across the globe is bringing about significant challenges in housing, mobility, governance, environment, energy and water, safety, and economic and social welfare, among others.

On the other hand, advances in science and technology present enormous opportunities in addressing these challenges and building novel IoT-enabled services for cities and communities. In fact, Internet of things and related technologies such as AI, 5G, Big Data, Blockchain, etc. unlock the full potential of Science and Technology Innovation (STI) [Ref] in smart cities and communities. STI-based innovative urban solutions are becoming pervasive in cities and communities to provide much needed support.

Data as an Enabler

Digital transformation and ICT initiatives together with the application of STI have proliferated substantial data creation in cities. Public and private sector organizations have accumulated data which can be turned into practical use and benefit to address cities' and communities' challenges.

In fact, aforementioned STI-based solutions almost all require data for problem solving. In this context, data acts as a strategic asset and enabler. Data allows problem identification, understanding, diagnostics, and solution. It acts as a raw material, input and, also output depending on problem and innovation selected. It also helps determine cities' and communities' problems, shape appropriate solutions and validate outcomes. Hence, it acts as a highly critical ingredient in innovating solutions.

Data and its Challenges

Utilizing data in cities and communities context poses its own data related challenges. It is important to identify data specific capabilities since they play a critical role throughout the lifecycle of data. It is also vital to ensure trust in data for accurate, ethical and secure processing while preserving and respecting privacy. Additionally, appropriate commercialization and monetization schemes might be needed to achieve tangible business benefits while contributing to sustainability of data and creation of a viable economy around it. Furthermore, it is also crucial to establish a prolific innovation ecosystem around data. This entails among others availing requisite data skills, potentially providing business support for data related innovation, conducting R&D programs, formulating policies, regulations and standards where needed. It is also highly critical to establish an appropriate governance framework around data processing and management.

The data processing and management framework depicted in this document aims to address these data related challenges from a holistic perspective. This all-encompassing approach enables policy makers as well as practitioners to appreciate the broad scope inherent in data related activities and capabilities.

Data Value chain

In the context of DPM framework, different capabilities applied on the data since its creation to its use will follow a value creation process where different stakeholders interacts at different stages of the data lifecycle. It is important to identify various, potentially commercial, value adding activities related to data, which is commonly referred to as data value chain. For instance how Data creation, collection, storage, transmission, trust and quality check, data sharing, …etc can potentially be monetized. In other words, each and every activity in the data value chain identified can potentially be monetized. In that perspective [b-FG-DPM TS D5] propose the data value chain with two parts related to what is named "data core activities" and "data support activities" that contribute directly or indirectly to value creation from data. Data support activities complement and supplement the core data activities and therefore help in enhancing value of data and / or provisioning of data. An

example of data core activities in that value chain is data creation, collection and aggregation. On the hand examples of support data activities are data laws, regulation and policies.

An organization can conduct one or more activities in the data value chain proposed in [b-FG-DPM TS D5]. In some cases, even multiple entities or organizations can participate in a single activity in the data value chain.

7 Data Processing and Management framework

This clause presents a comprehensive DPM framework. The framework is composed of dimensions and capabilities for each dimension.

7.1 Dimensions of DPM framework

The proposed DPM DPM framework is composed along five dimensions, namely:

- Data lifecycle Dimension
- Data Trust Dimension
- Data Commercialization dimension
- Ecosystem Dimension
- Governance Dimension

As depicted in Figure 1 It designed flexibly to accommodate future expansion in terms of dimensions and capabilities.

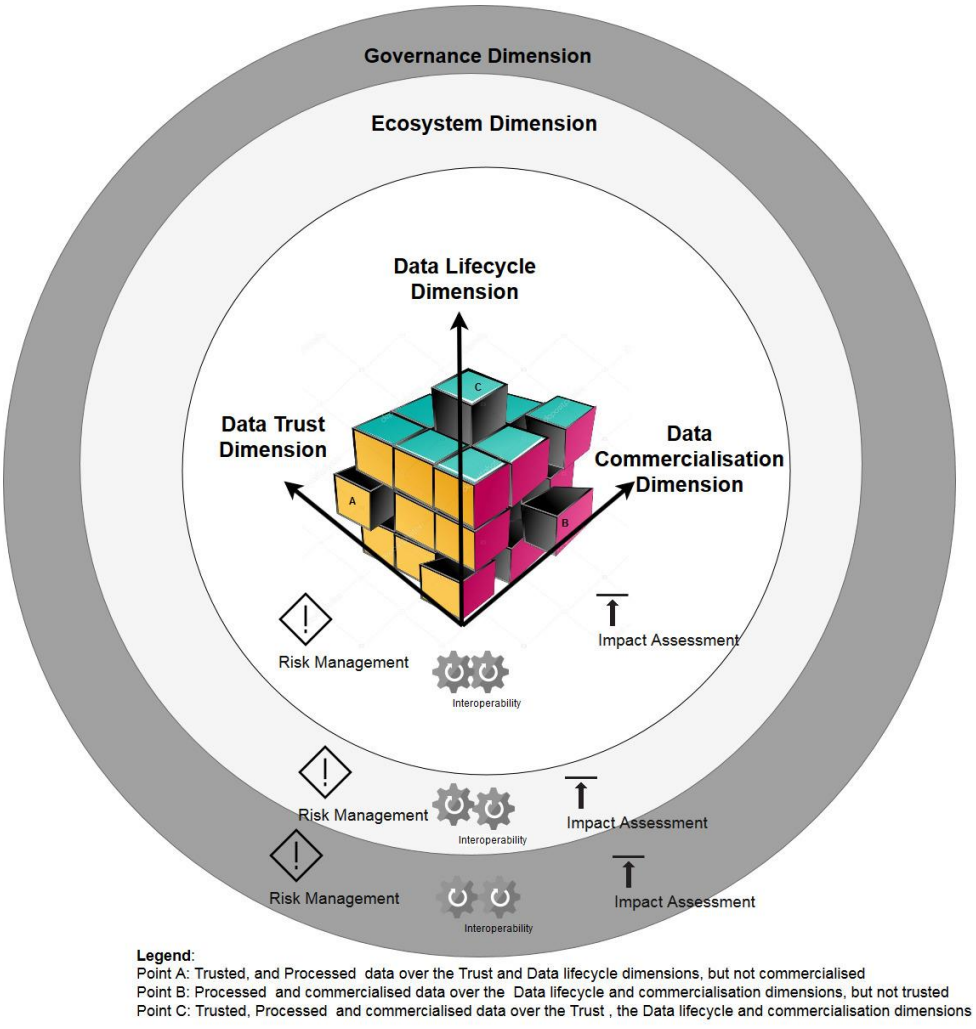


Figure 1: Data Processing and Management framework

Each dimension offers a particular perspective on data processing and management. In principle, they are independent, however taken together they complement each other.

