DATA SPACE FOR
SMART AND SUSTAINABLE
CITIES AND COMMUNITIES

Deliverable 3.1

Catalogue of Specifications

WP 3 – Technical blueprint

Authors: Clara Pezuela (FIWARE), Chandra Challagonda (FIWARE), Sophie Meszaros (OASC), Thimo Thoeye (OASC), Justine Gangneux (Eurocities), Gaber Terseglav (CCIS), Andreja Lampe (CCIS), Flavio Fuart (CCIS), Dennis Wendland (FIWARE), Jesús Ruiz (FIWARE), Alberto Abella (FIWARE)

Reviewers: Martin Brynskov (OASC), Martin Traunmuller (AIT)

Delivery date: 31/03/2023

Dissemination level: Public

Type: Report

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or of the granting authority. Neither the European Union nor the granting authority can be held responsible for them.

Revision History

Author Name, Partner short name	Description	Date
Clara Pezuela, Chandra Challagonda	Draft deliverable	30/11/2022
Andreja Lampe	Methodology chapter	09/01/2023
Several authors	Contributions to chapters (first round)	17/02/2023
Several authors	Contributions to chapters (second round)	03/03/2023
Clara Pezuela	Draft version for internal review	24/03/2023
Reviewers	Revision and feedback	29/03/2023
Related authors	Integration of comments	30/03/2023
Clara Pezuela	Final version to submit	31/03/2023

Abbreviations

SSCC	*smart and sustainable cities and communities*	*EU*	*European Union*
DSSC	*Data Spaces Support Center*	*MIT*	*Massachusetts Institute of Technology*
BB	*Building Block*	*EBSI*	*European Blockchain Services Infrastructure*
MIM	*Minimal Interoperable Mechanism*		
WP	*Work package*		
DSBA	*Data Spaces Business Alliance*		

Table of Contents

Executive Summary

This Catalogue of Specifications aims at providing an overview of the identified building blocks (BBs) (technical and non-technical) required to set up and operate the data space for smart and sustainable cities and communities (SSCC). The Catalogue follows the taxonomy proposed by the OpenDEI project and adopted by the Data Spaces Support Center (DSSC). The specified BBs are mechanisms to implement the Minimal Interoperable Mechanisms (MIMs Plus) promoted by the Living-in.eu[1] initiative. The Catalogue (Section 3) is accompanied by an introductory chapter (Section 1) which introduces all these concepts and initiatives to understand the background of the context. Section 2 summarises the collected information through the survey and several interviews which have been conducted. All this information has been analysed and properly integrated into the Catalogue (Section 3) in a reasonable and homogeneous manner. Section 4 describes the meaning of MIMs in the SSCC data space and establishes the mapping of MIMs onto the BBs. Finally, Section 5 compiles all the content in this document by extracting some conclusions and identifying next steps which link this document to the next outcome of WP3, the Reference Architecture and Cookbook. Complementary to this document, an online Catalogue has been developed to facilitate access to information and the evolution of the Catalogue in the following stages.

Several Appendices have been added for:

- Appendix I: Relevant regulations and legislations in the scope of the project.
- Appendix II: The detailed assessment of the input received through the survey to cities and suppliers.
- Appendix III: Detailed information about MIMs Plus and their application in data spaces.
- Appendix IV: The methodology that has been used to gather information for the Catalogue and all the involved stakeholders.

[1] https://living-in.eu/

1 Introduction

This deliverable is the first outcome of *WP3 Technical Blueprint* combining the work carried out in T3.1 to T3.4. In the scope of each task, we have identified through desk research most relevant standards, specifications and reference implementations for each of the pillars proposed by the OpenDEI framework. This work has been complemented with inputs gathered from a survey conducted at project level and also with a series of interviews with experts in the field. The result and the process are described in this document.

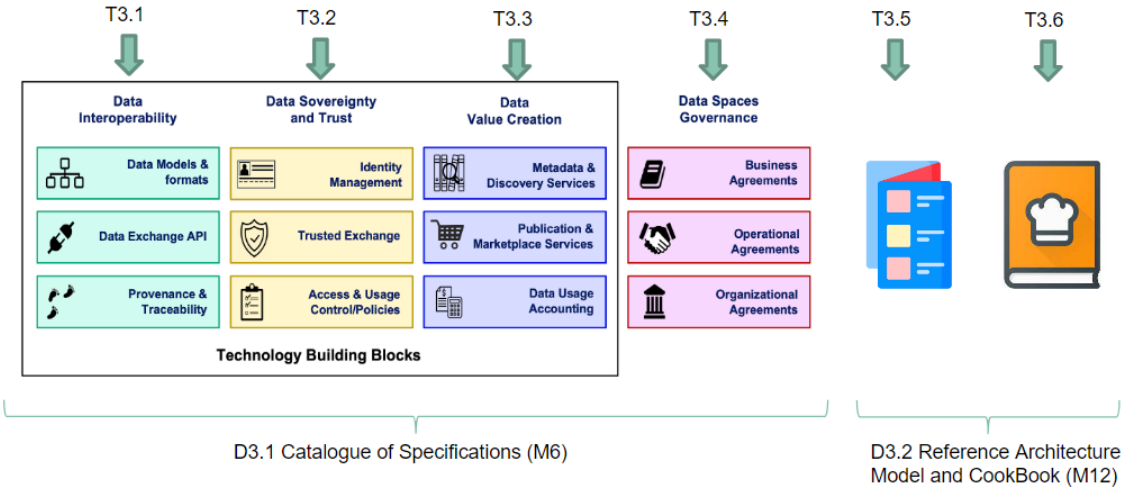

Figure 1. DS4SSCC WP3 structure and outcomes

Some of the activities, especially in relation to task T3.4 have been done in collaboration with WP2 in charge of the discovery of governance schemas for the data space for SSCC. Both WPs have been nurtured mutually during the process to come up with a consistent outcome.

The result of this document will be of high relevance for WP4, since this Catalogue together with the Reference Architecture under development, will be validated in the context of the WP4 by the application of real world use cases..

The content of this document is fully aligned also with the DSSC. Despite the timeline constraints (DSSC started at the same time as the preparatory actions for the data spaces), a continuous relationship and alignment with DSSC guidelines and recommendations has been maintained. In this regard, several considerations need to be taken into account:

- DSSC has not delivered at this time a reference Catalogue of BBs, so DS4SSCC has produced its own Catalogue, which might overlap with the DSSC Catalogue when available. As this will be the case for all data spaces, there will be a discussion with DSSC about how to federate the Catalogues of BBs across DSSC and all data spaces. At this point, it is an open topic.

- DSSC has published a collection of standards and technologies relevant to the data spaces. DS4SSCC will provide feedback, endorsement and contribution to those items listed in the collection which are in the scope of SSCC.
- The DS4SSCC Catalogue is the compilation of BBs based on the project's' experts' research, complemented by the collated inputs from cities and suppliers, organised under the taxonomy of OpenDEI framework and mapped into the Living-in.eu MIMs.

1.1 Data Space Introduction & References

The DS4SSCC leverages and builds on existing initiatives, standards and frameworks, adopted by the stakeholders in the domain, such as:

- OASC MIMs
- Living in.EU MIM+
- FIWARE Reference Architecture for Smart Cities
- Smart Data Models
- GAIA-X Technical Specifications
- OpenDEI Design Principles for Data Spaces Position Paper
- Data Spaces Business Alliance Technical Convergence document
- IDS-RAM specifications
- SITRA Rulebook for a fair data economy

The European data strategy aims to speed up the development of the European data ecosystem and economy, to harness the value of data for societal benefit, and to ensure Europe's global competitiveness and data sovereignty. The European Commission is investing in common European data spaces in strategic economic sectors and domains. There are many fundamental technical, organisational, legal, and commercial challenges that exist in developing and deploying data spaces to support data ecosystems. The DSSC is coordinating all these common European data spaces to ensure all of them are interoperable. DSSC has just released the final version of the Starter Kit for Data Space Designers to provide foundation to data space designing. It describes the ongoing and future work of DSSC in the 5 dimensions of a data space driven by BLOFT approach: Business, Legal, Operational, Functional and Technical. The Functional dimension is based on the building block taxonomy of the Open DEI model; and the Technical dimension refers to the standards, specifications and implementations that are recommended for implementing the building blocks. Additionally, the DSSC has also released the proposed glossary of terms around a data space.

Data Space is defined by the DSSC in its glossary as "*an infrastructure that enables data transactions between different data ecosystem parties based on the governance framework of that data space. Data space should be generic enough to support the implementation of multiple use cases*". There are many other definitions

but DS4SSC sticks to the DSSC definition as it has been co-created with other data space initiatives and endorsed by a relevant team of experts in the field.

In the same glossary, a BB is defined as "*a basic unit or component that can be implemented and combined with other building blocks to achieve the functionality of a data space*", thus a Catalogue of BBs is "*an organised inventory of data space building blocks recommended by the Data Space Support Centre. The Catalogue defines the building blocks and provides references to multiple options for their implementation*". Thus, this document contains the Catalogue of BBs which are recommended by the DS4SSCC project to build a data space in this domain.

The DSSC[2] is *the virtual organisation and EU-funded project which supports the deployment of common European data spaces and promotes the reuse of data across sectors*. DSSC has established the role of Relationship Manager as a bridge between the DSSC and each of the data spaces. The nominated person in this role for DS4SSCC maintains biweekly meetings and punctual communications with the DS4SSCC project coordinator to update each other about progress and plans.

DS4SSCC has also established relationships with other data spaces like Mobility, Green Deal or Energy. Through the DSSC workshops and thematic groups, DS4SSCC is engaging with these data spaces to share objectives, plans and stakeholders. In relation to the Catalogue, DS4SSCC is ahead in the definition of concrete BBs, so we expect that other data spaces can learn from our experience and leverage on our work. There is a plan to extensively promote this Catalogue across the DSSC and other data spaces.

1.2 Taxonomy of BBs: learnings from previous work

Several initiatives have been working on data spaces during the last years. This previous work has covered topics ranging from the data space concept to the concrete elements that are required to build a data space. In order to organise all these elements, a taxonomy is required. Thus, a taxonomy is defined in short as "a hierarchically ordered controlled vocabulary".

OpenDEI project[3] did a useful exercise in 2021 to come up with a set of design principles and a taxonomy of building blocks which covers all the aspects in relation to data spaces. The work was published in a widely known report named Design Principles for Data Spaces and the specifications of the proposed building blocks are available publicly onGitHub.

Other further projects, such as i4Trust[4], a DT-ICT-05 project focused on enabling trustworthy and effective data sharing, have started to implement some of these

[2] www.dssc.eu
[3] https://www.opendei.eu/
[4] https://i4trust.org/

building blocks, providing open reference implementations that can be perused by anyone who wants to build a data space.

Based on this taxonomy, the Data Spaces Business Alliance[5] (DBSA) released a Technical Convergence Discussion Document, a paper which aimed at defining the trust anchor framework, shared Catalogues and marketplaces and a policy definition language for data spaces.

The DSSC is also relying on the OpenDEI framework and DSBA Technical Convergence as baseline references for developing the blueprint for data spaces.

DS4SSCC follows also this Data Spaces Building Blocks taxonomy as shown in the picture below. This taxonomy is evolving continuously with many details from various projects like i4Trust and initiatives like DSBA and DSSC.

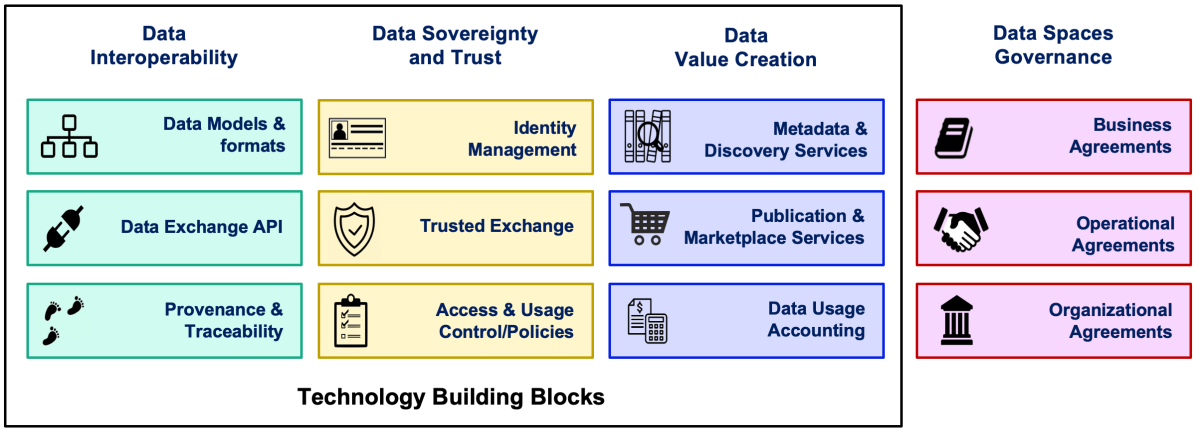

Figure 2. Building Blocks taxonomy recommended by OpenDEI and adapted by the DSBA Technical Convergence Discussion Document

The current description of the taxonomy for every block according to the OpenDEI paper is defined below. The DS4SSCC Catalogue described in section 4 will come back to these definitions to provide more details on specifications and available implementations.

Data Exchange API

Data providers joining data spaces must be able to publish data resources at well defined endpoints knowing that data consumers, unknown by them a priori, will know how to retrieve and consume data through those endpoints. Data consumers, on the other hand, must know how data available through discovered endpoints can be consumed. This is achieved by adopting domain-agnostic common APIs for data exchange.

Data Models & Formats

[5] https://data-spaces-business-alliance.eu/

Combined with the data exchange APIs, achieving full interoperability also requires the adoption of common data models to be represented in formats compatible with the API.

Provenance & Traceability

This provides the means for tracing and tracking in the process of data provision and data consumption/use.

Identity Management

The building block of Identity Management allows identification, authentication, and authorization of organizations, individuals, machines and other actors participating in a data space.

Trusted exchange

Trusted data exchange among participants provides certainty that participants involved in the data exchange are who they claim to be, and that they comply with defined rules/agreements.

Access & Usage Control / Policies

Access and usage control guarantees enforcement of data access and usage policies defined as part of the terms and conditions established when data resources or services are published or negotiated between providers and consumers.

Metadata & Discovery Services

This building block incorporates publishing and discovery mechanisms for data resources and services, making use of common descriptions of resources, services, and participants.

Publication & Marketplace Services

To support the offering of data resources and services under defined terms and conditions, marketplaces must be established. This building block supports publication of these offerings, management of processes linked to the creation and monitoring of smart contracts (which clearly describe the rights and obligations for data and service usage), implementation of clearing house functions.

Data Usage Accounting

This building block provides the basis for accounting access to and/or usage of data by different users. This in turn is supportive of important functions for clearing, payment, and billing (including data-sharing transactions without involvement of data marketplaces).

1.3 Data Space for SSCC

This section explains the need of data spaces for cities and communities to foster the digitalization of their infrastructures and services. It describes the current challenges cities and communities are facing and how data spaces may provide a way of addressing them. It also links the data spaces paradigm with the MIMs Plus, which were developed in the context of cities trying to build a minimal set of capabilities to enable interoperability among cities and communities, and governed through the Living-in.eu initiative.

1.3.1 The role of data spaces in the digital transformation of cities

The EU defines "cities and communities" as geographic areas that have legal status, representation, self-governance, and are recognized by the member state. They are typically defined by location and face complex, multidimensional problems that require cross-disciplinary solutions.

A "smart city or community" leverages all its resources, including people, organisations, infrastructure, and finance, to effectively tackle these issues through the use of data. It improves decision-making for both citizens and managers by providing them with accurate, up-to-date information.

Traditional decision-making processes in cities and communities often lack information and may lead to ineffective solutions. Smart cities and communities address these problems by utilising technology to gather and analyse more comprehensive information. This enables better alignment of city objectives, more informed decision-making, and effective problem-solving, ultimately improving the lives of citizens.

Many cities already use open standards to break the silos (e.g. exchange information across departments, sectors and services) and to integrate different verticals. Many tech providers are developing smart solutions for multiple cities, and, through standardisation, cities can avoid vendor lock-in. Cities are used to making their data public (open data), but it is still challenging to transition to a data ecosystem with other external organisations to share data that is not always public and implement a business model benefitting all parties. This new scenario introduces additional challenges, such as defining clear access control policies over the data and services, or developping precise monetisation mechanisms to motivate vertical solutions providers to participate.

Data spaces are the natural next step for cities in their digitalisation journey. Thanks to these, they can move from a scenario of systems of systems in which they integrate verticals by sharing data within the city towards a scenario in which they can go beyond and improve processes by sharing data with other organisations. In this way, cities can be consumers of services and data and offer data and services to third parties.

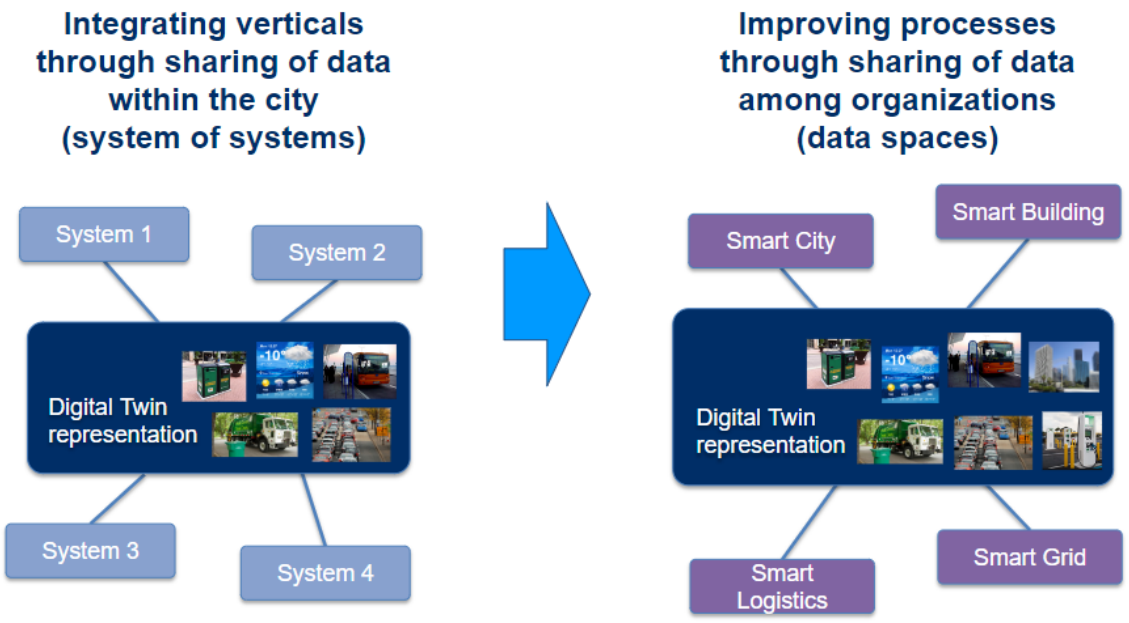

Figure 3. Expected transition of cities towards the data spaces

Cities are becoming enablers of the data economy, not only because of the massive amount of data they are able to collect and generate, but also due to the fact that they are well positioned to create ecosystems. Cities manage a great number of solutions, systems, data and devices from multiple providers that can be consumed and exploited by multiple users. This is the perfect environment in which to constitute a data space where different types of data and interests can converge and be harmonised to provide a common goal: better quality of life in cities for people and businesses and city's sustainability.

In this context, the digital twin concept becomes a mechanism to represent the physical world of the city virtually. In this representation of the real world, data collected by various devices and sensors can be constantly analysed and processed, and simulations, potentially based on artificial intelligence (AI) can be used to dry-run certain policies without affecting the real world. This can support policy makers to make smarter decisions. Even though there is no standardised way to define a digital twin for cities, common standards and building blocks can be identified and developed to collect and harmonise the data within different silos in the city and between cities and organisations to create data spaces in which urban ecosystems can develop innovative data-driven services.

The maximum potential of data spaces will be reached when the silos will be broken not just within one domain, but across sectors. Cities are impacted not just by local government, but by utilities, industries and actors within different sectors. Therefore, cities have the potential of being drivers and facilitators of cross-domain data spaces. For example, reservation estimations of the hotels in a city combined with

the number of passengers per day counted at the airport city could help the city in better dimensioning the public transportation of the city in certain days.

1.3.2 The technical challenges for SSCC data space

Many cities and smart communities around the world have been taking advantage of digitisation and e-governance to optimise the management of the cities, open data for developers to develop new innovative solutions and to cater to growing citizen's demands. Most of the cities have created Smart City platforms where they are able to manage their day to day activities. Smart City platforms have been collecting a lot of data in the due course and various innovative applications have been developed using that data. These applications are often limited to a single smart city platform, even though these applications and innovative solutions may also benefit from interfacing with other Smart City platforms and may be appreciated by other cities and communities as well.

Cities services have been growing and becoming very complex. Hence, there is a need for data to be exchanged between many similar platforms and sometimes even between the platforms of other cities or regional governments in order to manage the complex requirements.

The problem though is interoperability and portability. To enable this interoperability all platforms need to seamlessly exchange data and should enable interconnectivity. The concept of data spaces makes it possible to enable this interoperability and connectivity. It is paramount for the evolution of current Smart City and Smart Communities platforms.

Cities also need to be able to integrate new technologies and verticals available from cloud and other application and service providers. For instance, users of smart cities are often required to re-register to service providers' own identity management systems. The implementation of a city-wide data space may prevent this.

This paves the way for data spaces formed by cities to:

1. Usedata application services offered by third parties
2. Enable service providers offering services to cities
3. Allow cities to offer data services to other organisations from different domains (logistic operators, ports, buildings, …)

2 Summary of survey results and experts' interviews

Complementary to the desk research carried out by project's partners a survey was conducted among the supply and demand sides obtaining 85 answers in total, 46 out of them answering the technical questions (54%). 18 countries were represented in the answers and the majority of the participants that answered the technical questions came from governments (17), research organisations (10) and SMEs (12). Additionally, a set of 12 interviews were held with technology experts in the domain. This section summarises the inputs gathered from both instruments and explains how these inputs were included in the Catalogue described in the section 4.

The following table shows at a glance all the provided inputs about standards & specifications, used implementations and expected adoptions in the future, according to the nine technical BBs from OpenDEI taxonomy. The figures in brackets indicate the number of overall occurrences of an item in the respondents' answers. This is a general selection of the most frequently mentioned standards & specifications per BB:

- Data Models: Smart Data Models, INSPIRE, CityGML, Datex II
- Exchange API: NGSI, API REST, GC, OpenAPI
- Identity Management: Oauth2, LDAP
- Access & Usage Control: Oauth2, W3C ODRL
- Metadata & Discovery: CKAN, DCAT, INSPIRE
- Marketplaces: CKAN

In some cases the respondents mixed standards, specifications and implementations in the same answer, but this has been analysed while transposing these inputs to the Catalogue in section 3. We have filtered the nature of the input and placed it in the right area of the Catalogue where it makes sense. Unfortunately, on the question about Usage Accounting, no inputs were provided.

The following table provides an overview of the feedback that was provided through the survey:

	Data Interoperability			Data Trust and Sovereignty			Data Value Creation		
	Data Models & Formats	**Data Exchange API**	**Data Provenance & Traceability**	**Identity Management**	**Access & Usage Control**	**Trusted Exchange**	**Metadata & Discovery Services**	**Usage Accounting**	**Publication & Marketplaces**
Used Standards & Specifications	OGC MIMs (2) OPC NGSI-LD INSPIRE (3) OSLO ISO TC/211 CityGML (3) SensorThings OASIS NeTEx/Siri DatexII (3) GBFS/GTFS (2) Smartdatamodels.org (8) MaaS Data String TOMP BigQuery SQL PowerBI OPIN LDES OpenBanking & PSD2 EMF	JSON SQL LDES IoT Agents OGC (3) CitySDK NGSI (8) API REST (3) iShare OGC WFS Open API (2) OPIN PSD2 HL7 FHIR MQTT XMPP HTTP web sockets IoT Hub CityAPI CKAN API AWS SOAP Auroral Dutch API	DCAT-AP BIM OPC Geontology	SPID W3C WAC W3C ODRL W3C DID OAUTH2 (6) LDAP (4) AD ACM-IDM FairsFair.org CL@VE eIDAS ZVOP-2 AML-KYC GDPR X.500 DSML IDSA-IM OpenID Microsoft AD OS2 SAML2	W3C WAC W3C ODRL (2) OAUTH2 (2) ACM-IDM FairsFair.org eIDAS ZVOP-2 AML-KYC GDPR REST role access DPIA GDPR GEMMA KeyCloack API keys RBAC	OPC FairsFair.org eIDAS ZVOP-2 AML-KYC GDPR	CKAN (4) DCAT (4) INSPIRE (3) ISO Metadata Vlaanderen Datavindplaats OSGi NGSI-LD Wikidata Smartdatamodels ISO 19115		Datpublikatieketen CKAN (3) iShare MIM L3 i3Market

	CAD Excel Data Mesh DICOM CityJSON CGIAR W3C-WoT DCAT SAREF OSGi SensiNact MQTT REST APIs Dutch models Madrid models								
Used Implementations	Cesium (2) LDES FairsFair SDM (6) INSPIRE OSGi SCORE Water Turbinators OGC API Auroral Real State Core SAREF CGIAR Big Query SQL PowerBI OPIN	ChatGPT LDES GIS REST API DSBA Tech Convergence (3) GAIA-X	Metadata Vlaanderen portal D4CMMaia Geontology	SPID Microsoft ACM-IDM Tunnistamo CL@VE eIDAS ZVOP-2 AML-KC Rekono -SIPASS - Onfido/Ondata KeyCloack (3) KeyRock OpenID W3C Verifiable Credentials	ACM-IDM Openpolicyagent FairsFair.org Microsoft AD W3C ODRL DPIA Netherlands	DSMPM (i4TRust)	CKAN (2) SDDI Metadata Vlaanderen Datavindplaats SAML / LIberty INSPIRE Metadata Implementing Rules OSGi smartdatamodels		CKAN (2) i3Market

	PSD2								
Future consideration	ODIN OpenAPI Embedded Finance DATEX NETEX SIRI GIS BIM City GML IFC IDSA Connectors DSSC BBs DSBA Tech Covergence GAIA-X Auroral	OPC Flanders SDS NGSI-LD		GAIA-X DSBA Tech Covergence DSSC BBs SCIM SOLID iShare	SOLID Eclipse Context Broker DBSA Tech Convergence DSSC BBs CYPE GAIA-X	DSSC BBs	DSSC BBs UMM Joint GIS & BIM DCAT-AP-VL		DBSA Tech Convergence

Table 1 - Collected inputs from survey respondents with technical background

Additionally to the responses from respondents with a technical background, we have also considered inputs coming from respondents with no technical background. During the analysis of their answers, the experts from WP2 (governance) have identified some inputs that could be relevant for the technical BBs, and the following table shows what was extracted from there.

Protocols for data management	**Tools, processes and/or practices related to data quality assurance (implemented)**

DAMA DMBoK (5) Open vSwitch OVSDB Management Protocol OVS instances, Javascript Object Notation, JSON Remote Procedure Call [JSON-RPC]. Flemish Data Strategy: https://assets.vlaanderen.be/image/upload/v1647858968/Vlaamse_datastrategie_kacrph.pdf Flanders Smart Data Space: https://www.vlaanderen.be/digitaal-vlaanderen/onze-oplossingen/vlaamse-smart-data-space DPO - revizijska sled GGM (Gemeentelijk Gegevens Model) NGSI FAIR principles PETRA - reference architecture for provinces/ GEMMA - reference architecture for municipalities https://joinup.ec.europa.eu/sites/default/files/inline-files/Netherlands%20Factsheet%20Validated.pdf Kimball techniques csv, json, ngsi ld	Great Expectations tools (2) Data catalog, Data Excellence Central data management service Open data platform Linked open data platform APIs OPC, some control functions and checks in the programme code. DNA EQMS which we developed in our company software automated test, software coding styles, software reviews, software pair programming OSLO Toolchain: https://github.com/Informatievlaanderen/OSLO-toolchain FairsFair.org Logical & consistency controls geo database environments Some tools provided by the Azure environment Work in PISTIS will be upon relating value of data to quality, building on SafeDeed. Data ecosystem, Infrastructure Connector development, data cleansing, data analytics, Complex event processing commissions implemented for Data check on National Digital Agenda projects ISO 19157, ISO 19158, own data quality management procedures Organisation roles: e.g.: CDO, quality officers Process for data integrity checks and consistency checks between registrations Internal data quality assurance tool PowerBI muss unsere IT-Abteilung beantworten Sensor data filtering for values outside acceptable range.

Table 2 - Collected inputs from non technical background respondents of relevance for technical BBs

Complementary to the survey results, a total of 10 interviews were carried out by WP3 with technical experts, together with some inputs coming from interviews carried out in the context of WP2 that included useful technical information for WP3. The interviewed experts from WP3 were from well known organisations in smart cities domain: Martel Innovate (Netherlands), Universidad Politécnica de Madrid (Spain), IMEC (Belgium), Alastria (Spain), Libelium (Spain), Acatech (Germany), EGM (France), NEC (Germany), University of Sofia (Bulgaria), City of Kiel (Germany), Digital Flanders (Belgium), Porto Digital (Portugal).

In overall, we collected insights from well known experts in the field from academia, industry and some research centres across Europe. In most of the cases, the inputs were quite aligned with the ones gathered from the survey, which validated the results.

Intervieweeswere asked about:

- Which type of data (formats, models...) are you using/collecting in your city?
- Do you know the standards you are using for data modelling?
- Which standardised APIs are you using to exchange data within the city or with external entities to the city?
- Are you following any process for ensuring the traceability and provenance of your data? Could you describe a bit?
- Which IAM standards are you relying on?
- Which discovery and publication of data standards/mechanisms are you using?
- Are you following any standard for accounting the access of the users to the data? Which one?
- Which mechanism are you following to provide access to the data? Have you implemented any marketplace? Did you use an open implementation?
- Do you know what MIMs are? Which MIMs is your city/solution implementing?
- Any concrete reference implementation are you using for any of the functional blocks commented above?

All the participants signed a consent form to use the provided information in our reports and to allow the recording of the interviews. The interviews lasted between 30 and 45 minutes and were held during February and March 2023.

Following the summary of the answers per question:

Question	Answers summarised
Types of used data	Structured data from sensors (location, value), binary data, low level protocols, MQTT from LoRa sensors, SensorThings data, Copernicus data, GIS, BIM, HTTP and REST APIs, TCP (socket), MQTT,

	publish/subscribe protocols, CKAN, Open Data Portals, cameras, statistical data, CRM
Standards for data modelling	Smart Data Models, WFS, CityGML, BIM, DCAT-AP, RDF, NGSIv2, NGSI-LD, LDES, Data Privacy Vocabulary (DPV), OSLO
Standards for exchange APIs	NGSI v2, NGSI-LD, CKAN, Web services (Object Document Model, Object Relational Mapping), OpenAPI, IDSA Connectors, SOLID
Process for Provenance & Traceability	No traceability Some research with blockchain (DLT) Anubis Timescales DB
Standards for IAM	XACML, Keycloak, Oauth2, Open ID connect, XACML, DLT (wallets, verifiable credentials), data usage policies (UCOM), Anubis, eIDAS
Standards for Discovery & Publication	Context Broker, Open Data Portals, CKAN, Grafana (JSON, CSV, XML), DCAT-AP, SHACL, TM Forum Market place
Standards for Usage Accounting	Not really doing accounting, IUDX NGSI-LD format Metering API, TM Forum Accounting API, Counters for traffic of data Google Analytics from CKAN Anonymized data
Standards for Marketplaces	No really using marketplaces Grafana, CKAN, BAE (Business API Ecosystem)
Used reference implementations	Orion Context Broker, Scorpio, STF, Keycloak, CKAN, Draco, Keyrock, Cosmos, MongoDB, SparkScala, SOLID, Eclipse Data Space Components, RUDI project

Table 3 - Collected and summarised inputs from interviews

Practically all of the interviewees were aware of MIMs, and even used mechanisms for implementing them, especially MIM1 and MIM2, and were interested in MIM3 and MIM4.

3 Building Blocks Catalogue

By aggregating the desk research with inputs from survey and interviews, the following Catalogue of specifications for Building Blocks has been developed. The Catalogue has been structured according to the taxonomy of OpenDEI framework recommended by the Data Spaces Support Center. By following the same taxonomy, we ensure the compatibility of this Catalogue with other Catalogues that will be created by other data spaces, with the ultimate goal of federating all of them in a global and navigable Catalogue. The BBs definitions are inspired by the Design Principles for Data Spaces paper from OpenDEI.

Materialising data spaces requires making choices and adopting a minimum but sufficient set of technology standards. In the context of Living-in.EU we propose to refer to the Minimal Interoperability Mechanisms (MIMs Plus) as a guide in this decision.

This Catalogue will be made available online to facilitate its use and consultation by any city or community aiming to create a data space. Digital versions of this Catalogue will be easily accessible and maintainable along its progression when required.

The diagram below summarises the contents of the Catalogue at a glance. Each element will be described in detail in this section.

Building Block	Related Standards	Related Industry Body Specifications
Data Models and Formats	SAREF	Smart Data Models
	Data Privacy Vocabulary	Indian Urban Data Exchange
	OSLO	INSPIRE
	NeTEx/Siri	Crop Ontology
	Datex II	OPC UA
	OPIN	MaaS Data String
	PSD2	
Data Exchange API	NGSI-LD	LDES
		MQTT
		JSON-LD
Provenance and Traceability	ETSI-CIM	DCAT-AP
Identity Management	LDAP	CEF eID Building Block
	OAUTH2	OpenIdConnect
	X.500/X.509	SAML 2.0
	W3C DID	Solid
	W3C Verifiable Credentials	
Trusted Exchange	EUDI	EBSI
Access and Usage Control	XACML Policy Definition Language	Rego
	W3C ODRL	Open Policy Agent
	W3C WAC	
Metadata and Discovery Services	Dublin Core	
	DCAT	
	INSPIRE	
	ISO 19115-1:2014	
Publication and marketplace services		OpenAPI
		ICT Innovation Network reference architecture
Data usage accounting		TM Forum Accounting API
		IUDX Metering & Audit API
Business Agreements		Sitra Rulebook for a fair data economy
		Contract for Data Collaborations (C4DC)
		Open Data Institute
Organizational and Operational agreements	OSLO	DAMA DMBoK
	ISO 19157	Open vSwitch Database Management Protocol
	ISO 19158	

Figure 4 - Summary of collection

Looking at collected information available for each BB, we have carried out a qualitative assessment of their maturity based on the existence of widely adopted standards, availability of reference implementations and level of adoption by the cities and communities. Here below is the conclusion of the analysis, which may evolve in the future depending on the progress of cities and communities transiting to data spaces.

PILLAR	BUILDING BLOCK	MATURITY
Data Interoperability	Data Models	
	Data Exchange API	
	Provenance and Traceability	
Data Trust & Sovereignty	Trusted Exchange	
	Identity Mng	
	Access and Usage Control	
Data Value Creation	Publication & Marketplaces	
	Metadata & Discovery	
	Usage Accounting	

Figure 5 - Maturity assessment of BBs in SSCC domain

3.1 Data Interoperability

Data spaces should provide a solid framework for an efficient exchange of data among participants, supporting full decoupling of data providers and consumers. This requires the adoption of a "common lingua" that every participant uses, materialised in the adoption of common APIs for the data exchange, and the definition of common data models. Common mechanisms for traceability of data exchange transactions and data provenance, are also required.

3.1.1 Data Models & Formats

Functional description

The data models define the structure of the data to be shared across the data space. As a minimum requirement they have to describe the name of the different attributes present at the data sources. These attributes have a data type and, eventually, some restrictions to the possible values (I.e. between 0 and 1 for numeric values, or a specific list of strings for others). Finally, a written description explaining the meaning of the values has to be provided. This description could also point to some external resources with further information, description of the recommended units for those magnitudes with units, or other aspects.

These data describing the data sources have to be available in the data space to allow their search and eventually sharing, as denominated metadata.

The Data Models and Formats building block establishes a common format for data model specifications and representation of data in data exchange payloads.

Combined with the Data Exchange APIs building block, this ensures full interoperability among participants.

The role and scope of the Data Models and Formats building block is to facilitate a common format for data model specifications and representation of data. An example of usage would be the Smart Agrifood domain which needs a common representation of agricultural data (e.g. crops, sensor data from the field, multispectral imagery from UAVs, geolocation data, fertilisation logs, …). This common data model shall be used for all data exchanged between software components.

Baseline standards and industry body specifications

SAREF (ETSI, https://saref.etsi.org/)
SAREF (abbreviation for Smart Appliances REFerence) is a reference ontology or a common model that enables the integration of various components (standards, protocols, data models) in the field of smart solutions. The SAREF ontology is based on the concept of a "device" (e.g. a switch), where devices are tangible elements designed to perform one or more functions, such as in households, public buildings, industrial buildings, etc. The SAREF ontology defines basic functions that can be combined into more complex functions and was initiated by the European Commission in cooperation with ETSI[6].

The SAREF family of standards enable interoperability between solutions from different providers and among various activity sectors in the Internet of Things (IoT) and therefore contribute to the development of the global digital market. These standards are designed to run on top of the oneM2M system, the global IoT partnership project of which ETSI is a founding partner. OneM2M provides the communication and interworking framework to share the data among applications; SAREF provides the semantic interoperability necessary to share the information carried by the data.

In June 2019, the ETSI Technical Committee ETSI SmartM2M issued three new specifications based on the SAREF ontology, namely in the field of smart cities (SAREF4CITY: ETSI TS 103 410-4), in the field of industry and production (SAREF4INMA: ETSI TS 103 410-5) and in the field of smart agriculture (SAREF4AGRI: ETSI TS 103 410-6). The full list of SAREF extensions can be found at the SAREF site.

Smart Data Models (FIWARE/IUDX/OASC/TMForum, https://smartdatamodels.org/)

[6] http://ontology.tno.nl/saref

The FIWARE Foundation, TM Forum, IUDX and OASC are leading a joint collaboration initiative to support the adoption of a reference architecture and compatible common data models that underpin a digital market of interoperable and replicable smart solutions in multiple sectors, starting with Smart Cities. A smart data model includes three elements: The schema, or technical representation of the model defining the technical data types and structure, the specification of a written document for human readers, and the examples of the payloads for NGSIv2 and NGSI-LD versions. The schemas are coded in json schema, while the specification are open licensed and available at GitHub[7]

Smart data models include different business domains including Smart cities, Energy, Environment, Agrifood, Water, Health, Manufacturing, Logistics, etc.

India Urban Data Exchange (IUDX, https://catalogue.iudx.org.in/)
IUDX is the transformative initiative of the Ministry of Housing and Urban Affairs, Government of India to provide a data exchange platform to Indian cities. FIWARE and IUDX have been working closely in recent years to foster the adoption of open standards and to create a standard architecture for data exchange open source software. A standard for data exchange interfaces modelled after IUDX and based on ETSI's NGSI-LD was adopted in June 2021 by the Bureau of Indian Standards as a standard API for Indian cities. A shared vocabulary and data models are also available[8] based on them are available and mapped in the Smart Data Models initiative.

INSPIRE data specifications (EU,
https://inspire.ec.europa.eu/data-specifications/2892)
The INSPIRE Implementing Rules on interoperability of spatial data sets and services (IRs) and Technical Guidelines (Data Specifications) specify common data models, code lists, map layers and additional metadata on the interoperability to be used when exchanging spatial datasets.

Datasets in scope of INSPIRE are ones which come under one or more of the 34 spatial data themes (below) set out in the INSPIRE Directive. Interoperability in INSPIRE means the possibility to combine spatial data and services from different sources across the European Community in a consistent way without involving specific efforts of humans or machines. Interoperability may be achieved by either changing (harmonising) and storing existing data sets or transforming them via services for publication in the INSPIRE infrastructure.

Crop ontology (CGIAR, https://cropontology.org/)
The CGIAR or Consultative Group on International Agricultural Research is a global partnership of 15 research centres working to reduce poverty, improve food and

[7] https://github.com/smart-data-models
[8]https://voc.iudx.org.in/

nutrition security, and protect natural resources. CGIAR centres conduct research in various areas of agriculture, including livestock, forestry, fisheries, and crops. It is the organisation responsible for several sources of data models like the crop ontology.

Data Privacy Vocabulary (DPV) (W3C, https://w3c.github.io/dpv/dpv/)
The Data Privacy Vocabulary (DPV) is a data model primarily derived from GDPR. DPV provides a standardised vocabulary for describing and managing personal data and privacy-related concepts and terms.

It has been created in the W3C Data Privacy Vocabularies and Controls CG (DPVCG), it develops a taxonomy of privacy and data protection related terms, which include in particular terms from the new European General Data Protection Regulation (GDPR), such as a taxonomy of personal data as well as a classification of purposes (i.e., purposes for data collection), and events of disclosures, consent, and processing such personal data.

OPC UA (OPC Foundation,
https://opcfoundation.org/about/opc-technologies/opc-ua/)
The OPC foundation is an independent entity for the OPC Unified Architecture (OPC UA). It creates the OPC UA, which defines a set of data models for different types of industrial data, such as process data, alarm and event data, and historical data. It also includes security features, such as authentication, encryption, and access control.

OSLO (Flanders Gov, https://lov.linkeddata.es/dataset/lov/vocabs/oslo)
The OSLO repository of ontologies extends much of the work done by the ISA² programme on a regional level. While containing many ontologies and application profiles for Flanders, some of these have been more widely adopted than others. The ontology for mobility has become a standard for the Flemish region in Belgium. It defines data models for the description of the mobility, specifically it describes data models for traveller, booking, trip, service supply, licence and network.

NeTEx/Siri (UITP, https://www.transmodel-cen.eu/netex-standard/,
https://www.transmodel-cen.eu/siri-standard/)
NeTEx (Network Timetable Exchange) and SIRI (Service Interface for Real Time Information) are two related standards for exchanging public transport data, developed by the International Association of Public Transport (UITP) and adopted by the European Committee for Standardization (CEN). NeTEx is a standard for exchanging public transport schedules and timetables, including information about public transport network topology, scheduled timetables and fare information. Siri specifies a European interface standard for exchanging information about the planned, current or projected performance of real-time public transport operations between different computer systems.

Datex II (EU, https://www.datex2.eu/)

Datex II (Data Exchange for Traffic and Transportation) is a standard protocol for the exchange of traffic and travel data between different intelligent transportation systems (ITS). It is a standardised format for exchanging data related to traffic management, such as traffic flow, incidents, and roadworks, among others.

Datex II was developed by the European Union as a successor to the original Datex (Data Exchange for Traffic Telematics) protocol. The Datex II protocol defines a set of XML (eXtensible Markup Language) schemas for data exchange, which can be used to ensure compatibility between different ITS systems. The standard covers a wide range of data categories, including traffic management, public transport, weather information, and traveller information.

OPIN (Open Insurance, https://docs.google.com/spreadsheets/d/1Y0Gk_LpTvTNEfoDMdIxeD7juv3E8FKcbE3mHUJNV5JY/edit#gid=504262159)
Open Insurance Data Standard is a set of guidelines and standards for developing application programming interfaces (APIs) in the insurance industry. It covers data models for Motor insurance, Trade Credit insurance, Pet insurance, Property insurance, Business Interruption insurance, Cyber Liability insurance, Term Life insurance, Travel insurance

PSD2 (EU, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02015L2366-20151223)
PSD2 (Payment Services Directive 2) is a European Union directive that aims to regulate payment services across the EU, with the goal of increasing competition and innovation in the payments industry. PSD2 requires banks to open up access to their payment systems to third-party providers, subject to customer consent, and to provide a secure API for payment initiation and account information services. PSD2 defines a set of data models and APIs for data exchange, including payment initiation, account information, and authentication. It defines data models for Balances, Transactions, Beneficiaries, Direct Debits, Standing Orders, Products, Offers, Parties, Scheduled Payments, and Statements.

MaaS Data String (Ferdinand Burgersdijk, https://www.frcb.nl/maas/)
The three layers of Mobility as a Service (MaaS) for providing MaaS services. The first tier gives MaaS its direction and democratic accountability. The second and third tiers involve collaboration among market players to deliver MaaS services to customers. The approach emphasises the need for sharing common standards in order to share real-time data on mobility.

Available implementations

The smart data models program compiles data models based on actual use cases or adapted from open and adopted standards. They are open-licensed.

Smart data models are already adopted in different initiatives and companies: AWS (Smart Territory Framework), Microsoft (some of their DTDL classes extend them), or Atos (Urban data platform). Following some existing implementations.