The Data lifecycle dimension concerns the processing and management activities conducted on data from its creation to its use and disposal.

The Data trust dimension includes various actions taken to safeguard the security, privacy and quality of data and enhance trust for it by also including ethical requirements.

The Data commercialisation dimension includes all activities regarding the monetization and commercialization of data.

The Ecosystem Dimension includes all factors and mechanisms that directly or indirectly impact DPM activities.

The Governance dimension will cover all the policy related aspects that will be applied on each dimension.

Note that **Interoperability aspect** is needed at each dimension to enable the added value creation and new innovations following the DPM framework.

7.2 Data Lifecycle Dimension

This dimension reflects all the steps and the related data processing and management capabilities followed by data from its creation to its use and disposal. From a data point of view, the listed capabilities might affect the state and structure of data, the location of the data, its combination with other data, its transformation, its use and its disposal.

In the data lifecyle dimension, a list of data processing and management capabilities is provided in table 1 bellow.

Dimension Capabilities	Capability description
Data source identification	The process of identifying the IoT data source type (sensor actuators, ..), its location and eventually its owner.
Data categorization	The process of identifying the related security, privacy, trust and governance level of the data which can be private, close, open or public.
Data creation	The process of an IoT device to generates data when it is available. For example a sensor monitoring environment phenomena generates data either when the phenomena changes (event data), or on a regular monitoring basis (Time Series data).
Data acquisition/retrieval/capture	The process used by IoT application that runs a logic of requesting the data from the sensors and proceeds with the data acquisition.
Data collection	The process used by IoT application to run business logic might need to collect data from different sources that might be the sensors or the databases where datasets are stored.
Data masking	The process of making the data not possible to link with its owner. Techniques as anonymization and pseudonymisation are possible to enable the privacy of data during its lifecycle.
Data organization	The process of enriching the data during its structuring with contextual information following a common Metadata model.

Dimension Capabilities	Capability description
Data transmission	The process to move data from one location to another. Require communication technologies between diferent entities as the sensor devices to the gateway, the gateway to the cloud or the servers…etc.
Data storage	The process to accumulate data for future processing and use. The duration of storage depends on the application and the security, privacy and governance rules.
Data securing	The process to control the access to the data, to ensure its confidentiality and integrity.
Data validation	The process to check the quality, the correctness and trustworthiness of data.
Data cleaning	The process of removing wrong data by using for instance Data anomaly detection, remove useless data for the application.
Data filtering	The process of removing duplicate data.
Data aggregation	The process of selecting data from different sources and aggregates them to create new knowledge value.
Data sharing	Data sharing is not only about communication technology interoperability. It is the process of combining data from different sources (devices, data bases…) and this is possible only by identifying and applying the business intelligence related rules to allow data sharing, and the communication and metadata standards to enable the effective use of data owned by different players and stored and processed by different technologies.
Data ingestion	The process of absorbing a bulk of data specified by the applications without losing the value of the original raw data.
Data integration	The process of combining different datasets, it deals with the data format heterogeneity. For example combine a dataset using a certain database standard with another dataset using other database standards requires a syntactic and semantic interpretation of the datasets to be able to integrate them correctly.
Data processing	The ability to handle data as input and apply different treatment that might modify the data, or combine it without modifying it with other data in order to produce an output that is useful for a given application or service in the data lifecycle.
Data discovery	The process that will enable application to access the *IoT data* without the need to know the actual source of *data*, sensor description, or location. It is enabled by the data abstraction techniques that hide the device heterogeneity to the applications.
Data visualisation	The process of displaying the data to be analysed for the application business intelligence purpose.
Data analysis	The process used during the data visualization to derive new knowledge useful to the application.
Data use	After the data visualization and analysis, the data use is the process that application run to either to understand the monitored phenomena

Dimension Capabilities	Capability description
	by analysing the past and monitor the present, or predict the future and build new decisions.
Data disposal	The processes of deciding of using the data either to archive it, backup it, or destroy it.

Table 1: DPM Data lifecycle dimension capabilities.

Based on the IoT and SC&C scenario, the listed capabilities might apply fully or partially and in different order and at different phases of the data lifecycle. For that, it is recommended to apply data processing and management capabilities that respect the data category related trust, ecosystem and governance rules. In that sense, data category refers to the differentiation of personal data, closed data, public data, and open data for instance.

It is also important to align the DPM capabilities selection strategies following the agreed expectations from all the DPM dimensions.

Figure 2 depicts an example of data lifecycle phases where different DPM capabilities will be activated according to the data category allowed and agreed actions. In this example, the IoT and SC scenarios data are considered to fit into three data categories following the security, privacy, trust and quality and governance related rules. These data categories might use different IoT data types as Time series data which generates huge data volume or passive data which is acquired on demand by the application. The categories are defined as:

- Closed Personal data: Data that is generated by individuals, it is private and requires access control and masking techniques to be used or divulgated, it also requires consent of use and traceability of use accessible by its owner.
- Closed organisation data, Data that is generated and administered by companies, they are confidential and require security and governance rule to be accessed or used.
- Open data: Open data refers to any information that has been made available for anyone to access, alter, and share. It is open technically but also legally. It could be from a public source, e.g. government data, or from a business, e.g. company intelligence, and can be used for both commercial and non-commercial purposes.

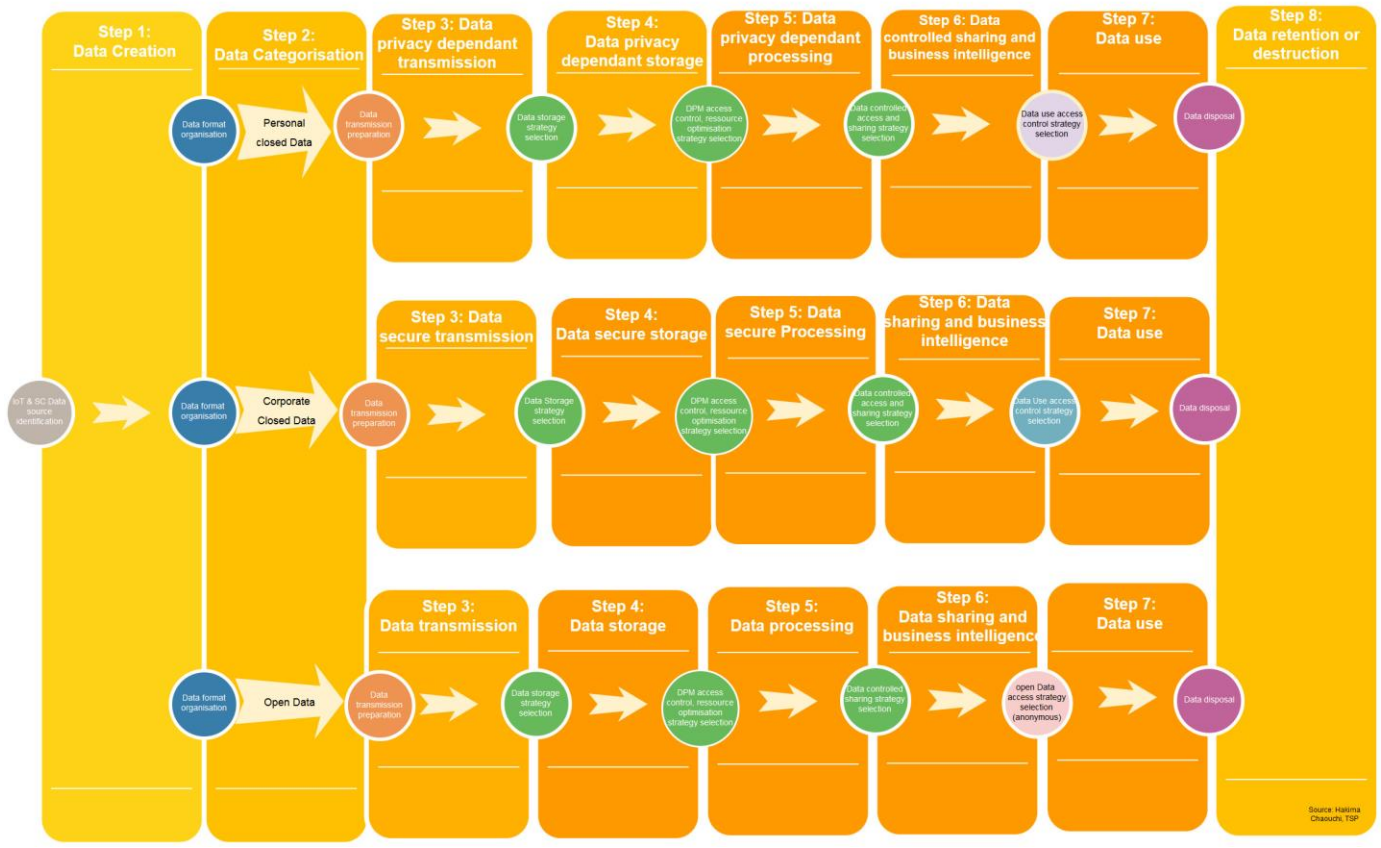
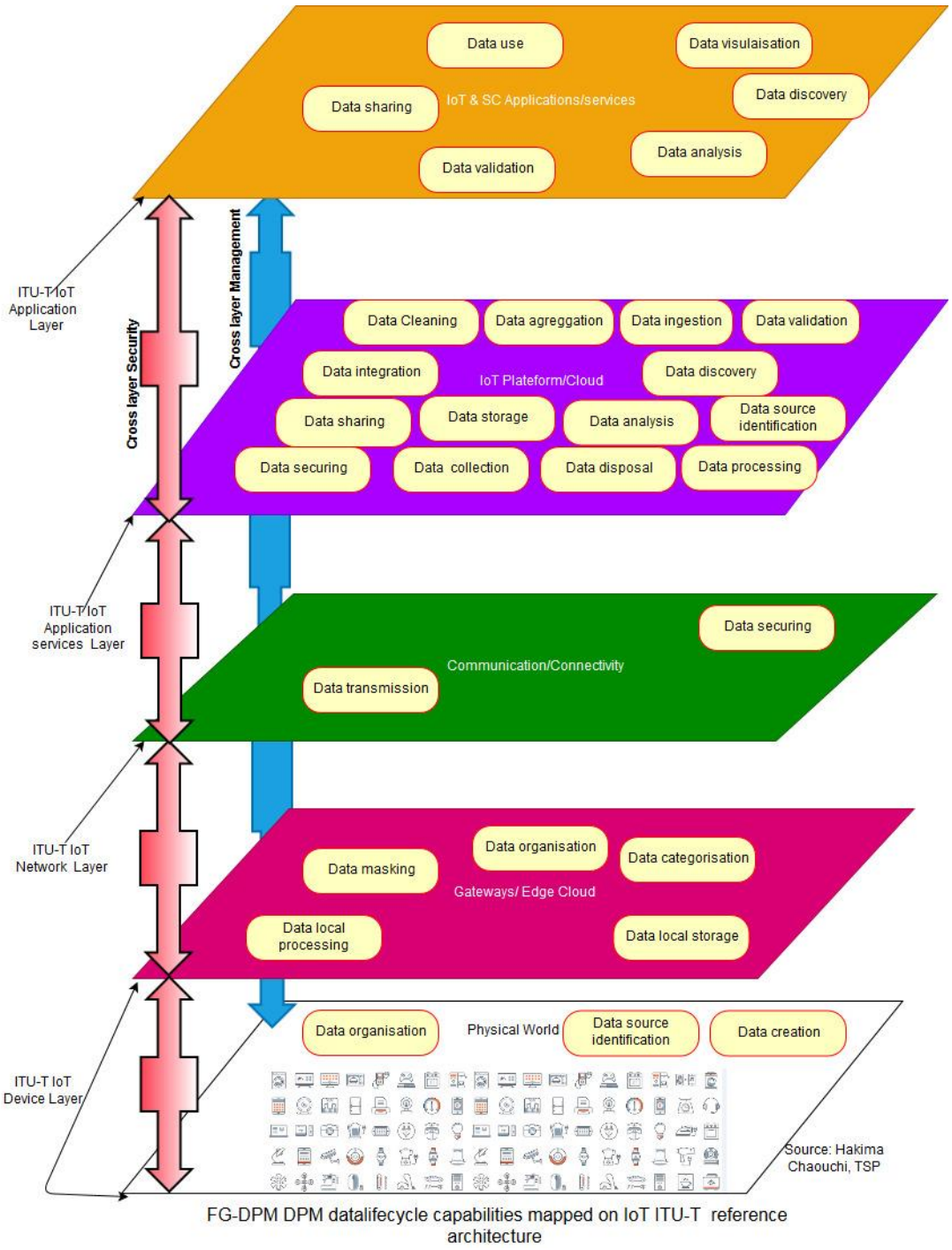


Figure 2: IoT & Smart Cities Data life cycle dimension phases Examples.