India Urban Data Exchange (IUDX, https://catalogue.iudx.org.in/)
IUDX is the transformative initiative of the Ministry of Housing and Urban Affairs, Government of India to provide a data exchange platform to Indian cities. FIWARE and IUDX have been working closely in recent years to foster the adoption of open standards and to create a standard architecture for data exchange open source software. A standard for data exchange interfaces modelled after IUDX and based on ETSI's NGSI-LD was adopted in June 2021 by the Bureau of Indian Standards as a standard API for Indian cities.

IoT Big Data Harmonized Data Model (GSMA,
https://www.gsma.com/iot/wp-content/uploads/2018/07/CLP.26-v5.0.pdf)
The latest version of the document entitled "IoT Big Data Harmonized Data Model" was released on June 19, 2018 and is marked with version 5.0. The document was created under the auspices of the Global System for Mobile Communications (GSMA). The document defines data models in the field of IoT Big Data.

Data interoperability has been identified as a technical barrier that prohibits the realisation of the full potential value of IoT Big Data. To help address that problem, in referenced document data models are defined as entities or things that are commonly used in IoT Big Data applications. The definitions of the data entities have been developed through contributions from participating mobile operators and aligned with existing industry work and namespaces where possible, for example, oneM2M in Smart Home, OASC for Smart Cities and schema.org for generic entities. These collaboratively developed harmonised data models, together with the accompanying documents "IoT Big Data Framework Architecture" and "IoT Big Data NGSIv2 Profile", aim to define a framework of how mobile operators can approach the delivery of IoT Big Data services.

SynchroniCity data models (OASC,
https://gitlab.com/synchronicity-iot/synchronicity-data-models)
The data models being developed within the SynchroniCity project are the focal point of the SynchroniCity interoperability framework and are a concretization of the Minimum Interoperability Mechanisms of the OASC initiative. These mechanisms are known as MIMs and are promoted by the OASC as simple and transparent mechanisms that can be used quickly by any city (large or small), thus achieving speed and openness in introducing innovations, while reducing costs and inefficiencies. In practice, MIMs are realised as a set of APIs for data access,

content definition of data, and can also be a common platform for storing and transmitting data.

CityGML and CityJSON (OGC, https://www.cityjson.org/specs/1.1.3/)
CityGML is an open data model and XML-based format for the representation and exchange of 3D urban models. It is maintained by the Open Geospatial Consortium (OGC) and is used for a wide range of applications, such as urban planning, environmental modelling, and disaster management.

CityGML defines a set of data models for describing the geometry, topology, semantics, and appearance of 3D urban models. Some of the key data models defined by CityGML include buildings and city objects. They could be described with different levels of detail and a topology and relationships can be set between the elements. It also allows the use of textures, colors and materials. Finally it can set up the type of use of the different areas.

CityJSON is a subset of CityGML coded in JSON format, it is also created in the OGC. CityGML and CityJSON can be converted to each other. CityJSON reduces the size of the objects compared with CityGML.

Sensorthings (OGC, https://www.ogc.org/standard/sensorthings/)
SensorThings is an Open Geospatial Consortium (OGC) standard that provides an open and unified way to model, manage, and share data from IoT devices and sensors. It defines some entities related to the IoT devices and it provides a specific API to manage them.

Dutch models (Netherlands Gov,
https://www.rijkswaterstaat.nl/en/mobility/roads-and-waterways)

The National Model System (LMS) and the Dutch Regional Model (NRM) are both traffic models used in the Netherlands for transportation planning and traffic management.

The National Model System (LMS) is a macroscopic traffic model that is used for national transportation planning and policy-making. It is based on a set of statistical and mathematical models that simulate traffic flows and travel behavior across the entire country. The LMS takes into account factors such as population density, land use patterns, economic activity, and infrastructure capacity to predict traffic volumes and congestion levels on different roads and highways. The LMS is used to inform decisions about investments in transportation infrastructure, such as the construction of new highways, the expansion of public transit systems, and the implementation of congestion pricing policies.

The Dutch Regional Model (NRM) is a mesoscopic traffic model that is used for regional transportation planning and traffic management. It is based on a similar set

of statistical and mathematical models as the LMS, but it operates at a smaller scale and provides more detailed information about traffic patterns and congestion levels within specific regions of the country. The NRM is used to support decisions about local transportation policies, such as the optimization of public transit networks, the implementation of traffic management measures, and the planning of bike and pedestrian infrastructure.

Madrid models (Madrid Regional Gov, https://www.ciudades-abiertas.com/vocabularios/#Cat%C3%A1logoVocabularios)

The Madrid models are a group of vocabularies compiled in the site ciudades-abiertas.com to describe different aspects of a smart city. It describes elements like the public agenda, Census of premises and terraces, as well as their economic activities and associated opening licences, the register of inhabitants, public sharing bike system, public buses, Traffic, employment, budget, noise pollution, subsidies, etc.

3.1.2 Data Exchange API

Functional description

Data providers joining Data Spaces must be able to publish data resources at well defined endpoints knowing that data consumers, a priori unknown to them, will know how to retrieve and consume data through those endpoints. Data consumers, on the other hand, must know how data available through endpoints they discover can be consumed. This is a key principle which was observed in the design of the world wide web: content providers publish web pages on web servers (endpoints) knowing that web browsers will be able to connect to them and retrieve web pages whose content they can render and display to end users. It means that all participants in Data Spaces should 'speak the same language', which translates into adopting domain-agnostic common APIs and security schemas for data exchange (the way of constructing sentences) together with data models represented in data formats compatible with those APIs (the vocabulary used in constructed sentences). This requires the definition of data exchange APIs supporting:

1. Semantic interoperability, ensuring that the meaning of the data model within the context of a subject area is understood by the participating systems.
2. Behavioural interoperability, ensuring that the actual result obtained from usage of data exchange APIs achieves the expected outcome
3. Policy interoperability, i.e. interoperability while complying with the legal, organisational, and policy frameworks applicable to the participating systems.

The Data Space publishes Digital Twin data, very much like web servers publish html content on the world wide web. Data Spaces powered should enable near real-time (right-time) exchange of Digital Twin data which is fundamental in the design of innovative value chains demanding a very dynamic exchange of data

among participants. Just think about scenarios like a city managing traffic lights in streets close to a given train station in order to facilitate that travellers arriving and taking a taxi can leave faster to their destinations.

Baseline standards and industry body specifications

NGSI-LD (ETSI, https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.05.01_60/gs_CIM009v010501p.pdf)

The NGSI-LD API is domain-agnostic. Actually, many different systems have been developed using NGSI-LD in domains such as Smart Cities, Smart Manufacturing, Smart Energy, Smart Water, Smart AgriFood, Smart Ports, or Smart Health, to mention a few. This facilitates data sharing because each system participating in a Data Space will be publishing data that simply enriches a Digital Twin data representation of the world that the rest of systems connecting to the Data Space will know how to access. Systems participating into the Data Space don't know a priori what other systems may consume the data they publish (although they will be able to set up concrete terms and conditions for accessing/using data). NGSI-LD brings very simple and therefore easy to use operations for creating, updating and consuming context / Digital Twin data but also more powerful operations like sophisticated queries, including geo-queries, or the subscription to get notified on changes of Digital Twin entities.

LDES (SEMIC, https://github.com/SEMICeu/LinkedDataEventStreams)
Linked Data Event Stream (LDES) is a new data publishing approach which allows publishing any dataset as a collection of immutable objects defined in RDF. The focus of LDES is to allow users to replicate the history of a dataset and efficiently synchronise with its latest changes. It is used to maintain and open up reference datasets to foster interoperability by advocating the reuse of the identifiers for which they are the authoritative source. Compatible to other specifications such as Activity Streams, DCAT-AP, LDP or Shape Trees.

MQTT (OASIS, https://mqtt.org/mqtt-specification/)
MQTT is a lightweight, open, simple, and designed to be easy to implement Client Server publish/subscribe messaging transport protocol. It is usually used in the IoT domain as a mechanism for exchanging data from devices to data platforms and vice versa, like a broker of messages over TCP. By using a REST API you can read and write MQTT data via JSON over HTTP.

JSON-LD (PaySwarm, https://json-ld.org/)

JSON-LD is a lightweight Linked Data format easy for humans to read and write. It is based on JSON format and provides a way to help JSON data interoperate at Web-scale. JSON-LD is an ideal data format for programming environments, REST Web services, and unstructured databases such as Apache CouchDB and MongoDB.

Available implementations

CEF Context Broker (FIWARE, https://github.com/FIWARE/context.Orion-LD)
Orion-LD is a Context Broker and CEF building block for context data management which supports both the NGSI-LD and the NGSI-v2 APIs. It is currently a fork of the original Orion Context Broker extending support to add NGSI-LD and linked data concepts. Orion-LD follows the ETSI specification for NGSI-LD and has been tested to be a stable and fast NGSI-LD broker with close compliance to the version 1.3.1 of the NGSI-LD API specification.

Indian Urban Data Exchange (IUDX, https://nudm.mohua.gov.in/wp-content/uploads/2021/02/IUDX-Booklet-FINAL.pdf)
ETSI's CIM NGSI-LD has been adopted by Indian Urban Data Exchange as a national standard for Data Exchange and Open APIs.

3.1.3 Provenance and Traceability

Functional description

This BB includes components which provide the means for tracing and tracking in the process of data provision and data consumption/use. It provides the basis for a number of important functions, from identification of the provenance of data to audit-proof logging of NGSI-LD transactions.

While self sovereignty identity is about proof of the digital identity of an individual, trust is not about just giving an identity to an asset, but to provide the provenance of the asset, and in this case of the data.

Provenance frameworks aim to establish decentralised authenticity solutions by delegating and transferring trust. One prominent framework is the Coalition for Content Provenance and Authenticity (C2PA), which has released its provenance specifications. C2PA utilises DID (a type of SSI) and Verifiable Credentials from W3C. Despite still being in their early stages, the W3C Verifiable Credentials and the C2PA initiative seek to bridge a gap and create a coordinated effort to standardise technical specifications for provenance, which can effectively link content to its producers.

Timestamp is another mechanism to provide traceability over the data, as it does not allow values from the past, through the use of timescale data bases.

There is some research around the topic which proposes the use of blockchain to ensure the provenance of the data, for example in the context of BIM; or projects like Geontology from Ontochain.

Baseline standards and industry body specifications

ETSI-CIM (ETSI, https://www.etsi.org/deliver/etsi_gr/CIM/001_099/018/01.01.01_60/gr_CIM018v010101p.pdf)

This specification describes how the W3C Data Integrity Model uses digital signatures for providing the integrity of the URL; hashlinks and IPFS links for the integrity of the linked content; and how they can be used in a NGSI-LD framework. It also refers to JSON Web Signature, a compact signature format.

The specifications of provenance in NGSI-LD refer to guarantee that data values will not be altered through all its cycles, so that a data consumer, without further contact with the data provider, can be sure of the integrity. The preferred solution in both literature and industry, to the data integrity problem, is the implementation of a digital signature system.

DCAT-AP (EU, https://joinup.ec.europa.eu/release/dcat-ap-how-model-and-express-provenance)
DCAT-AP provides an optional property in the description of the datasets called "provenance" but it does not provide guidelines for describing instances. Therefore, only few national implementations are providing this information for their data.

Available implementations

Canis Major (FIWARE, https://github.com/FIWARE/CanisMajor)
For those data spaces with strong requirements on transparency and certification, FIWARE brings components like Canis Major that ease recording of transaction logs into different Distributed Ledgers / Blockchains. CanisMajor is a blockchain adaptor that supports persistence and verification of NGSI-LD Entity-Transactions(e.g. create/delete/update- operations) in blockchains.

3.2 Data Sovereignty and Trust

Data Spaces should bring technical means for guaranteeing that participants in a Data Space can trust each other and exercise sovereignty over data they share. This requires the adoption of common standards for managing the identity of participants, the verification of their truthfulness and the enforcement of policies agreed upon data access and usage control.

3.2.1 Identity Management

Functional description

The Identity Management (IM) building block allows identification, authentication, and authorisation of stakeholders operating in a data space. It ensures that organisations, individuals, machines, and other actors are provided with acknowledged identities, and that those identities can be authenticated and verified, including additional information provisioning, to be used by authorisation mechanisms to enable access and usage control. The IM building block can be implemented on the basis of readily available IM platforms that cover parts of the required functionality.

Creation of federated and trusted identities in data spaces can be supported by European regulations such as eIDAS.

Role and scope of the Identity Management building block is to provide authentication and authorisation of data space participants. An example of usage would be a user within an organisation registered with a data space. User would provide his/her log-in credentials to the IM module in order to gain access to the data of the data space in line with his/her role in the organisation.

Traditionally, identity management was treated in a centralised manner, but now with the emergence of data spaces, decentralised identity management is required and new standards are arising in this respect. Thus, we have decoupled the standards and specifications from centralised to decentralised identity.

Baseline standards and industry body specifications - centralised identity management

CEF eID (EU, https://ec.europa.eu/digital-building-blocks/wikis/display/CEFDIGITAL/eID)
The CEF eID program includes a set of specifications, standards, and guidelines that provide a common framework for interoperability between different national eID systems. It is not a single standard or protocol, but rather a set of technical and operational requirements that ensure that eID systems can work together seamlessly. Therefore, CEF eID can be considered both a standard and a specification. It defines a common set of technical requirements and specifications for eID systems, but also includes guidelines for their implementation and deployment.

eID is a set of services provided by the European Commission to enable the mutual recognition of national electronic identification schemes (eID) across borders. It allows European citizens to use their national eIDs when accessing online services from other European countries.

The eID block provides a suite of standards and services for electronic identification across the European Union. With this system in place, EU citizens can use their national electronic IDs to avail services across the EU without having to get a new eID if they travel or move to another EU country. The basic idea is to take the existing eID systems across member states and make them all work together in a seamless manner. The eIDAS regulation provides the technical standards that help achieve exactly this.

LDAP (IETF, https://ldap.com/)
The Lightweight Directory Access Protocol (LDAP) is an open, vendor-neutral, industry standard application protocol for accessing and maintaining distributed directory information services over an Internet Protocol (IP) network. It is both a

protocol and a standard, as it is defined by a series of technical specifications and Internet Engineering Task Force (IETF) standards documents.

LDAP was originally designed to be a lightweight alternative to the X.500 directory access protocol, which was used in large-scale enterprise directory services. LDAP provides a simpler and more flexible mechanism for accessing directory services, and it has become a popular protocol for managing user and group information in a wide variety of applications and systems. Because of this relationship, LDAP is sometimes called X.500-lite.

OAUTH2 (IETF, https://oauth.net/2/)
OAuth 2.0 is the industry-standard protocol for authorization. OAuth 2.0 focuses on client developer simplicity while providing specific authorization flows for web applications, desktop applications, mobile phones, and living room devices. This specification and its extensions are being developed within the IETF OAuth Working Group. OAuth 2.1 is an in-progress effort to consolidate OAuth 2.0 and many common extensions under a new name.

OpenID Connect (OpenID Foundation, https://openid.net/connect/)
OpenID Connect 1.0 is a simple identity layer on top of the OAuth 2.0 protocol. It allows Clients to verify the identity of the End-User based on the authentication performed by an Authorization Server, as well as to obtain basic profile information about the End-User in an interoperable and REST-like manner.

OpenID Connect allows clients of all types, including Web-based, mobile, and JavaScript clients, to request and receive information about authenticated sessions and end-users. The specification suite is extensible, allowing participants to use optional features such as encryption of identity data, discovery of OpenID Providers, and logout, when it makes sense for them.

SAML 2.0 (OASIS,
https://docs.oasis-open.org/security/saml/Post2.0/sstc-saml-tech-overview-2.0.html)
The OASIS Security Assertion Markup Language (SAML) standard defines an XML-based framework for describing and exchanging security information between on-line business partners. This security information is expressed in the form of portable SAML assertions that applications working across security domain boundaries can trust. The OASIS SAML standard defines precise syntax and rules for requesting, creating, communicating, and using these SAML assertions.

The OASIS Security Services Technical Committee (SSTC) develops and maintains the SAML standard. The SSTC has produced this technical overview to assist those wanting to know more about SAML by explaining the business use cases it addresses, the high-level technical components that make up a SAML deployment,

details of message exchanges for common use cases, and where to go for additional information.

X.500/X.509 (ITU-T, https://www.itu.int/rec/T-REC-X.500/en, https://www.itu.int/rec/T-REC-X.509)
X.500 and X.509 are both ITU-T standards that are related to the management and exchange of digital certificates and public key infrastructure (PKI). X.500 and X.509 are both important standards in the field of digital identity management, with X.500 providing directory services for managing identities and X.509 providing a standard for digital certificates used in PKI.

X.500 is a directory services standard that defines a hierarchical directory structure for storing and retrieving information about network resources and users. It is often used for managing user and device identities and for providing a central location for storing and retrieving public key certificates in a PKI. X.500 comprises of a series of computer networking standards covering electronic directory services: Directory Access Protocol (DAP), Directory System Protocol (DSP), Directory Information Shadowing Protocol (DISP), Directory Operational Bindings Management Protocol (DOP), Certificate Authority Subscription Protocol (CASP), Authorization Validation Management Protocol (AVMP), Trust Broker Protocol (TBP). The X.500 directory structure was also the basis for later models of directory structure such as LDAP.

X.509 is a standard for digital certificates that specify a format for electronic certificates that are used to verify the identity of individuals, organisations, and devices in a PKI. CASP and AVMP are related with the ITU Recommendation X.509 specification (X.509v3 digital certificates). X.509 certificates contain information such as the identity of the certificate holder, the digital signature of the certificate issuer, and the public key of the certificate holder.

Baseline standards and industry body specifications - decentralised identity management

W3C DID (W3C, https://www.w3.org/TR/did-core/)
Decentralised Identifiers (DID) is an official web standard. DIDs are cryptographic digital identifiers not tied to any central authority. They provide individuals and organisations with greater security and privacy, along with more control over their online information.

Instead of having your identity tied to an email address or a social media account controlled by a big tech company, you can have a DID that can be stored and transferred across different types of digital infrastructure, including blockchains.

DIDs can represent individuals, organisations, online communities, governments, IoT devices, or anything else that needs an online identity.

W3C Verifiable Credentials (W3C, https://www.w3.org/TR/vc-data-model/)
W3C Verifiable Credentials are a set of standards and protocols created by the World Wide Web Consortium (W3C) for creating, issuing, and verifying digital credentials in a decentralised and secure manner.

Verifiable Credentials are a way to represent and exchange information about a person, organisation, or thing in a digital format that can be cryptographically verified. They can be used for a variety of purposes, such as proving identity, qualifications, and permissions.

The W3C Verifiable Credentials specifications include a data model for representing credentials, a syntax for encoding and exchanging them, and a set of protocols for creating and verifying them. These specifications are designed to be interoperable with existing technologies, such as decentralised identifiers (DIDs) and blockchain-based systems.

Solid (MIT, https://solid.mit.edu/)
Solid is a specification and protocol for building decentralised identity management systems. It provides a set of standards and guidelines for how personal data can be stored, accessed, and shared in a secure and privacy-respecting way.

Solid is based on linked data, a set of best practices for publishing and interlinking structured data on the web. It defines how data should be structured, how access control should be implemented, and how data can be shared between different applications and platforms.

As a specification and protocol, Solid is not a complete solution in itself. It is meant to be implemented by developers and organisations to build their own decentralised identity management systems. Solid provides a foundation and a set of building blocks that can be customised and extended to meet specific needs and use cases.

Solid (https://solidproject.org) is a project created by Tim Bernes-Lee from MIT to allow storing personal data securely in decentralised data stores called Pods, kind of secure personal web servers for data.

Available implementations

The IM building block can be implemented on the basis of readily available IM platforms that cover parts of the required functionality. Integration of the IM building block with the eID building block of the Connecting Europe Facility (CEF), supporting electronic identification of users across Europe, would be particularly important. Following section shows an overview of existing business and government oriented identity management solutions on the market:

Business focused identity management implementations:

Solution name	Category	Description
Apache Syncope	open-source	Apache Syncope is an open-source identity management system that provides a centralized way to manage user identities, access rights, and provisioning across different systems and applications.
Anubis	open-source	Anubis is an identity management solution provided by FIWARE, an open-source platform that aims to facilitate the development of smart applications in various domains such as smart cities, agriculture, and healthcare. Anubis provides a set of identity management services that enable secure access to resources and data within FIWARE-based systems.
Keycloak	open-source	Keycloak is an open source identity and access management solution. Keycloak provides user federation, strong authentication, user management, fine-grained authorization, and more.
KeyRock	open-source	Keyrock is the FIWARE component responsible for Identity Management. Using Keyrock enables you to add OAuth2-based authentication and authorization security to your services and applications.
Okta Identity Cloud	open-source	Okta is an Identity and Access Management (IAM) solution that provides a cloud-based platform for managing user identities, authentication, and authorization.
Onfido	open-source	Onfido is an identity verification solution. Onfido integrates with several industry-standard protocols and specifications, including OAuth, OpenID Connect, and SAML, to enable seamless identity verification experiences for users.
Rekono	open-source	Rekono is a family of solutions and services for electronic identification, electronic signatures and other trust services, enabling secure and robust authentication, verification and authentication.
Ubisecure	open-source	Ubisecure is an identity management solution that provides unified identity management across all cloud, on-premises, legacy and modern applications.
Amazon Cognito	open-source	Amazon Cognito provides an identity store that scales to millions of users, supports social and enterprise identity federation, and offers advanced security features to protect your consumers and business.

Microsoft AD	open-source	Active Directory (AD) is a directory service developed by Microsoft for Windows domain networks. It is included in most Windows Server operating systems as a set of processes and services.

Public sector focused identity management implementations:

Solution name	**Category**	**Description**
ACM-IDM	open-source	ACM/IDM is an OpenID Provider for the Flemish Government.
CL@VE	closed-source	CL@VE is an identity management system used in Spain that provides secure access to online public services. The system is based on federated authentication and supports multiple authentication methods, such as eID, digital certificates, and one-time codes.
FAIRsFAIR	closed-source	The FAIRsFAIR project has developed an identity management system that supports the FAIR (Findable, Accessible, Interoperable, and Reusable) principles for research data management.
Shibboleth	open-source	Shibboleth is an open-source project that provides a standards-based way for organizations to manage access to online resources and services, using the Security Assertion Markup Language (SAML) protocol.
SI-PASS	closed-source	SI-PASS is a single point for verifying the identity of different users (citizens, business entities, public officials) and for electronically signing applications and other documents in Slovenia. SI-PASS is generally used in the context of the provision of individual electronic services (e.g. eGovernment, eVem).
SPID	closed-source	SPID identity management is the Italian government's digital identity system that provides citizens and businesses with a secure and reliable way to access digital services.
Tunnistamo	open-source	Tunnistamo is a service that authenticates users against several different identity sources, including Espoo, Vantaa and Helsinki internal accounts, Facebook, Google and Github. Tunnistamo also keeps a running session for the user, meaning the user won't need to login again to access another service using Tunnistamo.