Note that the above mentioned data lifecycle phase are not fixed in that order for all IoT and SC&C scenarios. Also the data categorisation phase is provided here as a possible way to proactively know what would be the DPM capabilities allowed to be applied on the data, but this is not mandatory, it fully depends on the DPM framework related capabilities selection strategies.

Note: There are other data categories are available as Data commons [b-1] which are mainly covered in this framework under the Open Data category. The related AI commons [b-1] is out of the scope of this document, but worth mentioning as the strategy of open data might lead not only to open data for better data valuation but also open AI analytics methods and algorithms as a framework for the benefit of all humanity following the united nations SDG objectives roadmap [b-2].

Figure 3 bellow propose an example of mapping of the proposed Data lifecycle capabilities on the existing IoT & SC ITU-T reference architecture [b-ITU-T Y2060]. This is to show the DPM capabilities that might be activated at each layer of the IoT and SC&C architecture. It is important to mention that the same DPM capability might be activated at different layers following specific requirements related to the IoT and SC&C Scenarios objectives. For instance the Data storage capability might appear closer to the devices and also at the service layer. In the case of Closed Personal Data, the trust and governance dimension might lead to the strategy of storing the personal data locally and not transmit it remotely.



FG-DPM DPM datalifecycle capabilities mapped on IoT ITU-T reference architecture

Figure 3: DPM capabilities mapping example to ITU-T reference architectures.

In this example, The ITU-T device and network layers will host the data creation, structuring, and gathering. Different gateways are needed to facilitate the data gathering process. These gateways might be deployed at the Edge and at the IoT cloud platform.

The ITU-T application layer will allow users and consumers to get useful knowledge to take actions and decisions in their business or personal processes.

7.3 Data trust Dimension

This dimension proposes a broad perspective of data trust which covers all the aspects of security, privacy, and quality among others in the DPM framework including relevant capabilities to enhance data trustworthiness.

Security and Privacy capabilities will refer to all the activities that allow identification, confidentiality, integrity, availability and access control to the data. Examples are data encryption, data authentication, data integrity check, data masking, data access control…etc.

Quality capabilities will refer to all the activities that will allow the assessment of the data quality from the technical and non-technical point of views. This will enhance confidence and the reliability of data in accordance with the data quality indicators such as completeness, and uniqueness. An Example of data quality capability is fit for use check that ensures data is adequate for its intended purpose.

[b-Y 3052] provides recommendation on trusted environment in IoT where different capabilities are provided. The following table provides non exhaustive capabilities needed in this trust dimensions.

Dimension Capabilities	**Capability description**
Data Confidentiality	Ensures that information is not made available or disclosed to unauthorized individuals, entities, or processes.
Data Integrity	Ensures the accuracy and completeness of data over its entire life cycle.
Data Availability	Ensures accessibility and usability upon demand by an authorized entity
Data Unlinkability	Ensures that a user may make multiple uses of resources or services without others being able to link these uses together
Data Transparency	Ensures that an adequate level of clarity of the processes in privacy-relevant data processing is reached so that the collection, processing and use of the information can be understood and reconstructed at any time.
Data Intervenability	Ensures that users, data controller, data processors and supervisory authorities can intervene in all privacy-relevant data processing

Table 2: Data Trust Capabilities

One of the major concerns in IoT & Smart cities applications is the lack of an integrated standard framework to ensure “trusted data” which is fundamental to quality of services. In case of trusted data most of the issues are from non functional aspects as introduced the table bellow [b-4.3]. There are several considerations.

Environment enablers	The process to take external factors as supporting sources to enable trusted data, including social context, macro-economy, physical environment, infrastructure, security threats and organizational change.

Policy enablers	The process to include top management, evidence-based decision making, leadership and accountability and customer and other stakeholders' focus into data trust enhancing activities, including trusted assessorship and authority control of technology.
Technology enablers	The process to use security technology to promote the systematic and appropriate use of data, including data authentication, data authorization, data anonymization, data encryption, data ethics, data privacy by design, security labels and personalized I/O device
Data quality enablers	The process to enhance traceable actions and appropriate metadata to promote the accuracy, completeness, consistency, continuity, timeliness, uniqueness and validity of data to make it trustworthy, including accessibility data quality, contextual data quality, intrinsic data quality and representational data quality of technology .

Table 3: Trusted Data enablers

Note: In the context of IoT and SC&C data security, trust and sharing, emerging approach namely Blockchain with new concepts and methods are introducing distribution and decentralised authority access control. This can better respond to IoT and SC& C scalability requirement and also new trusted processes at different dimensions of the DPM framework. [b-FG-DPM TS D3.5 to D3.8] propose a detailed description of this approach and its potential use.

7.4 Data Commercialisation dimension

Individuals (consumers) and organizations routinely provide and exchange information while interacting and consuming online. This collected massive amount of information is potentially commercially valuable and is enabling new business models to emerge.

Data can be sold and purchased (i.e. traded or commercially exchanged) just like raw materials in the form of goods and services. Data can also be directly or indirectly monetized; it can be an organization's main or supplementary offering.

In this context, data commercialization is the process of creating commercial value from data, and it encompasses a set of capabilities, including but not limited to, monetization, valuation, pricing, licensing, distribution, marketing and sales as briefly described in the below table.

Dimension Capabilities	Capability description
Data Monetization	The process of generating incoming money flow with and out of data and data-derived information products and services.
Data Valuation	The process of estimating the worth of data from a data consumer perspective. Note: Contextualizing data to identify applicable use case(s) and to determine an appropriate valuation method are significant issues in data valuation.
Data Pricing	The process of determining price of data by an organization for selling it as a product / service.
Data Licensing	The process of determining data related terms and conditions for the legally binding agreement between the data licensor and the data licensee.

Data Distribution Channel	The channel through which data will be sold (distributed) by a seller to the buyer(s).
Data Marketing	The process of determining and conducting activities to create awareness for data and to incentivize its usage.
Data Sales	The process of conducting activities to fulfill a data sales order, including the receiving, processing and delivering the order.

Table 4: Data Commercialisation capabilities.

7.5 Data Ecosystem dimension

DPM activities and more generally data economy exist in a broad context and are influenced by several factors which directly and indirectly shape or impact it. These factors may be considered peripheral; however, they tend to affect how various DPM activities are in reality conducted.

Hence, the term data ecosystem incorporates all other value adding peripheral and non-specific factors (in addition to the ones addressed by other dimensions).

Below table briefly describes capabilities included in the Data Ecosystem dimension.

Dimension Capabilities	Capability description
Data Laws, Regulations and Policies	Include all the laws, regulations and policies that directly impact or indirectly shape DPM activities. Note: Data laws and regulations can be thought of as data rules promulgated by legally authorized bodies such as a government agency or an appropriating agency. Data policies are deliberate system of principles to guide decisions and achieve certain intended outcomes related to data.
Data Standards	Include all the commonly agreed norms, specifications and procedures related to data developed by Standards Development Organizations (SDOs).
Data Skills	Refer to various DPM related skills and their availability, including but not limited to, problem solving by data, collecting and drawing insights from data, performing data analytics and quantitative reasoning.
Data R&D Programs	Include all Research and Development programs undertaken by government, private sector and academia related to data, various DPM capabilities and requirements for it.
Data Entrepreneurship	Includes all activities and support schemes to create a business environment that fosters DPM start-ups and expansion of new firms.
Data Economy Financial Incentives	Refers to monetary benefits to motivate or encourage data economy related activities.
Data Platforms	Refers to readily available centralized systems for, among others, collecting, aggregating, analysing and processing large sets of data. Note: Data platforms provide connectivity and DPM capabilities by architecting and implementing DPM related ICT infrastructure, solutions and services.

Table 5: Data Ecosystem capabilities.

7.6 Governance Dimension

This dimension establishes a method for the construction of a governance framework for DPM. It takes consideration of both formal and informal drivers and influencers of governance framework. It also takes consideration of the need for governance framework to accommodate a wide range of activity that occurs from the conceptual through to detailed operations. As recommended in [b-FG-DPM TS 4.1] the goals for governance include, but are not limited to, the following:

- Create a data-driven and evidence-based informed culture.
- Promote appropriate regulation to enable, and not unreasonably contrained, the development of the sector.
- Promote equity, inclusion and transparency in the data economy including optimum access to rights of use and ownership.
- Provide a transparent and equitable system of DPM ethics
- Continue and strengthen the data governance programme by adopting a collaborative innovation community capacity building ideas.
- Provide effective ways of enabling appropriate data privacy and security protection.
- Adopt suitable and effective data management tools.
- Define the roles and responsabilities related to DPM and data analytics to improve evidence based decision making.

The framework described in figure 4 may be seen as a sequential process, albeit one that is expected to have feedback loops:

- The first step, "Strategic decision making and oversight" relates to the definition of a strategic vision.
- The second step, "Principles, safeguards and standards", relates to the definition of a roadmap and of the tools allowing to reach the strategic goals defined in step 1, as in term of operational data governance ("application layer") that enables efficiency of the data processes, than in term of strategic data governance ("ecosystem layer") that fosters innovation and economic development through the setting of rules and the diffusion of good practices, while favoring social development, ethic and systemic resilience (graph 2).
- The third step, "Management", relates to the operational framework that needs to be deployed to implement the strategic vision, given the principles, safeguards and standards defined by the Ecosystem.

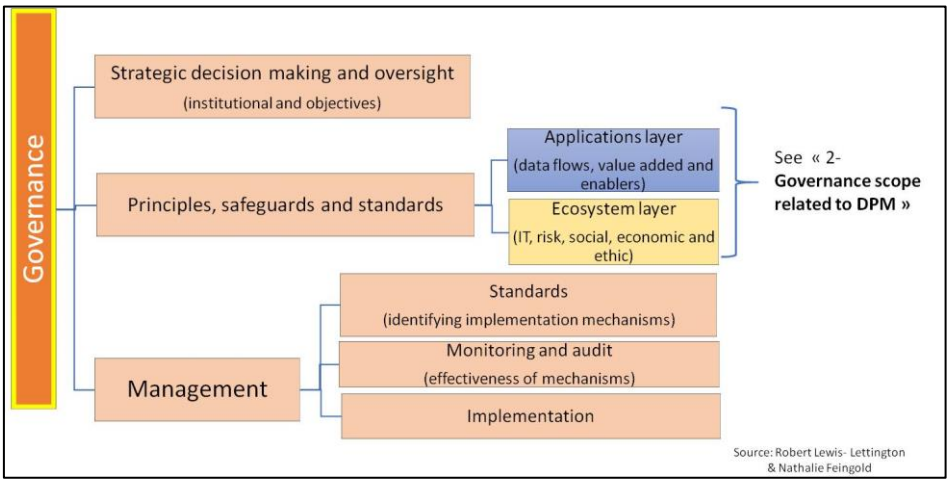


Figure 4 – Governance framework for DPM in SC&C

[b-FG-DPM TS 4-1] introduces the Jurisdictional framework for the governance of DPM in IoT and SC&C depicted in Figure 5, the sources of the drivers and influencers of governance framework, that need to be accommodated for this framework to be both legitimate and locally relevant.