3.2.2 Trusted exchange

Functional description

Any data space requires a Trust Anchor Framework and associated Decentralised Identity and Access Management Framework to enable the trusted operation of the system without requiring a central entity intermediating in all interactions among participants. This is required to ensure trust in the information published on the data space by providers, as well as to enable users and customers to access the dataspace portal services, manage their profile and seamless login into federated marketplaces where they can benefit from a tailored experience.

The Trust Anchor Framework defines and enforces a set of rules that different organisations agree to follow to deliver one or more of their services. This includes legislation, standards, guidance, and other rules. By following them, all services and organisations using the Trust Anchor Framework can use their digital identities and attributes in a consistent and trusted manner. This makes it easier for organisations and users to complete interactions and transactions or share information with other participants.

A decentralised Trust Anchor Framework is essential as a base mechanism to implement on top of it a Decentralised Identity and Access Management Framework to provide an efficient, scalable, and Decentralised IAM that participants can use not only to interact with the data space and marketplaces, but they can also adopt for interactions between themselves and their product/service consumers. The most appropriate and efficient implementation of a Decentralised IAM is one based on W3C Verifiable Credentials as the one described in the European Digital Identity Wallet Architecture and Reference Framework, which provides the specifications needed to develop an interoperable EUDI Wallet Solution based on common standards and practices and that was adopted by the eIDAS Expert Group on 26 January 2023.

For this reason and without loss of generalisation, we will assume that the supporting Decentralised IAM is based on Verifiable Credentials.

The Trust Anchor Framework addresses the following issues:

- **ID Binding**: How to verify that a given identifier corresponds to a valid legal identity of an entity in the real world?
- **Proof of participation**: How to verify that the entity is trusted because it is a subscribed participant in a given ecosystem (e.g., to check the trust of the Shared Catalogue of Product Specifications and of Product Offerings)?
- **Proof of Issuing Authority**: How to check that the credentials presented by a participant have been issued by another entity that can be considered a Trusted Issuer of that type of credentials? This enables the verifier to put the

right amount of trust in the facts attested by the Verifiable Credentials presented by a participant.

To enable transactional activity in the data space, the Decentralised Identity and Access Management Framework leverages on the above and provides an IAM system addressing additionally:

- **Identification**: How to verify that an identifier sent by a participant to another entity has been sent by the participant and not by an impostor that knows about the identifier? In addition, we need to cryptographically bind the identifier to the Verifiable Credentials sent by the participant so the facts attested in the credentials can be used for authentication and authorization.
- **Authorization**: How to use the attested facts in the Verifiable Credentials presented by a participant to perform advanced RBAC/ABAC access control and policy enforcement?

ID Binding

At the root of any trust framework there is the requirement to verify the identity of an entity in the real world and the assignment of some identifier that can be used later in representation of the real entity in the online processes. This association between an identifier (including some metadata) and the real identity of an entity is what we call *ID Binding*.

Please note that at this level, ID Binding states only who the entity is in the real world, not any additional properties that may be interesting for other purposes. For example, ID Binding establishes that the entity is a business incorporated in the EU, but it does not say what products it provides or the characteristics of the product, or the markets in which it operates, or in which data spaces it participates.

Many ecosystems assign a proprietary identifier to entities when they are onboarded in the ecosystem, creating silos of identifiers, and making very difficult the interoperability across ecosystems.

If eIDAS compliance is important, a good option is to rely on identifiers already used in digital certificates issued by the Trust Service Providers (TSPs) authorised by the relevant European laws. The combination of digital certificates issued by TSPs, and Verifiable Credentials contributes to the legal validity and interoperability of the data-related transactions in the European Union facilitating the validation of eSignatures, eSeals, and more. Essentially, Verifiable Credentials and Presentations used in the ecosystem will be signed using digital certificates.

Proof of Participation

When verifying a Verifiable Credential/Presentation, we must address the following:

- How do we know that the issuer of the Verifiable Credential is a participant in the concrete ecosystem (e.g., a given Data Space) where we are also participants?

- How do we know that the subject of the Verifiable Credential is a participant in the concrete ecosystem (e.g., a given Data Space) where we are also participants?

Proof of Issuing Authority

Given that anyone can have access to the technology needed to create Verifiable Credentials and anybody can issue credentials and digitally sign them with their eIDAS digital certificate, the problem is how a verifier knows that the Verifiable Credentials received from the subject have been issued by an entity which is entitled or authorised to issue that type of credential.

The primary mechanism to solve this problem is the use of Trusted Issuer Lists (there may be several lists, one per domain or type of credential). A Trusted Issuers List is a register of trusted public entities which can issue Verifiable Credentials belonging to a given domain or of a given type. It is assumed that an entity must be first in the Trusted Participant List before it appears in the Trusted Issuers List. This list includes the identifiers, public keys for verification of signatures and their accreditations in the form of Verifiable Credentials/Presentations from third parties, enabling the entity to issue credentials of a given type. All information in the registry is validated and signed by trusted legal entities of the corresponding domain (Conformity Assessment Bodies and third-party auditors).

Additionally and in compliance with the upcoming eIDAS2, the specific status of a role in the ecosystem shall need to be verified in a trustworthy manner. Such roles are:

- Wallet Providers
- Person Identification Data Providers
- Electronic Attestation of Attributes (EAA) providers, also known as Issuers
- Relying Parties
- Catalogues of attributes and schemes for the attestations of attribute providers

All these status are verified by participants in the ecosystem via Trusted Lists which jointly compose the Trust Anchor Framework.

Baseline standards and industry body specifications

EUDI (EU, https://digital-strategy.ec.europa.eu/en/library/european-digital-identity-wallet-architecture-and-reference-framework)

The most relevant standard is the mentioned European Digital Identity Wallet Architecture and Reference Framework, which defines the reference framework for the future digital identity wallet in the context of the new version of the eIDAS regulation, including the required Trusted Lists.

ESBI (EU, https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/Home)

The most important initiative is EBSI (European Blockchain Services Infrastructure). EBSI was born in 2018 when 29 countries (all EU member states, Norway and Liechtenstein) and the EU Commission joined forces to create the European Blockchain Partnership (EBP). EBP's vision is to leverage blockchain to create cross-border services for public administrations, businesses, citizens and their ecosystems to verify information and make services trustworthy.

Available implementations

European Blockchain (EBSI, https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/EBSI+Verifiable+Credentials)
The most relevant implementation is the one from EBSI, which is in pre-production state and with a target date for production in 2023. It proposes a new decentralised paradigm for the web called Web3, where users can own and control their data. They have created an architecture for decentralised trust, making information verifiable and empowering EU citizens.

i4Trust (Consortium of partners, https://github.com/i4Trust/building-blocks)
An implementation of the trust framework in data spaces was developed in the project i4Trust, relying on iShare and implementing the EBSI APIs.

3.2.3 Access and usage control

Functional description

Access and usage control guarantees enforcement of data access and usage policies defined as part of the terms and conditions established when data resources or services are published or negotiated between providers and consumers. Therefore, data spaces must provide means for guaranteeing organisations joining data spaces that they will be able to exercise sovereignty on their data. That requires the definition of a framework using common building blocks for access and usage control, based on mature security standards that will be used by all participants in the data space.

Such a framework needs to define an architecture for components dealing with the access and usage management. Commonly, access to data services needs to be protected by entities which enforce authorization decisions, e.g., Policy Enforcement Points (PEPs) in combination with Policy Decision Points (PDPs). PEPs intercept requests that are sent to a data service endpoint and enforce access according to decisions made by PDPs. These decisions are following certain rules based on access policies issued to the service consumers which are administered and managed in certain Policy Administration Points (PAPs) and Policy Management Points (PMPs). Depending on the use case and required complexities, it should be possible to define access policies for specific data service points on simple role-based models (RBAC) up to more complex attribute-based access models

(ABAC). PEP, PDP, PAP and PMP are terms and roles taken from eXtensible Access Control Markup Language (XACML) from the OASIS consortium.

Participants in the data space can be organisations themselves, and users or human individuals connected to certain organisations. Both, organisations and users, might participate in the data exchange within the data space. Therefore the framework should allow the delegation of access rights by organisations to other parties, which can be organisations and users, authorising them to act on others behalf.

The framework can be used by participants not just to interact with the data space but they can adopt it and use it for peer-to-peer interactions between participants in the ecosystem without the involvement of central entities (except for initial onboarding and certification processes).

In the end, a framework for access control in a data space will rely on certain policy languages, API specifications for the exchange of policies (e.g., with PAPs and PMPs or with service providers) and rules for processing of policies.

Finally, the framework should keep in mind the legal and business aspects of the fact that data consumers will abide by the terms of data usage that is defined by means of data licence.

Baseline standards and industry body specifications

Rego (Styra, https://www.openpolicyagent.org/docs/latest/policy-language/)
A Policy Definition Language is required to define and agree on access and usage policies. The defined and agreed policies can be used directly or translated into an executable language, e.g., Rego. Two such policy languages are ODRL and XACML.

XACML Policy Definition Language (OASIS, https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml)
It allows defining fine-grained and attribute-based access control policies and has been standardised by the Technical Committee of the OASIS consortium. It defines an access control policy language, a processing model for requests and responses and even an architecture based on PEPs, PDPs and PAPs.

Open Policy Agent (Styra, https://www.openpolicyagent.org/docs/latest/)
The Open Policy Agent (OPA) is an open-source, general-purpose policy engine which will be taken into consideration. It provides a high-level declarative language to specify policies and offload policy decision-making into other components.

W3C ODRL (W3C, https://www.w3.org/TR/odrl-model/)
The Open Digital Rights Language (ODRL) is a policy expression language that provides a flexible and interoperable information model, vocabulary, and encoding mechanisms for representing statements about the usage of content and services. The ODRL Information Model describes the underlying concepts, entities, and

relationships that form the foundational basis for the semantics of the ODRL policies.

Policies are used to represent permitted and prohibited actions over a certain asset, as well as the obligations required to be met by stakeholders. In addition, policies may be limited by constraints (e.g., temporal or spatial constraints) and duties (e.g. payments) may be imposed on permissions.

It is an interoperable standard for the negotiation and acceptance of Access and Usage Policies and is one of the W3C recommendations for policy languages. In general, the ODRL is defined by a model and a common vocabulary in an abstract manner. ODRL can be encoded in different formats, like XML or JSON.

W3C WAC (W3C, https://www.w3.org/wiki/WebAccessControl)
Web Access Control (WAC) is a decentralised cross-domain access control system providing a way for Linked Data systems to set authorization conditions on HTTP resources using the Access Control List (ACL) model. WAC has several key features: the resources are identified by URLs, it is declarative, users and groups are also identified by URLs, it is cross-domain. WebAC enforces access control based on the Access Control List (ACL) RDF resource associated with the requested resource.

Available implementations

i4Trust (Consortium of partners, https://github.com/i4Trust/building-blocks)
An implementation of an access and usage control framework in data spaces was developed in the project i4Trust.

The policy language is based on a JSON port of XACML and allows defining attribute-based policies and supports the delegation of access rights. Access policies, precisely called delegation evidences in the framework, can be requested at Authorisation Registries via REST APIs, which represent the PAP and PMP.

The FIWARE Identity Provider Keyrock implements PAP and PMP functionalities for standard XACML and the mentioned XACML JSON port functions.

The i4Trust framework relies on exchange of context data via NGSI-LD. Therefore, rules have been developed to match certain NGSI-LD requests on the attribute-based access policies, allowing to restrict access to context entities via the NGSI-LD API by parameters like the type of the operation or entity IDs, types and attributes.

Parts of the i4Trust framework are also integrated into the technical convergence of the DSBA.

3.3 Data Value creation

This building block facilitates the dynamic enlargement of data spaces with more stakeholders, data resources, and data-processing/analytics services (such as

big-data analysis services, machine learning services, or services based on statistical processing models for different business functions). It comprises capabilities for publishing data resources following the broadly accepted DCAT (Data Catalogue Vocabulary) standards, and for harvesting data from existing open-data publication platforms.

3.3.1 Metadata and discovery services

Functional description

Metadata can simply be defined as data about other data. Metadata is arguably more important than the data itself if value is to be created by exchanging and manipulating data. It can help both human and non-human actors to discover, analyse, combine and use data sources in any way. We quickly describe the most important types of metadata and their relation to data spaces here:

1. **Structural metadata**
   This is data describing the structure of a data source. For instance, it can describe how a data source interacts with others and which types of patterns and internal hierarchies there are within the data source. Perhaps the easiest way to describe it is by comparing this to a non-digital data source: a book. The structural metadata in this case would be the table of contents, an index, the page numbering, etc.
2. **Descriptive metadata**
   This is metadata that describes the content of the data source, by giving some information on the concepts contained within. For instance, one can think of labels, tags and classifications. Of course, unambiguously describing concepts is a whole field of research in itself, especially when the data is to be discovered by non-human actors (see further). In this latter case we will often turn to linked data and the semantic web.
3. **Preservation metadata**
   Preservation metadata is especially of value for non-digital data sources, such as physical archives. It describes the steps taken to preserve data, for instance the packaging, environmental parameters, and so on. For digital data sources one could consider the storage period as such a type of metadata.
4. **Administrative metadata**
   Administrative metadata is all data describing how data is to be accessed and used. It allows the definition of a governance structure around a data source. Examples are the copyright notices, other licence agreements, access control, and so on. This is described in more detail in the previous section, 4.2 Access and usage control.
5. **Provenance metadata**
   Provenance data describes the origins of the data contained within a data

source. It helps to define the data lifecycle, any transformations the data may have undergone, and potentially the trustworthiness of the data.

6. **Definitional metadata**
   This type of metadata allows defining the meaning of the data contained within a data source. This may include definitions, thesauri, and in the case of linked data, the ontologies to which the data source adheres. This is what we call *semantic* definition. This type of metadata may also be also used to *schematically* describe a data source, for instance, by defining the structure of a database.

Metadata management is a key aspect of data governance, and is essential to allow for data discovery. An important trend to take into account is Automated Metadata Extraction (AME), which uses AI to tag datasets automatically, based on their contents. However, AME does not always yield good results, especially not for administrative or provenance metadata.

Data discovery is the term used for the business process of discovering new data points, patterns, outliers and insights from existing data. This may be done by combining different data sources, trend analysis, or even visual interpretation. Data discovery is arguably the most effective way to create value from data, and is one of the key drivers for any business intelligence (BI) implementation.

When considering data spaces, data discovery can not only be applied to the data sources of a single actor, but to all data sources in the data space to which an actor has access. Therefore, the potential of value creation within a data space can be considered much higher.

Baseline standards and industry body specifications

The importance of metadata management has been understood for a long time, and thus, many different initiatives have developed different approaches to the subject over time. The number of existing metadata catalogues, thesauri and taxonomies is staggering, and would be impossible to cover in this deliverable. Moreover, many of these have been developed for a specific domain or purpose. Since the data space for smart and sustainable cities and communities has to encompass many different domains and applications, we will limit ourselves to initiatives that can be considered as overarching and crossing multiple domains.

An important approach to metadata management and data discovery is Linked Data[9]**.** This approach that has known a growing interest over the last decade is Linked Data or the creation of a Semantic Web. In this approach, every datapoint is described as a fact, formalised as a "triple" consisting of a subject, predicate and object (consider for instance: "a city is geographically located within a country"). Each of the three parts of this triple is defined by a URI, making them unique by definition. All these facts can consequently and unambiguously be connected

[9] https://www.scientificamerican.com/author/tim-berners-lee-james-hendler-and-ora-lassila/

together, forming a “web of data” rather than a “web of documents”. This is a very basic definition of linked data, but more details would lead us far beyond the scope of this deliverable.

Next we describe a number of metadata standards that are widely applied. This overview is limited to standards that specifically describe metadata, not the data itself. For the data standards themselves, please refer to section 3.3.1 on data models. It is important to note that many of the standards below have both a “linked” and a “non-linked” variant.

Dublin Core (ISO, https://www.dublincore.org/)
Dublin Core is one of the most common metadata sets used in industry and academia alike. It has been standardised by the ISO as ISO 15836, and is a part of the Dublin Core Metadata Initiative, DCMI. At its core, it lists 15 of the most basic metadata elements (such as “Title”, “Author”, “Subject”, …) which are used in almost every application of data.

DCAT (W3C, https://www.w3.org/TR/vocab-dcat-2/)
DCAT is essentially a linked data ontology that allows for the description of data catalogues. It provides a means to list all the datasets contained within one source (a data portal, a marketplace, or a data space) according to linked data best practices. This way, all essential properties of a data set can be described using the full power of the semantic web. It is an international standard developed by the W3C, but is potentially best known for its use within the European Commission’s ISA² programme, now known as Interoperable Europe. To allow more application-specific use, it has been extended as “DCAT-AP”, an “application profile” which has been instrumental in harmonising many European open data portals and has driven discoverability of all datasets contained within through the European Open Data Portal. This application profile has also known a number of national extensions (which allow further specialisation without breaking European interoperability) such as DCAT-AP-VL and DCAT-AP_IT. There is also a geographical extension which is called GeoDCAT-AP.

INSPIRE (EU, https://inspire.ec.europa.eu/data-specifications/2892)
INSPIRE is a metadata standard that has evolved from the geographic community and that has been further developed and endorsed by the European Commission, eventually resulting in the creation of the INSPIRE Directive. It enables the collation of geographical data sources through the definition of 34 geospatial themes, greatly improving interoperability of geographical data across Member States, effectively creating the European Spatial Data Infrastructure (SDI).

ISO 19115-1:2014 (ISO, https://www.iso.org/standard/53798.html)
ISO 19115-1:2014 defines the schema required for describing geographic information and services by means of metadata. It provides information about the identification, the extent, the quality, the spatial and temporal aspects, the content,

the spatial reference, the portrayal, distribution, and other properties of digital geographic data and services.

Available implementations

Typically metadata management systems are conceived as "data catalogues". Many different such implementations exist. Some of the most frequently used are Colibra, Apache Atlas and Informatica. When we consider open data, there are specific systems optimised for publicly publishing open data, called open data management systems, which are described in more detail in the next section on publication and marketplace services.

Popular Data Discovery tools include Tableau, Qlik Sense and Atlan.

One great example of an implementation of a data catalogue can be found in Flanders, the "*Datavindplaats*" (literally "data finding place") where a lot of effort has gone into streamlining both geographical and non-geographical metadata, leading to the creation of a new metadata standard: GeoDCAT-AP.

3.3.2 Publication and marketplace services

Functional description

Loose coupling of participants is a fundamental principle in data spaces. Data providers and consumers do not necessarily know about each other. Therefore, it becomes essential to incorporate building blocks enabling the management of data resources as true assets with a business value. Assets which can be published, discovered and, eventually, traded. This way boosting the creation of multi-side markets where innovative services can be created.

This requires the adoption of common mechanisms enabling the description of services for accessing data or linked to applications processing data, the description of offerings associated with those services, the publication and discovery of both services and service offerings, and the management of all the necessary steps, including clearing, payment and billing functions, supporting the lifecycle of contracts that are established when a given participant acquires the rights to use a service, according to certain service offering.

Besides that, such building blocks should enable the publication of data resources linked to data assets and the federation of existing Open Data Management Systems (ODMS) based on different technologies providing a unique access point to search and discover open datasets coming from heterogeneous sources.

Publication and Marketplace Services building block is used to support the offering of data resources and services under defined terms and conditions, marketplaces must be established. This building block supports publication of these offerings, management of processes linked to the creation and monitoring of smart contracts

(which clearly describe the rights and obligations for data and service usage), and access to data and services.

Role and scope of the Publication and Marketplace Services building block is to provide a directory of the various data assets for dynamic access and discovery as well as management of established contracts. An example of usage would be a data space user that queries the data resources publication platform on specific data assets (e.g. based on content, theme, industry, etc.). Upon selecting the dataset she/he wants to access, she/he receives a link (e.g. an URL) to the dataset chosen.

Baseline standards and industry body specifications

OpenAPI (TMForum, https://projects.tmforum.org/wiki/display/API/Open+API+Table)
The TMForum Open APIs, provided by TMForum which is a global industry association in the telecommunications industry, enable a seamless connectivity, interoperability and portability across complex ecosystem services. It consists of a suite of more than 60 APIs that are widely adopted by the industry and allow to enable services to be managed end-to-end throughout their lifecycle.

TMForum provides, among others, the following open APIs that are useful in the implementation of marketplace services:

There are different ODMS available. Just to name a few of them:

- CKAN: The Comprehensive Knowledge Archive Network (CKAN) is an open-source open data portal for the storage and distribution of open data.
- DKAN: DKAN is a Drupal-based open data portal based on CKAN and is a community-driven, free and open source open data platform
- Socrata: Socrata, meanwhile acquired by Tyler Technologies, is an Open Data Network to make government data discoverable, usable, and actionable
- Context Broker: Also the FIWARE Context Broker via NGSI v2 and NGSI-LD can be used as API for open data management

ICT Innovation Network reference architecture (Slovenia, https://ikthm-en.gzs.si/)
The ICT horizontal network is governed by the Slovenian Strategic Research Innovation Partnership for Smart Cities and Communities (SRIP SCC) which follows the vision of the Government of the Republic of Slovenia of making Slovenia a green reference country in digital Europe in national Smart specialisation strategy. The referentiality is realised through an exemplary model of collaboration between the business sector, science and the government in the introduction of modern digital solutions into people's lives.

ICT horizontal network designed an ICT Innovation Network reference architecture, based on the FIWARE Smart City reference architecture. It expands the former in the field of business interoperability in a way that it is more focused in the business

ecosystem/marketplace integration than into data exchange capabilities. It is also aligned with the latest data spaces architecture building blocks.

The core technological building blocks used in the reference architecture include Identity Management building block (using OAUTH2 standard), Data Models and Formats building block (using CEF Context Broker in connection with Smart Data Models), Access, and Usage Control Policies. Regarding the Publication and marketplace services data space building blocks the reference architecture uses the TMF Open API interfaces and data models for data exchange between different participants: Party Management API, Catalog Management API, Customer Management API, Product Ordering Management API, Shopping Cart Management API to name a few. It also supports extended data spaces building blocks such as Data Analytics Engine, Data Visualization and Workflow Management Engine.