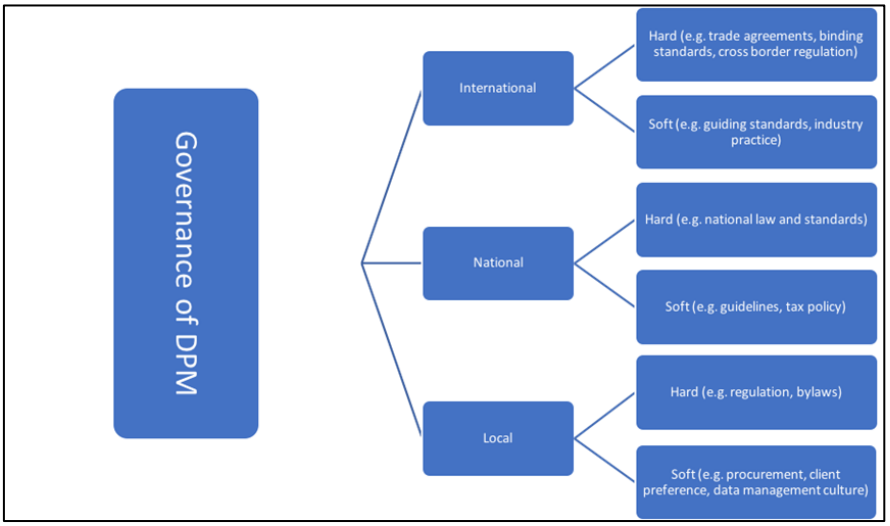


Figure 5. Jurisdictional framework for the governance of DPM in IoT and SC&C

8 Common considerations for dimensions

8.1 Interoperability

The FG-DPM has taken a holistic, ecosystem-based approach to understanding data processing and management to support IoT and smart cities & communities. In the context of the DPM Framework, ecosystem refers to a data ecosystem, which is comprised of the technical and non-technical factors and mechanisms directly or indirectly impacting DPM activities in an ecosystem, based on various degrees of interoperability. Factors and mechanisms include, but are not limited to, data laws, regulations and policies, data standards, data skills, data research and development programs, data entrepreneurship, data economy, financial incentives and data platforms.

Therefore, while interoperability is a core concept for the entire field of DPM, it is important to recognise different types and degrees of interoperability.

Degrees of Interoperability

In order to overcome the complexity and heterogeneity of IoT in smart cities and communities, the DPM Framework recognises the principle of minimal interoperability. It is an approach to establishing a set of modular mechanisms across multiple application domains and geographic territories, without having to specify everything in complete detail, and without requiring complete implementation of and compliance to the entire framework.

The DPM Framework does not directly establish such minimal interoperability mechanisms, but recognises the need for further work in this regard, preferably building on existing work, such as the *OASC Minimal Interoperability Mechanisms* (MIMs) [b-OASC].

While the principle of minimal interoperability can be applied to all types of interoperability, in the case of the OASC MIMs, interoperability would at least include: (1) context information management (cross-domain information meta-model), (2) shared data models (domain-specific information models) and (3) ecosystem transaction management (conditions for exchange).

Types of Interoperability

As shown in Figure 1 DPM framework, the interoperability applies to all the five dimensions described in this technical specification as shown in the figure bellow.

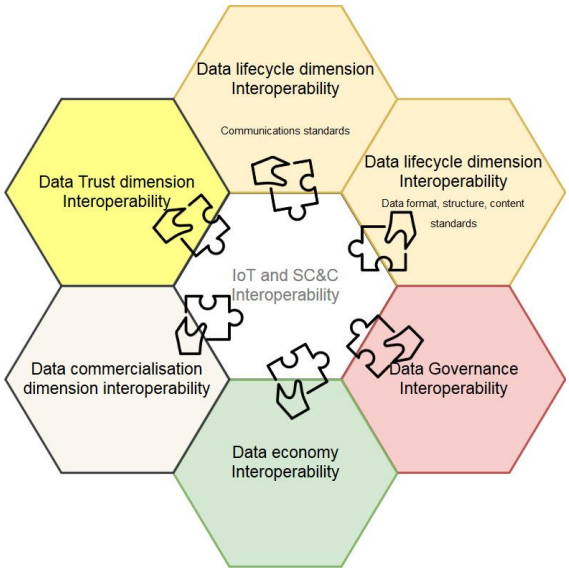


Figure 6: Types of interoperability in the DPM Framework

While Data lifecycle interoperability is concerned more by technical issues as controlling devices and communication technology heterogeneity [b-3], multi-domain data aggregation and integration standards used in case of data migration between databases or platforms, and Metadata and ontologies that are related to the semantic interoperability of data [b-FG-DPM TS D3.3], The trust dimension, the governance dimension interoperability, Ecosystem dimension interoperability and commercialisation dimension interoperability will be also concerned by other technical and non-technical issues.

Capabilities related to interoperability

To give an example drawn both from the work on requirements [TS D1.1] and on interoperability [b-TS D3.3], Annex I summarises capabilities related specifically to semantic interoperability with a focus on platform interoperability.

8.2 Risk Management

Concepts and components

Many unknown unknowns stem from data management and processing in IoT and SCC, with a direct impact on the ecosystems's resilience and stability. This uncertainty generates a variety of risks that should be identified and managed through a risk management process to ensure a safe development of cities.

To this purpose, existing standards and methodologies could be adapted to the specific needs, contexts and complexities of cities, communities and projects.

The common canvas of these standards and methodologies include the 4 following processes (definitions from ISO GUIDE 73:2009(E/F) Risk management — Vocabulary):

- Risk Identification : process of finding, recognizing and describing risks
- Risk Analysis : process to comprehend the nature of risk and to determine the level of risk
- Risk evaluation : process of comparing the results of risk analysis with risk criteria to determine whether the risk and/or its magnitude is acceptable or tolerable
- Risk treatment : process to modify risk

Moreover, as described in [b-FG-DPM TS D4.1], a new paradigm calls for faster implementation of a dynamic and global risk management approach, able to quickly anticipate risks, cyber risks and threats through an ongoing process.

To perform an efficient management of the risks, capabilities are required at a "micro" and at a "macro" level.

At a "micro" level, each project/city should be able to empower a risk management approach:

- Identify the different risks and threats that can appear at short, medium and long term;
- Analyse and assess the impact of each risk identified in term of consequences and likelihood;
- Evaluate the ability of the project/city to bear the risk or not, determine the risk tolerance;
- Treat each risk and threat by providing an adapted answer regarding the ability of the project/city to bear the risk or not (risk modification, retention, avoidance or sharing).

At a "macro" level, ecosystems and organizations should be able to forecast and manage risks at 360° and to provide quick answers. Thus, ecosystems and organizations should build a risk management framework able to quickly forecast and manage global risks, through an efficient decision-making process.

One important related capability is the ability to manage the global risks related to the ecosystem such as Systemic risk, and to analyse risk data and capture the big picture of the global activity through an efficient capacity of reporting and data aggregation. It also relies on the end-to-end visibility of treatments.

A second important related capability is the capacity to determine the ecosystem risk appetite and make sure that the stakeholders can afford the uncertainties generated by their strategies.

A third important related capability is the ability, at a global level, to facilitate risk management at each stage of the process : foster digital education or develop a data and risk culture for example.

The corresponding framework and its description

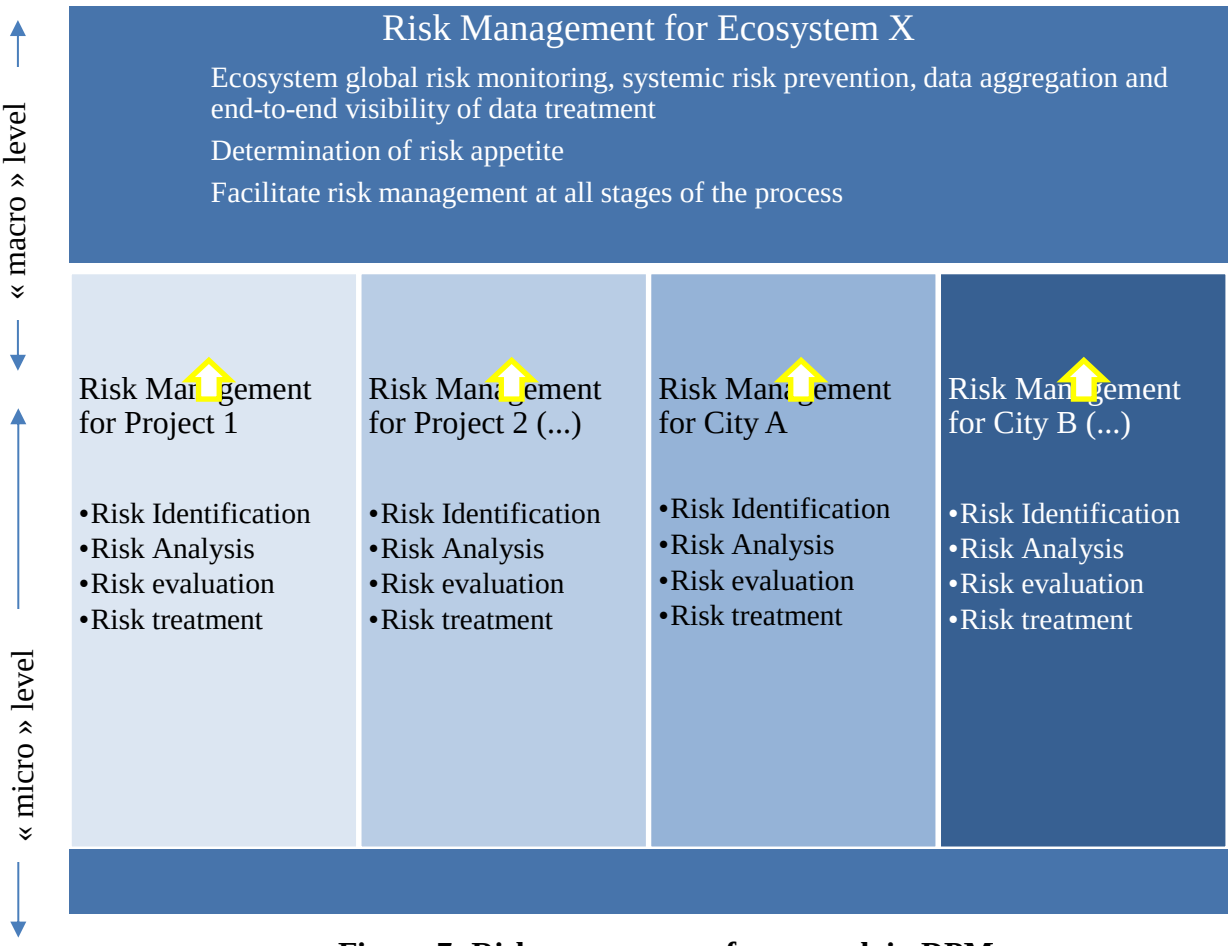


Figure 7: Risk management framework in DPM

8.3 Data Impact assessments

As recommended in [b-FG-DPM TS -D5], International Association for Impact Assessment (IAIA) defines impact assessment as "the process of identifying the future consequences of a current or proposed action". Impact assessment can help owners of DPM initiatives as well as their stakeholders determine the extent of impact for their data initiatives. It can be used to identify and circumvent the negative and unintended impacts while capitalizing on and enhancing the positive and sustainable impacts further.

Data economy impact assessment in smart cities [b-Y.4905] considers in general social, economic and environmental dimensions. [b-FG-DPM TS D5] describes in detail those impact level assessment concerns and approaches. The DPM framework clearly needs to offer at each dimension impact assessment related capabilities.

Appendix I
Semantic Interoperability Capabilities Example

(This appendix does not form an integral part of this Technical Specification.)

Data interoperability can be achieved in different ways to provide seamless integration of services in heterogeneous IoT environments. Several types of mediations can be considered to fully realize data interoperability in IoT platforms. These include semantic interoperability, syntactic interoperability, and object abstraction interoperability. These dimensions of data interoperability are briefly described as follows:

- Semantic Interoperability

The semantic interoperability is concerned with the meaning of data. Consensus on meaning is required while exchanging the data across systems. Semantic interoperability defines the true meaning of the contents that are generated by IoT devices and mutually agreed by a different system that use these contents. The semantic interoperability will enable different stakeholders to access and understand data unambiguously.

- Syntactical Interoperability

Heterogeneous IoT devices generate data that are stored and used in different formats. Syntactical Interoperability is concerned with the data formats, syntax and coding methods. Protocols used by IoT devices utilize standard syntax for communication of data. These are expressed in diverse formats such as XML, JSON or HTML.

- Object Abstraction Interoperability

Object abstraction interoperability provides the functional capabilities to support diverse object abstractions. Developing object abstraction by classification and categorization of heterogeneous data and providing metadata description and coding.

Moreover, the capabilities related to data interoperability dimensions have been described in the following table.