The goal of the ICT Innovation Network reference architecture is that future implementations become one-stop-shop for city services that allow stand-alone vertical solutions straight-forward integration into the platform's system services, while providing users with useful services in a simple and intuitive presentation interface.

Available implementations

Business API Ecosystem (FIWARE, https://business-api-ecosystem.readthedocs.io/en/latest/)
The FIWARE Business API Ecosystem (BAE) is a joint component made up of the FIWARE Business Framework and a set of APIs (and its reference implementations) provided by TMForum. It provides sellers the means for managing, publishing, and generating revenue of their products, apps, data, and services. The BAE enables the monetization of different kinds of assets (both digital and physical) across the whole service life cycle, from offer creation through to charging, accounting and revenue settlement and sharing. Its components enable creation of Marketplace services which participants in data spaces can rely on for publishing their offerings around data assets they own. Different types of data assets can be defined via plugins that can be installed in the BAE, taking care of data validation, provider permissions and service activation.

The BAE is incorporating the following TMForum APIs in order to implement marketplace services:

- Catalog Management API
- Product Ordering Management API
- Product Inventory Management API
- Party Management API
- Customer Management API
- Billing Management API
- Usage Management API

FIWARE also comprise components for publication of data resources linked to data assets around which offerings are managed through the BAE.

For this purpose, the Idra publication platform as well as extensions to the CKAN open data platform have been developed by FIWARE.

CKAN Extensions (FIWARE, https://fiware-ckan-extensions.readthedocs.io/en/latest/)
The CKAN extensions support enhanced data management capabilities and integration with FIWARE technologies including NGSI. In particular, data publication and discovery features provided by CKAN have been enhanced with the following features:

- Right-time (near real time) time data publication - due to this extension, CKAN is not limited to list data resources linked to static files as part of its catalogue but also data resources linked to NGSI-LD requests served by Context Broker components deployed in a data space. This brings the ability to discover data resources relying on DCAT capabilities publication platforms support,
- Identity Management, Authentication and Access Control functions based on Keyrock components (Keyrock is the FIWARE Identity Manager) - therefore supporting OpenId Connect, OAuth2 and XACML standards adopted at overall data space level.
- Publication of priced data resources - an extension making it possible to mark data resources listed as part of the catalogue as linked to offerings visible in the Data Marketplace. Users can therefore click on those data resources and navigate to the Marketplace to proceed with the acquisition of access rights
- Enhanced Data Visualization - an extension allowing the creation of rich visualisations for dataset resources by embedding WireCloud dashboards as resource views.

IDRA (Engineering Ingegnieria Informatica, https://idra.readthedocs.io/en/latest/)
Idra is a web application able to federate existing Open Data Management Systems. It unifies representation of collected open datasets, thanks to the adoption of international standards (DCAT-AP) and provides a set of RESTful APIs to be used by third party applications. Idra supports natively ODMS based on CKAN, DKAN, Socrata, Orion Context Broker (NGSI v2, NGSI-LD) and many other technologies: Idra provides also a set of APIs to federate ODMS not natively supported. In addition, it is possible to federate generic open data portals, that don't expose API, using the web scraping functionality or providing a dump file of the datasets in DCAT-AP format. Furthermore, Idra provides a SPARQL endpoint in order to perform queries on 5 stars RDF linked open data collected from federated ODMS and allows to easily create charts based on federated open datasets (through DataIEt-Ecosystem Provider DEE).

i4Trust (Consortium of partners, https://github.com/i4Trust/building-blocks)
i4Trust supports effective data exchange, publication and trading using a new way of working, partnering and creating new data-driven businesses.

Other features of the project are: trustworthiness (thanks to the unified framework for identification and the robust legal frameworks that i4Trust brings, participants can trust each other), sovereignty (i4Trust brings the means for enforcing the data access and usage policies participants want to define and gives them power to be sovereign of their data), effectiveness (designed for the exchange of data among Smart Solutions, i4Trust brings a standard data exchange API and data models guaranteeing participants to effectively share data), openness (i4Trust is opened, based on open-standard and implemented as Open Source, allowing participants to avoid vendor lock-in thus protecting their investment and reducing costs), cross-domain (i4Trust unleashes the potential of data sharing among different participants in multiple domains, allowing them to define cross-domain data value chains).

i3-MARKET (Consortium of partners, https://www.i3-market.eu/)
The H2020 i3-MARKET project aims to promote data market economy by providing support tools and avoiding to create another new marketplace but implementing a solution in the form of a backplane set of tools introduced as a framework of solutions that allow other data marketplaces and Data Spaces to expand their market functions, facilitating the registration and discovery of data assets and supporting the trading and sharing of data assets among providers, consumers, and owners for a better data sharing and trading processes.

The i3-MARKET platform is designed to enable secure and privacy-preserving data sharing across Data Spaces and marketplaces by deploying a backplane across operational data marketplaces. The i3-MARKET Backplane, on the one hand, can be seen as a set of tools that can be deployed and integrated as backend technologies in current running marketplaces facilitating and allowing to add the missing functionalities that current marketplaces lack, and, on the other hand, i3-MARKET acts as baseline technologies for stand- alone reference implementation(s) that facilitates the starting point for a modern data sharing economy. In other words, the i3-MARKET Backplane provides the tools for setting up the foundations of a data marketplace ecosystem.

Open banking (EU, https://www.openbankproject.com)
The Open Bank Project (OBP) platform is a middleware solution that allows financial institutions to easily create, secure, distribute, and monetise APIs. It comes with a catalogue of over 550 pre-built APIs available for immediate use. It defines accounts, branches, ATM, transactions, counterparties and payments and it incorporates an API for managing all these data models.

ICT Innovation Network reference implementations (Slovenia, https://ikthm-en.gzs.si/)
Currently, there are two implementations of the ICT Innovation Network reference architecture available: project iPOT[10] in the field of integrated smart mobility (cofunded by the European Regional Development Fund) and the project LokalnoGOR in the field of local food self-sustainability (cofunded by the European Agricultural Fund for Rural Development).

The **iPOT project** aims to create and integrate into a demonstration environment an integrated next-generation mobility platform, which will primarily enable the collection and processing of large amounts of data in real time, as a globally unique product. The project covers the area of mobility, transport, logistics, with the key objective of increasing the mobility of people and goods by enabling reliable, flexible, accessible, safer and greener urban and peri-urban services.

Architecture of the smart mobility platform is aligned with the ICT Innovation network reference architecture by using the following building blocks:

- Identity Management (Keycloak component using OAUTH2 standard),
- Usage Control Policies,
- Data Models and Formats (CEF Context Broker in connection with Smart Data Models),
- Publication and marketing (micro-service components compatible with the TMF Open API standard): Party Management API, Product Ordering Management API
- extended building blocks: Data Analytics Engine, Data Visualization and Workflow Management Engine.

The **LokalnoGOR project** aims to: 1) educate and support public educational staff, children, parents and persons with special needs, 2) empower food producers, 3) improve the agricultural sector through modern marketing, logistics and cooperatives to increase the supply of locally produced food, and 4) increase awareness of the benefits of locally produced food in local households. Project among other objectives, introduces a technology platform that supports local food self-sufficiency by connecting farmers and other local food providers with retail customers, powered by the ICT Innovation network reference architecture. Platform promotes local food supplier brands and enables viewing, ordering and delivery of fresh local food. The platform represents good practice in the field of food self-sufficiency: it reduces the carbon footprint, strengthens the local economy and community and revitalises town centres.

Architecture of the local self-sustainability platform is aligned with the ICT Innovation network reference architecture by using the following building blocks:

[10] www.ipot.si

- Identity Management (Keycloak component using OAUTH2 standard),
- Usage Control Policies,
- Data Models and Formats (CEF Context Broker in connection with Smart Data Models),
- Publication and marketing (micro-service components compatible with the TMF Open API standard): Party Management API, Catalog Management API, Customer Management API, Product Ordering Management API, Shopping Cart Management API,
- extended building blocks: Workflow Management Engine.

Open platform architecture design is compatible with the guidelines in the technical convergence document from DSBA and allows straight-forward on-boarding of the new participants in the ecosystem and also enables constructions of the new value chains with other business domains (e.g. tourism with Smart Destination)[11].

3.3.3 Data usage accounting

Functional description

This building block provides the basis for accounting access to and/or usage of data by different users. This in turn is supportive of important functions for clearing, payment, and billing (including data-sharing transactions without involvement of data marketplaces).

It has been seen on many occasions that Smart City platforms have been acting as a Data platforms for even the non-open data and enabling Data Economy. This document considers the only specifications which have a standard API defined.

The functional description as per TM Forum[12] as below

> *The Account API provides a standardised mechanism for the management of billing and settlement accounts, as well as for financial accounting (account receivable) either in B2B or B2B2C contexts.*
>
> *It allows creation, update and retrieval of account information either in a B2B2C relationship context (creation of mass market customer billing account within a "Billing on Behalf of" process for example) or in a B2B context (creation of a billing/settlement account for a partner or B2B customer).*
>
> *It also allows creation and query of bill items allowing partners or B2B customer to check their invoice*

The API could cover the end to end functionality of billing, accounting and payments as described in TM Forum API specifications. However, some of the

[11] www.optifarm.net
[12] https://www.tmforum.org/

implementations like IUDX in India, provide the data metering API based on ETSI's NGSI-LD API, enabling users to retrieve metering and audit data. Furthermore it offers ownership of user's Billing and Accounting capabilities or seamlessly integrates with them.

Baseline standards and industry body specifications

TM Forum Accounting API (TMForum, https://tmf-open-api-table-documents.s3.eu-west-1.amazonaws.com/OpenApiTable/4.0.0/user_guides/TMF666_Account_Management_API_REST_Specification_R19.0.0.pdf)
Provides a standardised mechanism for the management of billing and settlement accounts, as well as for financial accounting (account receivable) either in B2B or B2B2C contexts. The specification defines the standard APIs and models and relationships between different stakeholders.

IUDX Metering & Audit API (IUDX, https://rs.iudx.org.in/apis#tag/Metering)
In order to understand the usage of a resource and APIs, the Resource Access Layer, with the help of a metering and auditing layer, integrates with an immutable database for storing auditing information. This information helps IUDX Administrators for planning of new APIs based on usage, Data Exchange Consumers for usage planning and Data Providers for enabling flexible policies, understanding data consumption and usage. The API specification is at IUDX Metering & Audit API specification in ETSI's NGSI-LD standard.

Available implementations

Business API Ecosystem (FIWARE, https://business-api-ecosystem.readthedocs.io/en/latest/)
The FIWARE Business API Ecosystem (BAE) is a joint component made up of the FIWARE Business Framework and a set of APIs (and its reference implementations) provided by TMForum. It provides sellers the means for managing, publishing, and generating revenue of their products, apps, data, and services. The BAE enables the monetization of different kinds of assets (both digital and physical) across the whole service life cycle, from offer creation through to charging, accounting and revenue settlement and sharing. Its components enable creation of Marketplace services which participants in data spaces can rely on for publishing their offerings around data assets they own. Different types of data assets can be defined via plugins that can be installed in the BAE, taking care of data validation, provider permissions and service activation.

The BAE is incorporating the following TMForum APIs in order to implement marketplace services:

- Catalog Management API
- Product Ordering Management API

- Product Inventory Management API
- Party Management API
- Customer Management API
- Billing Management API
- Usage Management API

FIWARE also comprise components for publication of data resources linked to data assets around which offerings are managed through the BAE.

For this purpose, the Idra publication platform as well as extensions to the CKAN open data platform have been developed by FIWARE as described in Publication and Marketplace BB.

IUDX Metering API (IUDX, https://github.com/datakaveri/iudx-resource-server)

Implemented by the IUDX Resource Server (NGSI-LD Context Broker), the API's Resource Server is IUDX's data store allowing for publication, subscription and discovery of data. For search and discovery, the API allows users to search through temporal, geo-based and attribute queries. For publication and subscription, the API allows users to use AMQP streaming protocol over TLS. IUDX enables Providers of data sources to publish data as per the IUDX data descriptor. Furthermore, the API enables Consumers of data sources to search and query for data using HTTPs APIs. It enables Streaming Consumer a.k.a [Subscribers] of data sources to stream data using AMQP streaming protocol over.

3.4 Data Space Governance

The rationale of developing appropriate governance lies in the novelty of data spaces and in the ambition to upscale them. As such, the governance helps stakeholders to understand roles, responsibilities and value proposition of data spaces in the smart and sustainable cities and communities context.

This pillar from OpenDEI taxonomy is not a technical set of BBs, but a compilation of agreements, guidelines and recommendations that can be followed to build the non technical dimensions of a data space. Thus, the description of these elements in the Catalogue should not follow the same structure as the technical BBs (eg.: functional description, baseline standards and industrial body specifications, and available implementations).

The governance of the data space will enable a fair, transparent and trustworthy sharing and use of data in line with European values and with existing EU data-related legislations and provisions (e.g. General Data Protection Regulation, Free Flow of Non-Personal Data Regulation, ePrivacy Directive, Open Data Directive, etc). For more information, refer to Appendix I: 'Relevant EU regulations and legislations'.

As stated in the mapping of MIMs to Data Spaces Building Blocks, business (3.4.1), organisational (3.4.2) and operational (3.4.3) aspects are mainly covered within the description of MIM3 (contracts), MIM4 (Trust), and MIM6 (Security), providing guidance towards identifying pivotal interoperability points for governing data spaces.

3.4.1 Business agreements

Functional description

The business aspects of these agreements define the contractual terms in which one or more parties will cooperate. This includes the financial aspects, which detail how the value created is distributed among the parties, and which fees will apply to data exchange. They will also define how the cost of setting up and maintaining the data space are distributed.

Furthermore they will contain data sharing agreements, public procurement of data/data purchasing agreements, and Service Level Agreements (SLAs). These agreements specify the responsibilities of each of the parties in maintaining a specific level of service, including dispute resolution measures, such as fines that should be paid when a certain service level is breached. For instance, if a certain uptime cannot be guaranteed, or when a data breach occurs.

Finally, business agreements will also define how each of the parties conform to existing legislation, such as GDPR.

Baseline resources

- Sitra, Rulebook for a fair data economy: · https://www.sitra.fi/en/publications/rulebook-for-a-fair-data-economy/#preface-and-templates
- Contract for Data Collaborations (C4DC): library of examples of data sharing agreements between different types of stakeholders: https://contractsfordatacollaboration.org/library/
- Open Data Institute, Designing data sharing agreements, a checklist : https://gatesopenresearch.org/documents/2-44

3.4.2 Organisational and operational agreements

Functional description

Setting up a data space also requires a number of organisational and operational measures to be put in place. On one hand, they should prevent the misuse of data by ensuring data sovereignty, trust and security. On the other, they should tackle questions like what to do when new parties want to be involved in the dataspace. In many cases, it makes sense to organise the data space as a marketplace, in which a more dynamic exchange can happen. This also includes data management

mechanisms, data quality assurance, data strategy, dedicated data officer to manage contracts for data providers/consumers. Additional factors in this context are cultural enablers like data literacy, skills, and training to be able to work with data.

Baseline resources

- DAMA DMBoK: https://www.dama.org/cpages/home
- OVSDB (Open vSwitch Database Management Protocol): https://docs.openvswitch.org/en/latest/ref/ovsdb.7/
- OSLO (Open Standards for Linked Organisations): https://joinup.ec.europa.eu/collection/oslo-open-standards-linked-organisations-0/about
- Fairsfair : https://www.fairsfair.org/
- ISO 19157: https://www.iso.org/standard/32575.html
- ISO 19158: https://committee.iso.org/sites/tc211/home/projects/projects---complete-list/iso-19158.html
- Great Expectations platform: https://greatexpectations.io/

4 Data Spaces and MIMs

Given their complexity, enterprise architecture of (local) governments often differs from most conventional sector-specific enterprises. While most businesses specialise in product delivery for a specific market segment (typically requiring "business capabilities" such as *customer management*, *product development*, *production* and *delivery*), governments often have to cater to all customer segments within a geographical area and provide many different products and services. Therefore the set of business capabilities of (local) governments is far larger.

On top of that, governments are bound by public procurement rules, which can lead to a further differentiation in their ICT-architectures. Public procurement aims to enable a free and thriving market, but this also implies that it is difficult for public buyers to limit themselves to working with a single or limited number of vendors. That is why interoperability is such an important consideration within public administrations, and why it has to be considered at a higher level of abstraction than simply through standardisation.

4.1 Necessity of MIMs

This higher level of abstraction can be provided by the "Minimal Interoperability Mechanisms" (MIMs). Following the definition provided by the reference paper from Living-in.eu MIMs plus:

> "MIMs *are the minimal but sufficient capabilities needed to achieve interoperability of data, systems, and services between buyers, suppliers and regulators across governance levels around the world. Because the mechanisms are based on an inclusive list of baselines and references, they take into account the different backgrounds of cities and communities and allow cities to achieve interoperability based on a minimal common ground. Implementation can be different, as long as crucial interoperability points in any given technical architecture use the same interoperability mechanisms. The MIMs are vendor neutral and technology agnostic, meaning that anybody can use them and integrate them in existing systems and offerings, complementing existing standards and technologies*".

Thus, a MIM is a description of a common set of required tasks or processes to provide such minimal but sufficient set of capabilities that a city needs to achieve a certain objective, along with guidance to help provide a useful level of interoperability between different technical solutions or approaches that may be used to achieve that set of capabilities (Taken from the Recommendation Y-MIM of ITU Study Group 20).

Minimal here is used to describe something that can meet a specific objective without an unnecessary complexity. Typically, this may be based on an existing standard, but will focus on those requirements in the standard that will enable the

user to put in place a basic implementation of what is needed to achieve a city objective.

A key aim of a MIM is also to achieve an acceptable level of interoperability. The achievement of complete interoperability often requires a great deal of work, a high level of expertise, and time to implement. Therefore, there are many circumstances, where a less than perfect level of interoperability can provide a useful first step.

In order to create functioning data spaces in cities and communities with widely differing characteristics and policy priorities, and keeping in mind that local authorities have to be active across all sectors and customer (citizen) segments, it is unrealistic to expect standards and specifications to emerge for a deep level of interoperability across all domains. Therefore, in the context of data spaces for cities and communities especially, MIMs provide a first focal point through which minimal interoperability can be achieved.

In Europe, MIMs (in this case called MIMs Plus) are governed by the Living-in.eu movement, specifically by the "tech subgroup". This subgroup is one of 5 subgroups defined within Living-in EU. This implies that the development of the MIMs Plus adheres to these principles:

- Citizen-Centric approach
- A City-Led Approach at EU level
- The City as a citizen-driven And open innovation ecosystem
- Technologies as key enablers
- Ethical and socially responsible access, use, sharing and management of data
- Interoperable digital platforms with open standards, APIs and shared data models

Clearly, the Living-in.eu tech subgroup is governed by administrations and is designed to have public bodies' best interests at heart. It is public institutions which are driving the development of the MIMs.

4.2 Overview of MIMs

In total, the Living-in.eu community has identified 10 distinct MIMs which are necessary to achieve minimal interoperability:

MIM	Subject	Name	Status
MIM1	Context	MIM1: Context Information Management	Governance
MIM2	Data Models	MIM2: Shared Data Models	Governance

MIM3	Contracts	MIM3: Ecosystem Transactions Management	Capability
MIM4	Trust	MIM4: Personal Data Management	Capability
MIM5	Transparency	MIM5: Fair Artificial Intelligence	Capability
MIM6	Security	MIM6: Security management	Work item
MIM7	Places	MIM7: Geospatial information management	Capability
MIM8	Indicators	MIM8: Ecosystem indicator management	Work item
MIM9	Analytics	MIM9: Data Analytics Management	Work item
MIM10	Resources	MIM10: Resource Impact Assessment	Work item

Table 4 - List of MIMs

Following, we show in detail MIMs mapped to the BBs, as described in the Catalogue.

MIM1- Context Information Management

Context information management manages the context information coming from Internet of Things (IoT) devices and other public and private data sources, providing cross cutting context data and access through a uniform interface. It therefore ensures comprehensive and integrated access, use, sharing, and management of data across different solutions and purposes.

This feature is paramount for data spaces to enable interoperability and data value creation. It specifically supports data exchange API and metadata & discovery services given its focus on making information usable, discoverable, and accessible.

MIM2 - Shared Data Models

In order to be able to link data sets to other sets that add important context information, it is important that the data sets being used from elsewhere use precisely the same definitions for key terms as the original dataset. For instance, if the original data set defines “children” as people aged between 5 and 15 and the other data set defines children as people between the ages of 2 and 12, then a great deal of inaccuracy would result by combining them. More fundamentally, to enable data sets to be combined automatically, the terms used in each data set need to be defined in machine readable terms so that the APIs can “understand” how to handle them. Data models are machine readable definitions of key terms. And finally, the data models need to be in a format consistent with MIM1 to enable

Apps to link relevant context data with data sets. This feature ensures interoperability.

MIM3 - Ecosystem Transactions Management

Data spaces within cities and communities require easy and risk-free access to suitable local data sources that are already within those communities. A local data space can include a marketplace allowing for easy and risk-free access to relevant and available local data, solutions, and other resources so that new and valuable services and solutions, many of which have been already deployed in other cities can easily be implemented within the local area. The use and reuse realises new societal values, including new revenue streams, incentivising the stakeholders, including infrastructure owners, to share data, analytics, services and/or solutions in infrastructure partnerships based on key technology enablers.

MIM3 is the management layer that allows stakeholders:

- To provide data along with relevant information about its content and quality and any terms and conditions for use.
- To provide data processing services along with relevant information and terms and conditions for using the services.
- To find and access the data and data processing services and other services they need and to be able to gain relevant insights into what those data streams/data processing services/data applications consist of and how valuable they can be.

Hence, this MIM contributes to technical building blocks that define data traceability, data usage, publication and to governance to ensure appropriate business, organisational and operational agreements.

MIM4 - Personal Data Management

MIM4 focuses on Personal Data Management in other words how to provide easy to use methods for citizens/users to control which data sets/attributes they want to share with solution, application, or service providers under transparent circumstances, enabling trust between the different parties. There are many initiatives seeking to provide personal data management solutions, but these are primarily in the pilot or development phase, and this has led to a fragmented marketplace. Some projects focus just on personal data management, others, such as RUDI, aim to support wider data sharing ecosystems, but with personal data management being a key feature.

There are two networks of providers – MyData and Solid, which each follow different high- level methodologies. Even within each of these two networks, there are significant differences in the technical and processes used by different projects and so individual implementations are not necessarily interoperable. There are a number of initiatives outside of these networks developing their own technical solutions.