Data Interoperability Capabilities	**Capability description**
Data description registration	The process of data description registration provides the registration of heterogeneous data formats from diverse IoT platforms to support the interoperability provisioning process with maintained record of data formats.
Semantic data translation	The semantic translation process provides the translation of data formats to semantic formats that have been registered by semantic description registry. The translation of data to semantic formats is achieved through the defined domain ontology.
Semantic annotation	The process to provide annotation of the data based on the standard ontology and semantic data model using the selected annotation description language. It specifies the particular function of IoT resources, their information and their operations which can be understandable by other services.
Semantic alignment and linking	The alignment process is performed through an alignment strategy based on the defined semantic alignment algorithm, and is to align

Data Interoperability Capabilities	Capability description
	the information model or the semantic schema with another schema build to represent the same information but with different semantics. The Semantic alignment and linking capability enables to resolve semantic heterogeneity in systems where models have different meanings or semantics.
Semantic validation	The capability verifies the semantic conversion of data with defined schema. It constitutes of the mechanism to validate the semantic structure of the data with validation test case defined on the basis of semantic ontology.
Syntax interpretation	This capability provides conversion among diverse data formats. In case of API requests to individual platforms, this functional component translates the queries specific to the platforms, for which data or services are requested.
Schema translation	The capability to provide translation mechanism at schema level. Different platforms make use of different schemas to describe the data. The function provides the conversion to interoperate the schemas at platform level.
Syntactical validation	The capability used to verify the syntactical integrity of the translation. It constitutes the mechanism to validate the syntactical structure of the data based on the defined base schema.
Data classification	The capability provides the classification and categorization of data. It enables the classified representation to be understandable through abstract representation model. The sub capabilities of data classification process includes classifying the data based on the metadata available with the core data. Also tagging the data according to the abstract representation model.
Data integration	This capability enables mechanisms for the extraction and integration of data. As data can be from different sources, the function provides procedures to integrate the data in standard formats such as a common data model.
Metadata descriptions and coding	The capability allows the metadata to be assigned to the data converted to the representation format. It enables to assign codes and other metadata description to the core data in order to identify its sources.
Syntactic description storage	The capability to gather, manage and facilitate persistence of different data formats.
Semantic registry	The capability provides the functional process to register and manage the semantic ontology models.
Ontology discovery	The process used to search and provide matching of ontology records in the semantic description registry.
Semantic Data Model	The capability provides the ability to express the semantic meaning of the exchanged data using the information objects.

Data Interoperability Capabilities	Capability description
Base ontology translation	The capability delivers the functional process of translation of concepts from a domain ontology model to the common base ontology model.
Entity Loading	The process to load the entities from source and target ontologies in order to provide semantic alignment. To extract and load entities in semantic ontology models may comprise diverse type of information including concepts, properties, and individuals.
Template discovery	Template discovery capability provides the functionality to discover existing syntax templates which contain platform level description of syntactical formats.
Common data model Creation	The capability to develop a common data model that is a set of data elements which can be mapped among different interoperating systems. A core of a data service that is populated with different data elements and useful in multiple application domains.
API syntax convertor	The capability provides functionality to translate the API syntax description from diverse applications to standard format.
Base ontology repository	The capability provides the functionality for the persistent storage, retrieval on the entities of base ontology model. The semantic base ontology provides the generic concepts to support the semantic interoperability.

Bibliography

[b-ITU-T Y.4000/Y.2060] Recommendation ITU-T Y.4000/Y.2060 (2012), *Overview of the Internet of things.*

[b-ISO/IEC 30182:2017] ISO/IEC 30182: 2017, *Smart city concept model — Guidance for establishing a model for data interoperability.*

[b-ITU-T Y.2091] Recommendation ITU-T Y.2091 (2011), *Next Generation Networks – Frameworks and functional architecture models: Terms and definitions for next generation networks.*

[b-ISO/Guide 73] ISO/Guide 73:2009, *Risk management — Vocabulary.*

[b-ITU-T X.1601] Recommendation ITU-T X.1601 (2015), *Security framework for cloud computing.*

[b-ISO 8000-2] ISO 8000-2:2017, *Data quality — Part 2: Vocabulary.*

[b-ISO/IEC TR 10032] ISO/IEC TR 10032:2003, *Information technology — Reference Model of Data Management.*

[b-ISO 5127] ISO 5127:2017, *Information and documentation — Foundation and vocabulary.*

[b-ISO 31000] ISO 31000:2018, *Risk management — Guidelines.*

[b-ISO/IEC 120] Security aspects –Guidelines for their inclusion in publications

[b-ITU-T X.1252] Recommendation ITU-T X.1252 (2010), *Baseline identity management terms and definitions.*

[b-ISO/IEC 25010] ISO/IEC 25010:2011, Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models.

[b-ITU-T Y.2720] Recommendation ITU-T Y.2720 (2009), *NGN identity management framework.*

[b-ISO/IEC TR 26927] *ISO*/IEC TR 26927:2011, *Information technology — Telecommunications and information exchange between systems — Corporate telecommunication networks — Mobility for enterprise communications.*

[b-ITU-T Y.101] Recommendation ITU-T Y.101 (2000), *Global Information Infrastructure terminology: Terms and definitions.*

[b-IEV ref 101-12-03] Ref 101-12-03, *Mathematics/Concepts related to information.*

[b-PAS 185:2017] PAS 185, *Smart cities – Specification for establishing and implementing a security-minded approach.*

[b-ITU-T Y.3052: 2017] Overview of trust provisioning for information and communication technology infrastructures and services

[b-ITU-T Y.4905: 2019] Smart sustainable city impact assessment

[b-FG-DPM TS D0.1] Draft Technical Specification "Data Processing and Management for IoT and Smart Cities and Communities: Vocabulary"

[b-FG-DPM TS D3.5] Technical Report D3.5, "*Overview of Blockchain for supporting IoT and SC&C in DPM aspects*".

[b-FG-DPM TS D4.1] Technical Report D4.1, "Framework of Security and Privacy in DPM".

[b-FG-DPM TS D4.3] Technical Report D4.3, "Requirements and Concerns about Technical Enablers for Trusted Data".

[b-FG-DPM TS D4.4] Technical report D4.4, "Data Quality Management for Trusted Data".

[b-FG-DPM TS D5] Technical Specifications D5.1-D5.4, "Data Economy Impact, Commercialization and Monetization".

[b-1] *https://www.itu.int/en/ITU-T/AI/2018/Documents/Presentations/Session%20Day%203%20-%202pm.pdf*

[b-2] *https://sustainabledevelopment.un.org*

[b-3] PrIoT: Prototyping the Internet of Things, N Pawar, T Bourgeau, H Chaouchi, FICLOUD 20182018 IEEE 6th International Conference on Future Internet of Things and

[b-OASC] Open & Agile Smart Cities https://oascities.org/