The role of MIM4 is to identify the key capabilities required and identify pivotal points of interoperability between the different solutions to help build confidence and support implementation. This MIM relates to the governance of data spaces by including the management and consideration of personal data in the business, organisational and operational agreements. It also contributes to some of the technical building blocks of identity management and trusted exchange.

MIM6 - Security Management

MIM6 focuses on potential risks that can cause financial burdens or loss of services. In turn, it also looks at solutions and measures to be taken as a response to those. In the context of data spaces, this MIM helps 'governance' to include security considerations in the business, organisational, and operational agreements as well as it contributes to technical building blocks by underscoring identity management and trusted exchange.

MIM7 - Geospatial information management

MIM7 aims to provide Minimal Interoperability Mechanisms related to geo-temporal data. However, there are many existing geo-temporal data standards that are of relevance to cities and to propose the full list would not be compatible with the concept of MIMs. MIM7 is therefore being developed as a number of parts.

During the work on MIM7 it has become clear that there are considerable inconsistencies between MIM7 on one hand and MIM1 and MIM2 on the other. Those inconsistencies are related both to the scope of the respective MIMs, and also due to the fact that they are based on two different ecosystems of standards that do not seem to align at the moment. The geospatial world is strongly based on the OGC ecosystem of standards, whereas MIM1 & MIM2 are based on the ETSI ecosystem of standards. In order for the three MIMs to work together for a municipality this needs to align. MIM7 Part 1 has been developed to address this issue. MIM7 Part 1 comprises two minimal requirements and two recommendations.

Aligned with the Rules for the structure and drafting of International Standards endorsed by the ISO and OGC OGC (see sub-clause 5.3 of [OGC 06-121r9]). The verb form "shall" indicates a requirement to be strictly followed to conform to this MIM. Recommendations, in turn, are based on good practices and 'should' not be strictly followed. This MIM relates to Metadata & discovery services.

Full description of these MIMs can be found in the Appendix III.

4.3 Mapping between MIMs and Building Blocks

The proposed building blocks by DS4SSCC Catalogue will be the "mechanisms" to implement the MIMs. The following picture shows which MIMs will be implemented by each of the BBs.

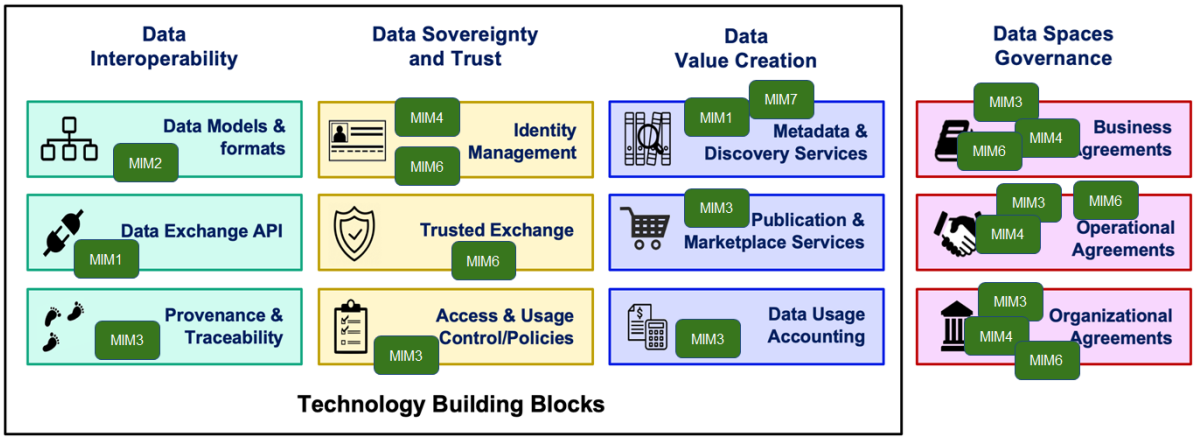

Figure 6. MIMs mapping into Data Spaces Building Blocks

The table below describes how the related MIMs are implemented in each BB:

Building Block	Relevant MIMs	Rationale
Data Models & formats	MIM2	MIM2 defines the shared data models to achieve interoperability
Data Exchange API	MIM1	Data exchange is happening only with the context management API to enable Data Spaces
Provenance & Traceability	MIM3	This is the closest match to what MIM3 is defining. Provenance & Traceability still needs definition in details and MIM3 could be expanded to address this.
Identity Management	MIM4, MIM6	This is about people identification and handing the personal data and also the Data Space security. So MIM4 and MIM6 are mapped.
Trusted exchange	MIM6	The exchange of the data needs to be done seamlessly and trust between parties exchanging the data is essential. MIM6 which is about Security Management can be the best fit to the Trusted Exchange.
Access & Usage Control	MIM3	Access to the data and usage policies where the data owner has full rights to control the data sharing can be mapped to MIM3, which

		is about ecosystem transaction management and contracts. Data owners should have defined rights, obligations and commitments to the contracts when sharing the data.
Metadata & Discovery	MIM1, MIM7	MIM1 is the closest match as the data items can be discovered using Marketplace services but it needs to be enabled by the data exchange capabilities, where it should enable cataloguing and discovery services. Given MIM7's objective to cover geo-temporal data it helps discovery.
Publication & marketplaces	MIM3	MIM3 needs to be further expanded to accommodate the marketplace services. At this stage it is the closest match.
Data Usage Accounting	MIM3	MIM3 which also defines the contracts and usage is mapped to this. Also Data usage and accounting at the MIM1 should be also defined to be able to extract the usage of the data and accounting of the data usage.
Business agreements	MIM3, MIM4, MIM6	All the governance related blocks can be mapped to MIM3 and needs further well defined guidelines for Data Space governance. MIM4 and MIM6 point to the non-technical need to regulate security and personal data management in the business agreements.
Operational agreements	MIM3, MIM4, MIM6	All the governance related blocks can be mapped to MIM3 and needs further well defined guidelines for Data Space governance. MIM4 and MIM6 point to the non-technical need to regulate security and personal data management in the operational agreements.
Organisational agreements	MIM3, MIM4, MIM6	All the governance related blocks can be mapped to MIM3 and needs further well defined guidelines for Data Space governance.MIM4 and MIM6 point to the non-technical need to regulate security and

		personal data management in the organisational agreements.

Table 5 - Rationale about MIMs mapping into Data Spaces BBs

5 Conclusions and next steps

This document aims at describing the Catalogue of Specifications for the Building Blocks to build data spaces in the scope of Smart and Sustainable Cities and Communities.

The Catalogue follows the **taxonomy** of BBs described in the **OpenDEI** framework and proposed by the DSBA Technical Convergence and DSSC. The proposed BBs are mechanisms for implementing MIMs proposed by Living-in.eu. A clear mapping between BBs of the Catalogue and MIMs (from **MIM1 to MIM7**) has been presented.

Each BB in the Catalogue includes a functional description, related standards and industry body specifications and available implementations. A level of **maturity** for each BB has been stated according to the degree of awareness and adoption of such BB in the SSCC domain.

The collection of **standards, specifications and implementations** have been carried out through desk research by project partners, interviews to experts in the field (12) and a survey to cities and suppliers (46 technical related answers). A methodology has been put in place to set up the process, the targets and the instruments for the collection.

All the received inputs through the **survey and interviews** have been analysed, processed and assessed. The complete assessment can be consulted in the Appendix II. Considered inputs have been included already in the presented Catalogue in the document, while future inputs will be considered in the online version. A total of **43 standards and specifications** and **more than 30 existing implementations** for technical building blocks have been identified so far.

Besides this narrative version of the Catalogue, DS4SSCC is publishing also an **online version** of the Catalogue aiming at:

- Facilitating the navigation through the Catalogue by filtered search and different views of content.
- Fostering the easy evolution of the Catalogue with the time being till the end of the project first, and during the deployment phase, later on.

Having this online version allows us to dynamically update it by adding new standards, specifications or implementations, at the same time as adapting it to the future recommendations coming from DSSC. The Catalogue is a vivid tool and the digital support permits a continuous evolution beyond this document frozen at the delivery time.

The online Catalogue will follow the same taxonomy as OpenDEI one and the information per BB as shown in Figure 2. For the development of the digital Catalogue, DS4SSCC has counted with the support of the DS4SKILLS project which has adapted its online inventory to our needs. There is a plan to extend the

online Catalogue feature with the inventory of use cases and datasets under collection in WP4 and WP2 in future stages of the project.

The online Catalogue is accessible at DS4SSCC web site.

During the collection and analysis process we have been aware of several initiatives and ongoing projects which can be relevant for the SSCC data space, especially in relation to the standards. Therefore, one of the next steps will be to have a look at the evolution of following actions:

- Rolling plan for ICT Standardisation for smart cities and communities
- European Interoperability Reference Architecture (EIRA)
- Common Assessment Method for Standards and Specifications (CAMSS)

In the next deliverable of this WP3, a Reference Architecture for all the referred BBs will be depicted. It will provide a reference framework for using and combining the different BBs, as well as the interactions amongst them. The architecture will be accompanied with a CookBook which will provide the necessary guidelines to use the BBs in the typical scenarios for SSCC. Several examples will illustrate the process by customising the architecture in the selected use cases by WP4.

6 Appendix I: Relevant EU regulations and legislations

In alphabetical order:

- **Data Governance Act** (DGA): Proposal for a Regulation of the European Parliament and of the Council on European data governance: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52020PC0767
- **Digital Markets Act** (DMA): Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828: https://eur-lex.europa.eu/legal-content/EN/TXT/?toc=OJ%3AL%3A2022%3A265%3ATOC&uri=uriserv%3AOJ.L_.2022.265.01.0001.01.ENG
- **Digital Services Act** (DSA): Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32022R2065
- **ePrivacy Directive**: Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058
- **eIDAS**: Regulation 910/2014 on electronic identification and trust services for electronic transactions in the internal market (eIDAS).
- **E-Commerce Directive:** Directive 2000/31/EC on certain legal aspects of information society services, in particular electronic commerce, in the internal market
- **General Data Protection Regulation** (GDPR) : General Regulation on data protection 2016/679, of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data: https://eur-lex.europa.eu/eli/reg/2016/679/oj
- **Open Data Directive**: Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024
- **Platform-to-Business Regulation**: Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32019R1150
- **Regulation on the Free Flow of Non-Personal Data:** Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European

Union: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1807

Upcoming

- **Data Act**: Proposal for a Regulation Of The European Parliament And Of The Council on harmonised rules on fair access to and use of data :https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN
- **Interoperable Europe Act**: Proposal for a Regulation of the European Parliament and of the Council laying down measures for a high level of public sector interoperability across the Union :https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0720

Main resources:

- Da Bormida in Cugurra, M. (2022) Does Everything Conform to Legal, Ethical, and Data Protection Principles? In Topham, S, Boscolo, P & Mulquin, M. *Personal Data Smart Cities*. Rivers Publishers: Denmark
- European Commission (2022) Staff working document on data spaces, 23 February: https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces

7 Appendix II: Survey inputs assessment

This appendix includes all the received inputs (standards, specifications and implementations) collected through the survey to stakeholders, and indicates the evaluation made for each of them by the project experts. In some cases, the input has been included in the Catalogue at the corresponding BB; and in other cases it has not been considered as it does not fit into the description of the BB or it is out of scope.

Data Exchange API			
Received input	Type	Comment	Recommendation
Web services (Object Document Model, Object Relational Mapping)	Specification	Basic technology	Not considered
IDSA Connectors	Specification	Not used for exchange API but for access control	Add in Access & Usage Control BB - refer to IDS-RAM 3.0
SOLID	Specification	Allow to store personal data securely in decentralized data stores called Pods, kind of secure personal web servers for data. Does not fit here.	Not considered here, but included in Identity Management BB.
JSON	Standard	It is a lightweight data-interchange format, evolved to Linked Data version.	Included, but use LD version
SQL	DB language	Does not fit here	Not considered
LDES	Specification	Linked Data Event Streams	Included
IoT Agents	SW component	Does not fit here	Not considered
OGC	Organization	Too generic, but in case refers to Publication	Include OGC-WFS in Publication & Marketplaces
CitySDK	Solution	Does not fit here	Not considered
API REST	Standard	Basic technology	Not considered
iShare	Solution	Does not fit here	Not considered
OpenAPI	Standard	Basic technology	Not considered
PSD2	Regulation	Does not fit here	Include in Regulations

			section
HL7	Standard	Does not fit here	Include in Data Models
OPIN	??	Not clear what is referring, several acronyms for OPIN	Not considered
FHIR	Standard	Evolution of HL7. Does not fit here	Include in Data Models
MQTT	Standard	Communications protocol, use as API sometimes	Included
XMPP	Standard	It is a set of open technologies for instant messaging, presence, multi-party chat, voice and video calls… Does not fit here	Not considered
HTTP	Specification	Basic technology	Not considered
Web sockets	Specification	Comms protocol. Does not fit here	Not considered
IoT Hub	Solution	Does not fit here	Not considered
CityAPI	Solution	Does not fit here	Not considered
CKAN API	Specification	API for open data portals, not for exchange data	CKAN is included in Publication & Marketplaces
AWS	Services	It provides cloud computing services. Does not fit here.	Not considered
SOAP	Specification	Basic technology	Not considered
Auroral	Project	Does not fit here	Not considered
Dutch API	Specification	Specific API for local domain	Not considered
ChatGPT	Commercial Product	It is a solution which provides a model trained to follow an instruction in a prompt and provide a detailed response. Does not fit here	Not considered
DSBA Technical Convergence	Specification	This document proposes a general approach to implement data spaces. Regarding Exchange API,	Included NGSI-LD

		it proposes NGSI-LD that has been already included	
GAIA-X	Specification	It provides specifications for different matters, not sure which one refers here, so it is not possible to fit it properly here.	Not considered

Data Models and Formats			
Received input	Type	Comment	Recommendation
CGIAR	Organization	It defines data models	Included
BIM	Specification.	Digital representation of a building or infrastructure	Not considered
DCAT-AP	Standard	For open datasets' metadata	Included in Metadata&Discovery
RDF	Standard	Description of web semantic resources. Basic technology	Not considered
NGSIv2	Standard	For generic API. Not defines data models. New version NGSI-LD	Included NGSI-LD in Exchange API
Data Privacy Vocabulary (DPV)	Standard	W3C	Included
OGC	Organization	Define multiple standards. Refer to data models	Included
MIMs	Specification	Defines recommendations for data models. Implemented in SDM	SDM included
OPC	Specification	Communication protocol plus companions with defined data models	Included
NGSI-LD	Specification	It does not define data models	Not include (in data models)
INSPIRE	Directive for geographic public information	It defines data models	Included
OSLO	Standard	Ontology partially mapped in SDM	Included

ISO TC/211	Standard	Standardization in the field of digital geographic information.	Not included
Cesium	Implementation	3D geospatial platform, it is a solution, not a data model	Not included
CityGML	Implementation	XML-based format used for the representation and exchange of 3D city models	Included
SensorThings	Implementation	It belongs to OGC for IOT representation	Included
OASIS	Organization of open standards	It is a publisher of standards	Not considered
NeTEx/Siri	Standard	EU format for data interchange	Included
DatexII	Standard	EU regulation for defining transport data models	Included
GBFS/GTFS	Standard	Data models standards for mobility	Mapped in SDM. Included
Smartdatamodels.org	Specification	Collaborative Data models repository.. Based on open standards or open contributions	Included
MaaS Data String	Specification	Defined by MaaS Alliance. Based on GTFS	Included
TOMP	Implementation	Shared elements with OSM and GTFS	Solution Not included. Data models Included
BigQuery	Implementation	Out of scope	Not included
SQL	Technology	Basic technology	Not included
PowerBI	Implementation	Out of scope	Not included
OPIN	Standard	Insurance data models	Included
LDES	Specification	Linked Data Event Streams	Included in Exchange API
OpenBanking & PSD2	Specification	Define data models for finance	Included PSD2; Open Banking moved to Publication and Marketplaces

EMF	Framework	Eclipse Modeling Framework	Not included
CityJSON	Specification	Data models and representation	Included
DICOM	Specification	Data models for Digital Imaging and Communications in Medicine	Included
Data Mesh	Data architecture	Out of scope	Not included
CAD	Concept	Computer-Aided Design	Not included
Excel	Implementation	Out of scope	Not included
W3C-WoT	Standards	Web of things. Too wide scope	Not considered
DCAT	Standard	Data Catalog Vocabulary	Not included
SAREF	Standard	Published by ETSI	Included
OSGi	Implementation	Open Service gateway initiative	Not included
SensiNact	Implementation	It can use and provide BBs. Out of scope	Not included
MQTT	Standard	Communications protocolo, use as API sometimes	Included in Exchange API
REST APIs	Standard	Basic technology	Not included
Dutch models	Implementation	traffic and transportation models	Included
Madrid models	Implementation	For different aspects of a Smart city	Included
FairsFair.org	Project	For FAIR identity management - close source	Included in Identity Mng as implementation
SCORE Water	Project	Possible use or definition of data models, TBC	Not included for now

Provenance and Traceability			
Received input	Type	Comment	Recommendation
DCAT-AP	Specification	Not very mature but provides some feature	Included

BIM	Specification	It is a language for modeling the construction information. Not specific about provenance	Mentioned that some research at the respect by using blockchain
OPC	Standard	An industrial interoperability standard. Provider of OPC-UA. We have not found in the specifications any specific process to manage the provenance& traceability of the data	Not considered until further exploration
Geontology	Project	A geo-aware network protocol for enabling trustable cross-border operations and data exchange in a global digital economy. Not specific about P&T	Not considered

Identity Management			
Received input	**Type**	**Comment**	**Recommendation**
W3C WAC	Specification	W3C WAC specification is a method for managing access control and permissions for web widgets. It provides a standardized method for defining permissions and access control policies for web widgets that can be implemented using different web technologies and platforms. The W3C WAC (Widgets Access Control) is a specification that defines a method for managing access control and permissions for web widgets. As a specification, it describes how web widgets can be secured and controlled in terms of access to resources and data.	Not considered here. Move to the access control section.

W3C ODRL	Standard	The W3C ODRL standard is a machine-readable language for expressing and managing rights and permissions for digital content. It provides a common framework and vocabulary for expressing rights and permissions that can be used by different platforms and applications.	Not considered here. Move to the access control section.
W3C DID	Specification	W3C DID specification is a technical specification that defines a standardized method for creating, resolving, and managing decentralized identifiers. Decentralized identifiers (DIDs) are a new type of identifier that enables verifiable, decentralized digital identity. A DID refers to any subject (e.g., a person, organization, thing, data model, abstract entity, etc.) as determined by the controller of the DID.	Included in the decentralized baseline standards and specifications
LDAP	Protocol	The Lightweight Directory Access Protocol (LDAP) is an open, vendor-neutral, industry standard application protocol for accessing and maintaining distributed directory information services over an Internet Protocol (IP) network. LDAP is specified in a series of Internet Engineering Task Force (IETF) Standard Track publications called Request for Comments (RFCs), using the description language ASN.1. The latest specification is Version	Included in the centralized baseline standards and specifications

		3, published as RFC 4511. LDAP is based on a simpler subset of the standards contained within the X.500 standard. Because of this relationship, LDAP is sometimes called X.500-lite.	
eIDAS	Regulation	eIDAS is a European Union regulation that provides a framework for electronic identification, authentication, and trust services that enables secure and seamless electronic transactions across borders within the EEA. The regulation sets out requirements and standards for electronic identification and authentication and establishes a trust framework for electronic transactions.	Included as regulation in appendix
ZVOP-2	Legislation	Personal data protection act of the Republic of Slovenia. Too specific	Not considered
AML-KYC	Regulation	Know Your Customer (KYC) are guidelines and regulations in financial services that require professionals to verify the identity, suitability, and risks involved with maintaining a business relationship with a customer. The procedures fit within the broader scope of Anti-Money Laundering (AML) and Counter terrorism financing (CFT) regulations.	Not considered
GDPR	Regulation	General Data Protection Regulation	Included as regulation in appendix
DSML	Technology	DSML is an XML-based language used to represent and exchange	Included as regulation in appendix

		directory service information. It provides a standardized method for accessing and manipulating directory services and facilitates interoperability between different directory service implementations and client applications.	
IDSA-IM	Implementation	IDSA identity management refers to the set of technologies and processes used by the International Data Spaces Association to manage the identities of users and entities involved in data exchange. The IDSA identity management system is based on open standards and interoperable technologies and includes digital identities, certificates, policies, and governance mechanisms that enable secure and trusted data exchange between organizations and across industries.	Not considered. I found no information about the actual identity management solution.
OS2	Implementation	OS2 service has been developed with the aim of giving small and medium-sized municipalities the opportunity to get a service that can be established quickly and without the municipality having to build competences on new IT technologies.	Not considered. No information available
W3C Verifiable Credentials	Standard/specification	W3C Verifiable Credentials (VC) is a standard for issuing, presenting, and verifying digital credentials, which was developed by the World Wide Web Consortium (W3C).	Included in the decentralized baseline standards and specifications

		W3C Verifiable Credentials can be considered both a standard and a specification. It defines a common set of technical requirements and specifications for digital credentials, but also includes guidelines for their implementation and usage.	

Trusted exchange			
Received input	Type	Comment	Recommendation
OPC	Standard	For interoperability, not related to trust	Not considered here.
FairsFair.org	Project	For FAIR identity management - close source	Not considered here
eIDAS	Regulation	eIDAS is a European Union regulation that provides a framework for electronic identification, authentication, and trust across borders	Included as regulation in appendix
ZVOP-2	Legislation	Personal data protection act of the Republic of Slovenia. Too specific	Not considered
AML-KYC	Regulation	Know Your Customer (KYC) are guidelines and regulations in financial services that require professionals to verify the identity, suitability, and risks involved with maintaining a business relationship with a customer. Too specific	Not considered
GDPR	Regulation	General Data Protection Regulation	Included as regulation in appendix
XACML	Specification	Access control language. Not related to trust.	Included in Access and Usage Control BB.

Access and Usage Control			
Received input	Type	Comment	Recommendation
W3C WAC	Specification	A method for managing access control and permissions for web widgets	Included
W3C ODRL	Standard	A machine-readable language for expressing and managing rights and permissions for digital content.	Included
OAUTH2	Standard	industry-standard protocol for authorization	Included in Identity Mng BB
ACM-IDM	Implementation	OpenID provider for Flemish Gov. Too specific	Included in Identity Mng BB
FairsFair.org	Project	For FAIR identity management - close source	Not considered here
eIDAS	Regulation	eIDAS is a European Union regulation that provides a framework for electronic identification, authentication, and trust across borders	Included as regulation in appendix
ZVOP-2	Legislation	Personal data protection act of the Republic of Slovenia. Too specific	Not considered
AML-KYC	Regulation	Know Your Customer (KYC) are guidelines and regulations in financial services that require professionals to verify the identity, suitability, and risks involved with maintaining a business relationship with a customer. Too specific	Not considered
GDPR	Regulation	General Data Protection Regulation	Included as regulation in appendix
REST role access	API	Similar to RBAC concept, so embedded somehow in the description of the BB	Somehow included
DPIA Netherlands	Implementation	It assesses the data	Not considered

		protection risks of the (professional) use of Microsoft Teams in combination with OneDrive, SharePoint Online and the Azure Active Directory. Too specific	
GEMMA	??	Several products in Internet with such name. Not clear relationship with this BB	Not considered
KeyCloak	Implementation	open source identity and access management solution.	Included in Identity Mng BB
API keys	Implementation	It is a token that a client provides when making API calls. A decentralised IAM with a-priori unknown API consumers in data spaces does not fit well	Not considered
RBAC	Concept	It is a concept which is mentioned in the description of the BB	Included
Open Policy Agent	Specification	Policy-based control for cloud native environments	Included
Microsoft AD	Implementation	directory service developed by Microsoft for Windows domain networks	Included in Identity Mng BB

Metadata & Discovery			
Received input	Type	Comment	Recommendation
CKAN	Implementation	open-source DMS (data management system) for powering data hubs and data portals	Included in Publication&Marketplaces
DCAT	Standard	Data Catalogue Vocabulary	Included
INSPIRE Metadata	Standard	For open datasets' metadata	Included
Metadata Vlaanderen	Implementation	Catalogue of Open Data	Not included

		in dutch, not metadata	
Datavindplaats	Implementation	Good example	Included
OSGi	Implementation	Open Service gateway initiative	Not included
NGSI-LD	Standard	API for data context exhange	Included in Exchange API
Wikidata	Implementation	Central storage for the structured data of its WikiX	Not considered
Smartdatamodels	Implementation	Repository of open data models.	Included in Data Models&Formats
ISO 19115	Standard	GIS metadata	Included
Context Broker	Implementation	Building block which implements NGSI-LD API specification	Included in Exchange API
Open Data Portals	Implementation	Web portals where the open data is published	Not considered
Grafana (JSON, CSV, XML)	Implementation	For visualization	Not considered
DCAT-AP	Specification	Based on DCAT	Included DCAT
SHACL	Specification	Language for validating RDF graphs against a set of conditions.	Not considered
SDDI	Specification	Transmit data between devices.	Not considered
SAML / LIberty	Standard	SAML allows an identity provider (IdP) to authenticate users. Liberty is a server supporting SAML	Not considered

8 Appendix III: Relevant MIMs description

MIM1- Context Information Management

Objectives

Context information management manages the context information coming from Internet of Things (IoT) devices and other public and private data sources, providing cross cutting context data and access through a uniform interface. It therefore ensures comprehensive and integrated access, use, sharing, and management of data across different solutions and purposes.

This feature is paramount for data spaces to enable interoperability and data value creation. It specifically supports data exchange API and metadata & discovery services given its focus on making information usable, discoverable, and accessible.

Requirements for conformance

At its core, the additional data that a data owner will want to access is data that provides useful information about the context of their own data set. To do this it needs to be possible to automatically link the relevant parts of the data in their data set with the relevant parts of the new data set.

Context information needs to use clear and accurate data models showing the properties of the entity described by the data and its relationships to other entities. See MIM2 for more details.

Appropriate APIs can then be used to link the context data appropriately with the original database.

The implementation across (and even within) the city, or any application ecosystem, can be very diverse and heterogeneous. An agreement on the interfaces is necessary to be able to access the information. This is enabled by the context management API and the data models.

The key requirements are:

- Use of Data models complying with MIM2
- MIMs Plus version 5.0 FINAL DRAFT June 2022 · https://living-in.eu/mimsplus
- Use of appropriate APIs and an Information model containing ...?
- The common data and data models need to be available in a catalogue, along with guidelines, so that different verticals are integrated in a holistic/integrated city data lake to enable interoperability for applications and systems among different cities. The catalogue should support structural interoperability, behavioural interoperability (representation, data mappings) and governance interoperability.

Recommended Specifications

- NGSI-LD, as specified by the ETSI Industry Specification Group on Context Information Management (ETSI ISG CIM), provides an API for managing and requesting context information and an underlying meta model based on entities - the core information elements, often the digital counterparts of real-world object - and their properties and relationships to other entities.
- Even though the NGSI-LD specification has been published relatively recently, there are already three Open-Source implementations (Scorpio, djane and Orion-LD). Orion-LD is the NGSI-LD version of the Connecting Europe Facility (CEF) building block Context Broker.

In addition, data models are needed that are, or can be made to be, compliant with NGSI- LD. See MIM2.

A relevant specification under development:

- INSPIRE: will further develop OAPIF by OGC as a driver linking to OGC APIs to enable access to complex geospatial context information that compliments the geospatial characteristics covered by NGSI-LD

Verification

ETSI organized a Testing Task Force (TTF) to create a Testing toolkit to validate context brokers towards the NGSI-LD specification. The result was a set of clearly defined test descriptions, test purposes and executable robot scripts. All this information can be found on the ETSI CIM Website https://www.etsi.org/comittee/cim.

Relevant European References and Specifications

- European Commission 2019 European Interoperability Reference Architecture, EIRA© https://joinup.ec.europa.eu/collection/european-interoperability-reference-architecture- eira/about
- European Commission 2020 Core Public Service Vocabulary Application Profile https://joinup.ec.europa.eu/solution/core-public-service-vocabulary-application-profile
- European Commission 2020 Core Vocabularies https://joinup.ec.europa.eu/solution/e-government-core-vocabularies/release/20
- European Commission 2017. Communication on The European Interoperability Framework- Implementation Strategy COM (2017) 134Annex 2, Retrieved from: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52017DC0134
- Nan Zhang, Xuejiao Zhao, and Xiaope He 2020 Understanding the relationships between information architectures and business models: An

empirical study on the success configurations of smart communities Government Information Quarterly v37 (2), https://doi.org/10.1016/j.giq.2019.101439

- The Berlin declaration on digital society and value based digital government (German): https://www.bmi.bund.de/SharedDocs/downloads/EN/eu-presidency/berlin-declaration- digital-society

MIM2 - Shared Data Models

Objectives

In order to be able to link data sets to other sets that add important context information, it is important that the data sets being used from elsewhere use precisely the same definitions for key terms as the original dataset. For instance, if the original data set defines “children” as people aged between 5 and 15 and the other data set defines children as people between the ages of 2 and 12, then a great deal of inaccuracy would result by combining them. More fundamentally, to enable data sets to be combined automatically, the terms used in each data set need to be defined in machine readable terms so that the APIs can “understand” how to handle them. Data models are machine readable definitions of key terms. And finally, the data models need to be in a format consistent with MIM1 to enable Apps to link relevant context data with data sets. This feature ensures interoperability.

Requirements for compliance

All the entities described by data in the data ecosystem should be described by a consistent set of data models using the Resource Description Framework (RDF) methodology, Resource Description Framework Schema (RDFS), and Web Ontology Language (OWL)

For spatial (and spatio-temporal) observation data the provisions of MIM-7 (Places) regarding data encoding have to be taken into consideration.

In order to ensure wider interoperability, it is recommended that data models should all be taken from one of the relevant existing Data model initiatives, see below.

Recommended Specifications

The preferred option is to follow the NGSI-LD compliant data models for aspects of the smart city. These have been defined by organisations and projects, including OASC, FIWARE, GSMA and the SynchroniCity project and there is an ongoing joint activity of OASC, TM Forum and FIWARE to specify more - the smart data models initiative: https://smartdatamodels.org/

Alternatively, existing data models and ontologies can be mapped for use with NGSI-LD by identifying what are entities, properties and relationships, which can be managed and requested by the NGSI-LD API. Some examples are as follows:

- oneM2M base ontology (that is compatible with SAREF). Additionally, oneM2M provides the means to instantiate ontologies as a means to provide semantic descriptions of the data exchanged (through the use of metadata)
- SAREF: Smart Appliances REFerence (SAREF) ontology specified by ETSI OneM2M committee with the extension of SAREF4Cities provides an ontology focused on smart cities
- Core vocabularies of ISA like Core Public Service Vocabulary Application Profile used as the basis for the Single Digital Gateway Regulation that touches local governments, Core Person, Core Organization etc
- DTDL is the Digital twin Definition Language developed by Microsoft. This language is based on top of json-ld and the existing Fiware data models are converted in this format.

MIM3 - Contracts

Objectives

Data spaces within cities and communities require easy and risk-free access to suitable local data sources that are already within those communities. A local data space can include a marketplace allowing for easy and risk-free access to relevant and available local data, solutions, and other resources so that new and valuable services and solutions, many of which have been already deployed in other cities can easily be implemented within the local area. The use and re-use realizes new societal values, including new revenue streams, incentivising the stakeholders, including infrastructure owners, to share data, analytics, services and/or solutions in infrastructure partnerships based on key technology enablers.

MIM3 is the management layer that allows stakeholders:

- To provide data along with relevant information about its content and quality and any terms and conditions for use.
- To provide data processing services along with relevant information and terms and conditions for using the services.
- To find and access the data and data processing services and other services they need and to be able to gain relevant insights into what those data streams/data processing services/data applications consist of and how valuable they can be.

Hence, this MIM contributes to technical building blocks that define data traceability, data usage, publication and to governance to ensure appropriate business, organisational and operational agreements.

Capabilities

The data space realises standardised exposure of data and data set offerings built on standard interoperability mechanisms (e.g., those result of combining MIM1 and MIM2) and mechanisms for guaranteeing security and privacy by design. The data space also realises access to services offerings that build on this data and transfer it to knowledge, intelligence, and information for the consumers.

A crucial aspect of a data space is ecosystem transaction management. These functionalities enable effective matchmaking of relevant data sources (e.g., urban IoT data) from providers with respective data consumers, facilitate trusted exploitation of such data based on enforceable data usage agreements and secure value flow between these stakeholders.

The data space needs to provide a number of capabilities which may include some or all of the following management of:

- **Catalogues** *This module provides functionalities to publish and search for different data service 1 offerings. Data offerings can be organized into groups/categories - in a hierarchical fashion when possible - to allow for an easy navigation and discovery of them. The module allows data providers to define the technical description of the data offerings they own as well as information related to the offering terms and conditions such as price, SLA, license, etc.*
- **Offers/Orders** *This module allows the ordering and acquisition of "data service" offerings and managing acquired rights on data services. More specifically, a data consumer interested in purchasing a data service offering available in the catalogue can place an order to finalize the purchase of that digital asset. It allows the performance of operations such as subscription un-subscription, activation, deactivation, and renewal.*[13]

With the term "data service" we include both data access and data processing services

- **Revenue sharing** *This module allows data providers to generate revenue for their offerings by charging data consumers for purchasing them. It provides tools to manage data service usage information in order to enable usage-based business models. It exposes an interface to interact with external charging platforms such as PayPal. It collects all the information required for the charging process (price, data service usage, consumer identifier, etc.), which may differ according to the pricing model associated with the data service offering and the outcome received by the external charging platform.*
- **SLAs and data licenses** *This module allows data service providers to set, define and customize different SLAs and licenses for data offering published*

[13] For more information on this, please refer to section 1.2.3: Publication and Marketplace services.

on the data marketplace, thus enabling the creation of a dynamic ecosystem in which data service providers can establish various business models. It provides an interface to retrieve predefined data usage license templates so that data providers can link a data usage license instance selected among the available templates to the related data service offerings.

- **Feedback and reputation** *This module provides user feedback management for the different data service offerings published on the marketplace. It also provides rating and reputation mechanisms to support data consumers in selecting the data service offerings and to promote an honest behaviour among users and providers.*
- **Party Management** *This module covers the identification and gathering of information associated to parties involved in the exchange of data through data services and which can play the role of consumers and providers of data services. Parties can be individuals or organizations playing the role of consumers and/or providers.*
- **Customer** *This module covers the identification and gathering of information about the users of the marketplace. It provides tools to manage customer information and related parties, which are the legal entities associated with the customer accounts. Depending on the access restrictions for the marketplace defined by the marketplace provider (e.g., city council, consortium, 3rd party), customers can be created and linked to specific roles (e.g., data provider, data consumer, administrator, etc.)*
- **Transparency and accountability service** *This module provides tools for auditing orders (including pricing model, license terms, SLAs) and tracking the parameters defined by SLAs.*
- **Federation** *This module manages a set of federation capabilities in accordance with the marketplace governance. Federation capabilities allow* different *marketplaces to interact with each other and access their resources to provide access to data offerings across them and enable the development of aggregated services.*

Specifications

These specifications are currently recommended by MIM3:

- Basic Data Marketplace Enablers SynchroniCity_D2.4.pdf
- Reference Architecture for IoT Enabled Smart Cities, Update: SynchroniCity_D2.10.pdf
- TM Forum Open APIs and component suites provide service and a technology-neutral suite of APIs that provide the minimum building blocks for interoperability across all operational management areas. Each API and component suite provide the specification, reference implementations and in most cases conformance test kits. Reference Implementations are available under the Apache2.0 license. These APIs have gained global adoption in the

Telecommunications industry and are proven to maximize reuse. They are designed to be extendable as required for specific services. The respective data models have been harmonised with FIWARE and GSMA data models. https://projects.tmforum.org/wiki/display/API/Open+API+Table

Examples of TM Forum specifications that link with the capabilities listed above:

- Catalogue management: TMF620 API, TMF633 API, TMF634 API, TMF637 API, TMF638 API, TMF639 API
- Offers/Orders management: TMF622 API, TMF641 API, TMF652 API
- Revenue (sharing) management including Payment Methods: TMF670 API, Payment Management: TMF676 API, Shopping Cart Management: TMF633 API
- SLA and data license management
- Feedback and reputation service
- Party Management: TMF632 API
- Customer management: TMF629 API
- Transparency and accountability service
- Federation management

An open-source implementation of these capabilities can be found in FIWARE (Business API Ecosystem framework) which was used in SynchroniCity and more recently in the i4Trust project

MIM4

Objectives

MIM4 focuses on Personal Data Management in other words how to provide easy to use methods for citizens/users to control which data sets/attributes they want to share with solution, application, or service providers under transparent circumstances, enabling trust between the different parties. There are many initiatives seeking to provide personal data management solutions, but these are primarily in the pilot or development phase, and this has led to a fragmented marketplace. Some projects focus just on personal data management, others, such as RUDI, aim to support wider data sharing ecosystems, but with personal data management being a key feature.

There are two networks of providers – MyData and Solid, which each follow different high- level methodologies. Even within each of these two networks, there are significant differences in the technical and processes used by different projects and so individual implementations are not necessarily interoperable. There are a number of initiatives outside of these networks developing their own technical solutions.

The role of MIM4 is to identify the key capabilities required and identify pivotal points of interoperability between the different solutions to help build confidence and support implementation. This MIM relates to the governance of data spaces by

including the management and consideration of personal data in the business, organizational and operational agreements. It also contributes to identity management.

Requirements for conformance

MIM4 will address needs and requirements from two perspectives:

- That of Individual citizens in terms of transparency & privacy preferences collection,
- That of Cities and Data Using Services (Data Controller/Processors/) in terms of Authorization and Data usage control and enforcement

The provisional sets of capabilities required are listed below:

For individual citizens

1. Citizens need to be able to choose the operator they wish to manage their data and to move from operator to operator
2. Citizens should be able to access their data through many different channels
3. Citizens should be able to use the identity of their choosing, in best cases a keychain of identities can be defined, so that users can choose the identity per service
4. Citizens should have insight what personal data is available, stored, shared, etc. by the providers of the applications and/or services they use
5. Citizens should be able to request changes to or deletion of part or all personal data available, stored, shared, etc. by the provider of the applications and/or services in use. The providers would need to comply with these requests unless there were legally justifiable reasons not to do so
6. Citizens should be able to indicate in which circumstances what personal data is ‘free’ to use for which parties through a 'permission arrangement’
7. Citizens should be able to grant consent to providers of the applications and/or services, be it governmental or businesses, that attribute based, decentralised storage and ‘revealing’ of personal data attributes provides full service and access to these applications and/or services
8. Citizens should be able to roam with their data between cities and internationally.

For cities and Data using services

1. Cities need to enable users to handle consent, allow and revoke access, and have full transparency on their personal data
2. Permission management needs to be handled preferably on the attribute level. Personal data processing should be described in a fine-grained manner, by covering all aspects (purposes, processing, types of data ...) in a standardized manner (see as example W3C dpv: https://dpvcg.github.io/dpv/
3. Personal Data Management needs to have an open API in line with MIM1 to broker data and standard data models MIM2. Data sources need to be open

and documented, and discoverable via MIM1, listing their data via MIM2. Operators may benefit from being groupable at joint initiative of cities with close ties

4. PDM systems need to manage the personal data to a high level of security. (The detail of how to do this will be dealt with by MIM6 - Security)
5. PDM systems need to be flexible enough to handle methodologies that require personal data pods to store the data as well as those that utilise personal data spaces or that allow the data to continue to be stored by the relevant organisation, but where the subject of the data is able to exercise rights as to its use.

Recommended Specifications

A detailed proposal for interoperability between Personal Data Management Operators was proposed to OASC in May 2021. This proposal has two pillars:

Pillar 1: One Connector for all Personal Data Management Operators

Pillar 2: Legal framework governance

The proposal is described in the paper "Towards Interoperable Personal Data Management within Smart Cities: Minimum Interoperability Mechanism 4" that can be accessed at: https://mims.oascities.org/mims/oasc-mim4-trust/references

Effectively, this defines a connector that enables any Personal Data Management provider that complies with the Legal agreement to be able to access data from any data source that is MIM4 compliant. In this way, each Personal Data Management provider can innovate freely around their technical solution, provided that it enables the capabilities defined in MIM4 while data providers only need to provide a single method for them to access the data.

While designed for the MyData network, the MIM4 proposal has now been reviewed in detail by MyData Global, Vastuu Group, Forum Virium Helsinki, RUDI (the Urban Data Initiative of the city of Rennes), the DataVaults and Kraken European Projects focusing on Personal Data Management and the CAPE personal data management solution developed by Engineering.

This review indicated that the proposed interoperability mechanism is a feasible way of enabling a level of interoperability between all of these and is likely to be relevant to all Personal Data Management solutions. All of the above initiatives have also agreed to work together over the next few months to develop demos to test the proposed MIM4 Part 1 in practice.

References

- MyData Declaration and Whitepapers
- MyData Architecture and Technical Specifications
- MIM 4 white paper: Preliminary description and validation by the City of Helsinki (MIM4 Champion) and its MyData Operator, Vastuu Group.

- MyData as MIM4 Presentation by Kimmo Karhu, Head of Data at City of Helsinki
- Ihan.fi as Testbed for Fair Data Economy and Blueprint 2.5
- Buyle, R., Taelman, R., Mostaert, K., Joris, G., Mannens, E., Verborgh, R., & Berners- Lee, T. (2019). Streamlining Governmental Processes by Putting Citizens in Control of Their Personal Data. In A. Chugunov, I. Khodachek, Y. Misnikov, & D. Trutnev (Eds.), Proceedings of the International Conference on Electronic Governance and Open Society: Challenges in Eurasia (Vol. 1135, pp. 346–359). Springer International Publishing.
- Solid project and apps and Inrupt supporting the Solid project ecosystem
- On Digital Trust Infrastructure, "Proper data use in the public space" publication (in Dutch) which calls for research into a generic trust infrastructure in the public domain. In addition to recommending the inventorization and evaluation of digital infrastructure in the public space, it recommends "investigating possibilities for the realisation of a national, impenetrable and open digital trust infrastructure for identification, authentication and authorisation of personal data, including the related governance."
- When working on project architecture and use cases, reuse I Reveal My Attributes (IRMA) architecture and apps, from the (Dutch) Privacy by Design Foundation

MIM6

Objective

MIM6 focuses on potential risks that can cause financial burdens or loss of services. In turn, it also looks at solutions and measures to be taken as a response to those. In the context of data spaces, this MIM helps 'governance' to include security considerations in the business, organisational, and operational agreements as well as it contributes to technical building blocks by underscoring identity management and trusted exchange.

Baseline Specifications

- International Standard ISO/IEC 27005:2018 Information technology – Security techniques – Information security risk management
- NIST Special Publication SP800-53, Security and Privacy Controls for Information Systems and Organizations
- REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27vApril 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

MIM7

Objective

MIM7 aims to provide Minimal Interoperability Mechanisms related to geo-temporal data. However, there are many existing geo-temporal data standards that are of relevance to cities and to propose the full list would not be compatible with the concept of MIMs. MIM7 is therefore being developed as a number of parts.

During the work on MIM7 it has become clear that there are considerable inconsistencies between MIM7 on one hand and MIM1 and MIM2 on the other. Those inconsistencies are related both to the scope of the respective MIMs, and also due to the fact that they are based on two different ecosystems of standards that do not seem to align at the moment. The geospatial world is strongly based on the OGC ecosystem of standards, whereas MIM1 & MIM2 are based on the ETSI ecosystem of standards. In order for the three MIMs to work together for a municipality this needs to align. MIM7 Part 1 has been developed to address this issue.

MIM7 Part 1 comprises two minimal requirements and two recommendations.

Aligned with the Rules for the structure and drafting of International Standards endorsed by the ISO and OGC OGC (see sub-clause 5.3 of [OGC 06-121r9]). The verb form “shall” indicates a requirement to be strictly followed to conform to this MIM. Recommendations, in turn, are based on good practices and ‘should’ not be strictly followed.

This MIM relates to Metadata & discovery services.

Requirements

1. Expose data through a service interface either through OGC wfs or OGC API features
2. Ensure that all published features have unique identifiers that follow the requirements of the Inspire directive data specifications, chapter 14 Identifier management: https://inspire.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4rc3.pdf or the work of W3C in the data on the web best practice: https://www.w3.org/TR/dwbp/#DataIdentifiers

Recommendations

1. If data is shared through wfs, a proxy OGC API could be considered on top of that
2. The use of standard-based encoding such as GeoJSON, GML, GeoPackage and CityGML

Rationale

- MIMs are Minimal Interoperability Mechanism that should be relatively easy for cities and communities to achieve.
- The Inspire Directive, leveraging data sharing, description principles and standards like WMS and WFS, has transformed the European geospatial landscape in the last decade, and is making geodata interoperable throughout Europe.
- A main recognised challenge for European municipalities is to integrate and transfer data between internal and external IT systems.
- That most municipalities with minimal effort can establish OGC services like WFS, WMS and OGC APIs with minor investments.
- Geodata-based features need to be accessed as linked data by many IT- and IoT- systems, and over a long period of time, thus persistent identifiers are vital for the integrity of IT- and IoT-systems over time.
- For municipalities with more technical and financial strength the OGC ecosystem of standards for both geodata and sensor data are a good basis for more complex services.

Understanding that:

- The Feature and Thing (in OGC and entity in NGSI-LD) is the essential item for integrating between the two ecosystems of standards.
- That context will be created from data from various sources, for example geodata and building information models.
- A main challenge for municipalities will be to both establish and maintain the number of connections between NGSI-LD entities and their representations in the SDI (identifiers, existence, location) over time and that this process will need to be automated, most probably based on geospatial techniques like geodata or in the more complex case a digital twin.

Means of verifications

An advantage of INSPIRE is the ability to validate metadata, services and data against the technical provisions listed above. To this end, the INSPIRE reference validator, fully based on open-source components, is being used. Local instances of the tool can be deployed within the city's own infrastructures in addition to the centrally available solution.

Relevant European References and Specifications

For the European Union context, non-binding technical guidelines and good practices are available for implementing the legal provisions of the INSPIRE Directive. Technical specifications are made available for each standard, which enable data providers to choose a particular solution based on the specific needs and concrete use cases. The governance of the technical specifications is ensured by the INSPIRE Maintenance and Implementation group (MIG), and its permanent technical sub-group (MIG-T).

9 Appendix IV: Methodology

9.1 Objective of the research

The objective of the research is to identify, specify, and document standards, various open-source and commercially available implementations and services in smart cities and communities domain for each technical building blocks of OpenDEI framework: Data Interoperability Building Blocks, Data Sovereignty and Trust Building Blocks, Data Value Creation Building Blocks and Data Spaces Governance Building Blocks for the creation of the Catalogue of Specifications.

9.2 Results

As a result of data collection and research, the Catalogue of Specifications in form of this deliverable D3.1 Catalogue of Specifications is created as collection of existing building blocks in the SSCC domain, including specifications (description) and reference implementations, classified according to the taxonomy proposed in OpenDEI Design Principles for Data Spaces Position Paper and leveraging on Data Spaces Business Alliance Technical Convergence document and in alignment with MIMs. The Catalogue will follow the same structure and templates as in DSSC, when available, with functional description, baseline standards and initiatives and available implementations.

9.3 Data Collection

9.3.1 Data collection research methods and process flow

Data collection research for Catalogue of specifications includes the following methods:

- Execute desk research to identify existing open source and commercially available standards and building blocks
- Perform an online survey (questionnaire)
- Perform 1-on-1 interviews with selected stakeholders for detailed information
- Workshops will be used for identification, verification, discussion about identified examples of best practice and gaps for alignment with main stakeholders and Stakeholders Forum
- Validation workshops and events will be used to present and validate Catalogue of Specifications with Stakeholders Forum, a broad group of stakeholders and the general public.

The process of data collection has resulted in the creation of a Catalogue of Specifications as Deliverable D3.1 Catalogue of Specifications.

In collaboration with WP2, WP4, and WP5 the deliverable D3.1 will be validated and presented at validation workshops and events.

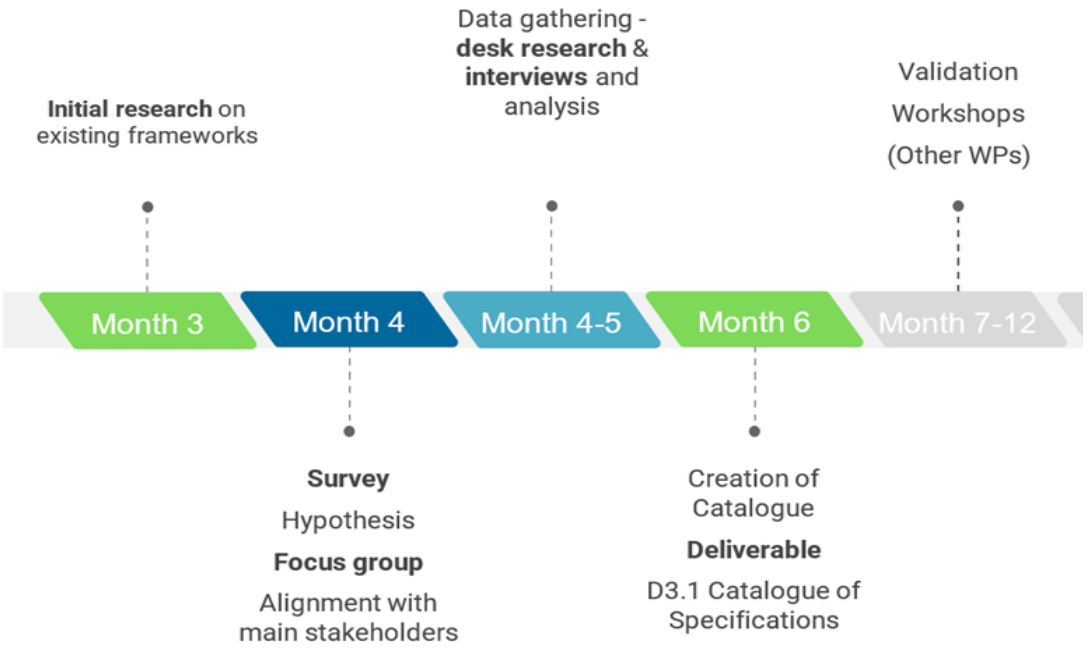

Figure 7 - D3.1 Data collection process flow

Initial desk research on existing frameworks

Initial desk research on existing initiatives, standards and frameworks has identified standards, open-source and commercially available implementations in smart cities and communities domain for each technical building blocks of OPENDEI framework, adopted by the stakeholders, to capture existing knowledge and identify gaps and needs. Desk research resulted in preliminary overview of candidate Data Space building blocks which have been validated and complemented by other research methods and mapped to the existing practise in the SSCC domain in four sections according to OpenDEI and OASC MIMs:

- Data Interoperability
  - Data Models and Formats
  - Data Exchange APIs
  - Provenance and Traceability
- Data Sovereignty and Trust
  - Identity management
  - Trusted exchange
  - Access & usage control / policies
- Data Value Creation
  - Metadata & Discovery protocol
  - Publication & Marketplace Services

    - Data Usage Accounting
- Data Spaces Governance
    - Business Agreements
    - Operational Agreements
    - Organizational Agreements

Each building block details will feature following structure:

- functional description,
- baseline standards and specifications
- available implementations

Proposed template for the Catalogue data is:

Metadata	Description
Name	Name of the BB
Description	Functional description of what the BB represents
Related standards	Set of standards that are relevant to implement the BB. Includes name, relevance, publisher, link to source for each identified standard.
Related specifications	Set of specifications that are relevant to implement the BB. Includes name, relevance, publisher, link to source for each identified specification.
Reference implementations	Set of available implementations that are relevant to implement the BB. Includes (brand name, provider, link to reference page, link to code repository (if OS) for each identified implementation.
Implemented MIMs	List of 10 OASC MIMs allowing multiple selection
Recommended by DSSC	Indicates if this BB is also recommended by DSSC (yes or no)
Scope	Indicates which domain the BB is applicable or if it is of generic purpose
Use Cases references	List of selected use cases (in WP4) that are using this BB, as a matter of examples
Maturity level	Level of maturity of the BB established in three degrees: quite mature, evolving and few mature

Table 6 - Template for BB description

Survey

A survey has been conducted to capture existing knowledge and to identify standards around Data Interoperability, Data Sovereignty and Trust, Data Value Creation and Data Governance and Legal Building Blocks for SSCC data spaces and various open-source and commercially available implementations and services for each technical building block. The survey was available online during the research. The list of survey questions is included in Appendix 9.5.1 Survey questions.

Figure 8 - Web page to access the DS4SSCC survey

Interviews

To complement Desk research, 1:1 Interviews with selected experts were conducted to capture existing knowledge and to identify Data Interoperability Building Blocks, Technical Data Sovereignty and Trust Building Blocks, Data Value Creation Building Blocks and Data Governance and Legal Building Blocks for SSCC data spaces and various open-source and commercially available implementations and services for each technical building block with special focus on gaps identification and having in mind large and small cities and communities as well as their digital transformation maturity.

Guidelines for interviews

1:1 Interviews with selected stakeholders will capture data for Catalogue. Guidelines for conducting interviews are in Appendix 9.5.2 Guidelines for interviews.

Workshops and events

Workshops were used for verification and discussion about identified and emerging examples for alignment with main stakeholders as well as to capture participants' existing knowledge as well as to identify their needs to complement the former data collection methods. A workshop will be developed in several steps:

- Preparation and creation of online board collaboration environment, when appropriate
- The workshop execution (Hybrid, Online, Offline)
- Presentations in powerpoint for the workshop with information about the project and instructions for the workshop,
- Outcomes of the workshop.

Scenario of the workshop:

Steps	Duration	Topics/Observation
Invitation, instructions for workshop	Distribute prior the workshop	Distribute 1-Page overview of DS4SSCC and Purpose of workshop, invitation
Open the session	5 minutes	State of play Explain purpose of the workshop, the results to be developed and the agenda
Tour de Table (If number of participants allow)	5 minutes	Name, Company and Job Role (Keep the introduction very brief)
Description of the main topic	20 minutes	Optional introduction of main topic to enable efficient work if needed
Work on main topics	60 minutes (4x15 mins)	Intro & instructions, operational details Each station (online, physical) will have a facilitator and subject matter expert on section to moderate the discussion and help participants.

		Participants are divided in 4 groups & circulate between the section stations. Participants provide inputs by post-it stickers in Miro board or physical board. Depending on the main topic. Outputs are digitized for further reference.
Wrap-up	10 minutes	Review the outputs of the workshop, clarify the use of outputs gathered and remove any areas of confusion
Total Time	80-90 minutes	

Table 7 - workshop structure

Example guidelines for workshop on data spaces building blocs

Workshop will be used for verification and discussion about identified examples for alignment with main stakeholders for Catalogue of specifications according to four sections of OpenDEI architecture. Miro boards were used for collaborative work.

- Data Interoperability
  - Data Models and Formats
  - Data Exchange APIs
  - Provenance and Traceability
- Data Sovereignty and Trust
  - Identity management
  - Trusted exchange
  - Access & usage control / policies
- Data Value Creation
  - Metadata & Discovery protocol
  - Publication & Marketplace Services
  - Data Usage Accounting
- Data Spaces Governance
  - Business Agreements
  - Operational Agreements
  - Organizational Agreements

Workshops will also be organized to validate the Catalogue of Specifications by the Stakeholders Forum. Partnership will organize seven workshops and events with the Stakeholders Forum and other stakeholders focusing on different topics to collect

the feedback from the community of practice during the creation of project outputs, starting in January 2023 and concluding in September 2023.

The work of DS4SSCC is being presented at several events to present D3.1. Catalogue of Specifications to the broadest group of stakeholders and the general public.

9.3.2 Data Metrics and geographic scope

	Survey	**Interviews**	**Focus Groups and workshops**	**Countries represented**
Planned number of actors involved	35	10	2	15
Achieved number of actors involved	46	10 (+2 from WP2)	2	18

Table 8 - Data collection metrics for technical BB

	Government, Public administration	**Industry**	**SME**	**Civil Society**	**Research Institutes & Academia**
Planned stakeholders involved	15	5	15	5	10
Achieved stakeholders involved	17	5	12	2	10

Table 9 - Categories of involved stakeholders for technical BB

	Data owner	**Data provider**	**Data user**	**Data intermediary**	**Software provider**	**Data platform provider**
Planned stakeholders roles involved	50	50	50	30	30	30
Achieved stakeholders roles involved	59	53	64	39	30	44

Table 10 - Roles of involved stakeholders overall (not just technical ones)

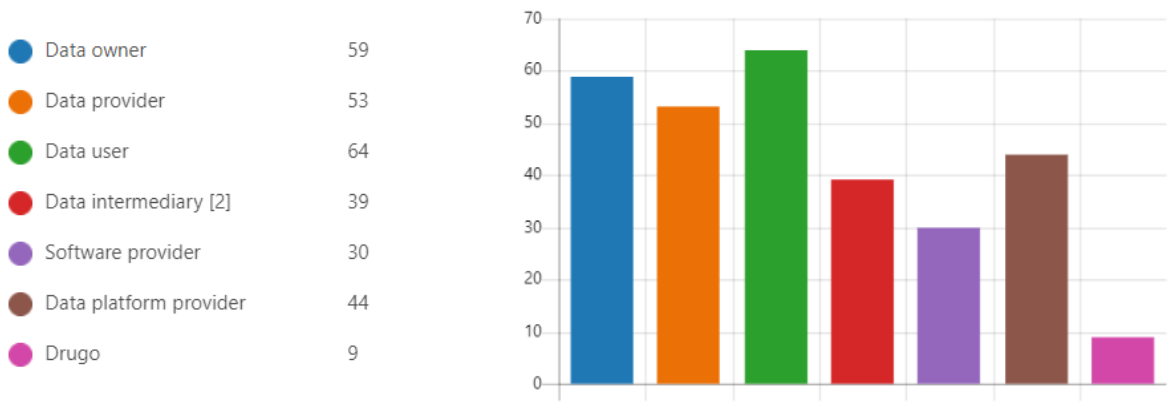

Figure 9 - Graph of involved stakeholders overall

9.3.3 Partner roles and responsibilities

For the development of deliverable D3.1 Catalogue of Specifications partners will organize work in tasks T3.1 – T3.4 with following responsibilities:

Task	Lead	Contributors
Data Interoperability Building Blocks (T3.1)	FIWARE	CCIS
Technical Data Sovereignty and Trust Building Blocks (T3.2)	OASC	FIWARE CCIS
Data Value Creation Building Blocks (T3.3)	FIWARE	OASC
Data Governance and Legal Building Blocks (T3.4)	OASC	EUC CCIS

Table 11 - Partner roles and responsibilities

9.4 Stakeholders engagement and timeline

Identification of the main stakeholders, their roles as well as interests has been identified and will be revised during the project as the work progresses. Depending on stakeholders' roles, several touchpoints are envisioned for each target group.

Data collection methods/Stakeholders touchpoints																		
Category	**Survey**			**Desk research**			**Interviews**			**Focus Group**			**Workshops**			**Events**		
	Perform an online survey (questionnaire) to gather data			Identify and collect data for Catalogue of specifications (using data from WP2 & Survey, if available)			1:1 Interviews to complement the survey inputs and desk research			Focus group to complement the survey inputs, interviews and desk research; more experts at the same time; ideally to reconfirm outputs from survey, desk research & interviews			Validation Workshops			For example: presentation in FIWARE Global Summit in June		
	1 common questionnaire, coordinated questions WP2/WP3/WP5			Template			Task leaders to decide when to combine Desk research with 1:1 interview			Physical or online; but not the hybrid; Miro as means to cocreate when online; prints/post-it when physical 1 planned			Cocreation with other WPs			Cocreation with other WPs		
	Stakeholders involved	When	Who	Stakeholders involved	When	Who	Stakeholders involved	When	Who	Stakeholders involved	When	Who	Stakeholders involved	When	Who	Stakeholders involved	When	Who
Policy makers	x	M4	WP2&5													x	M7-12	WP5
Cities, Municipalities, Regions	x	M4	WP2&5	x	M3-4	WP3	x	M4-5	WP3	x	M3-4	WP3	x	M7-12	WP5	x	M7-12	WP5

Technical and R&D Community	x	M4	WP2&5	x	M3-4	WP3	x	M4-5	WP3	x	M3-4	WP3	x	M7-12	WP5	x	M7-12	WP5
Civil Society and NGOs	x	M4	WP2&5										x	M7-12	WP5	x	M7-12	WP5
Industry & businesses	x	M4	WP2&5	x	M3-4	WP3	x	M4-5	WP3	x	M3-4	WP3	x	M7-12	WP5	x	M7-12	WP5
Standardization organizations	x	M4	WP2&5	x	M3-4	WP3							x	M7-12	WP5	x	M7-12	WP5
European alliances and associations dealing with dataspaces and AI	x	M4	WP2&5	x	M3-4	WP3				x	M3-4	WP3	x	M7-12	WP5	x	M7-12	WP5
European funded projects and policies	x	M4	WP2&5	x	M3-4	WP3										x	M7-12	WP5

General public	x	M4	WP2&5												x	M7-12	WP5

Table 12 - Data collection methods/Stakeholders touchpoints

9.5 Survey and interview questions

9.5.1 Survey questions

Data interoperability

1. Are you familiar with data models and formats? Required to answer.
   - Yes
   - No
2. Which standards and specifications for data models and formats are being used/developed by your organisation?
3. Which types of specifications are available for data models and formats?
   - Open
   - Proprietary
   - None
4. Can you please provide the name/brand and the description of this type of implementation? Please add a link to the public code repository in case of OS and link to reference in case of proprietary one.
5. Are there any new standards, specifications or results from data space initiatives (e.g., DSBA) that you would be considering for data models and formats?
6. Have you already used these standards or implementations in specific use-cases?
   - Yes
   - No
7. Please provide some use cases where they have been used and list any missing functionality you may have identified.
8. Are you familiar with Data Exchange APIs?
   - Yes
   - No
9. Which specifications for Data Exchange APIs are being used/developed by your organisation?
10. Which types of implementations are available for Data Exchange APIs
    - Open source
    - Proprietary
    - None
11. Are there any new standards, specifications, or results from data space initiatives (e.g., DSBA) that you would be considering for Data Exchange API?
12. Have you already used these standards or implementations in specific use-cases?
    - Yes
    - No
13. Please provide some use cases where they have been used and list any missing functionality you may have identified.
14. Are you using any processes for Provenance and Traceability?
    - Yes
    - No
15. Which types of processes are available for Provenance and Traceability?
    - Open source
    - Proprietary
    - None
16. Can you please provide the name/brand and the description of this type of implementation? Please add link to public code repository in case of OS and link to reference in case of proprietary one.
17. Are there any new standards, specifications or results from data space initiatives (e.g., DSBA) that you would be considering for Provenance and Traceability?
18. Have you already used these standards or implementations in specific use-cases?
    - Yes
    - No
19. Please provide some use cases where they have been used and list any missing functionality you may have identified.

Data Sovereignty & Trust

20. Are you familiar with identity management?
    - Yes
    - No
21. Which standards and specifications for identity management are being used/developed by your organisation?
22. 41.Which types of standards and specifications are available for identity management?
    - Open source
    - Proprietary
    - None
23. Please provide the name/brand and the description of this type of implementation. Please add a link to the public code repository in case of OS and a link to reference in case of a proprietary one
24. Are there any new standards, specifications or results from data space initiatives (e.g., DSBA) that you would be considering for identity management?
25. Have you already used these standards or implementations in specific use-cases?
    - Yes
    - No
26. Please provide some use cases where they have been used and list any missing functionality you may have identified.
27. Are you familiar with access & usage control / policies?
    - Yes
    - No
28. Which standards and specifications for access & usage control/policies are being used/developed by your organisation?

Which types of standards and specifications are available for access & usage control/policies?
    - Open source
    - Proprietary
    - None
29. Please provide the name/brand and the description of this type of implementation. Please add a link to the public code repository in case of OS and a link to reference in case of a proprietary one.
30. Are there any new standards, specifications or results from data space initiatives (e.g., DSBA) that you would be considering for access & usage control/policies?
31. Have you already used these standards or implementations in specific use-cases?
    - Yes
    - No
32. Please provide some use cases where they have been used and list any missing functionality you may have identified.
33. Are you familiar with trusted exchange?
    - Yes
    - No
34. Which standards and specifications for trusted exchange are being used/developed by your organisation?
35. Which types of standards and specifications are available for trusted exchange?
    - Open source
    - Proprietary
    - None
36. Please provide the name/brand and the description of this type of implementation. Please add a link to the public code repository in case of OS and a link to reference in case of a proprietary one.
37. Are there any new standards, specifications or results from data space initiatives (e.g., DSBA) that you would be considering for trusted exchange?
38. Have you already used these standards or implementations in specific use-cases?
    - Yes
    - No
39. Please provide some use cases where they have been used and list any missing functionality you may have identified.

Data value Creation

40. Are you familiar with Metadata & Discovery protocol?
    - Yes

- No

41. Which standards and specifications for Metadata & Discovery protocol are being used/developed by your organisation?
42. Which types of standards and specifications are available for Metadata & Discovery protocol?
    - Open source
    - Proprietary
    - None
43. Please provide the name/brand and the description of this type of implementation. Please add a link to the public code repository in case of OS and a link to reference in case of a proprietary one.
44. Are there any new standards, specifications or results from data space initiatives (e.g., DSBA) that you would be considering for Metadata & Discovery protocol?
45. Have you already used these standards or implementations in specific use-cases?
    - Yes
    - No
46. Please provide some use cases where they have been used and list any missing functionality you may have identified.
47. Are you familiar with Data Usage Accounting?
    - Yes
    - No
48. Which standards and specifications for Data Usage Accounting are being used/developed by your organisation?
49. Which types of specifications are available for Data Usage Accounting?
    - Open source
    - Proprietary
    - None
50. Please provide the name/brand and the description of this type of implementation. Please add a link to the public code repository in case of OS and a link to reference in case of a proprietary one.
51. Are there any new standards, specifications or results from data space initiatives (e.g., DSBA) that you would be considering for Data Usage Accounting?
52. Have you already used these standards or implementations in specific use-cases?
    - Yes
    - No
53. Please provide some use cases where they have been used and list any missing functionality you may have identified.
54. Are you familiar with Publication & Marketplace Services?
    - Yes
    - No
55. Which standards and specifications for Publication & Marketplace Services are being used/developed by your organisation?
56. Please provide the name/brand and the description of this type of implementation. Please add a link to the public code repository in case of OS and a link to reference in case of a proprietary one.
57. Are there any new standards, specifications or results from data space initiatives (e.g., DSBA) that you would be considering for Publication & Marketplace services?
58. Have you already used these standards or implementations in specific use-cases?
    - Yes
    - No

79.Please provide some use cases where they have been used and list any missing functionality you may have identified.

Data sharing models and governance

59. How does your organisation share data with other stakeholders? Please select all that apply.
    - Open data
    - Data marketplaces
    - Personal Data Stores
    - Data brokers or trusted third parties
    - Public administration partnership/agreements
    - Business to Government (B2G) partnership/agreements

- Business to Business (B2B) partnership/agreements

60. How does your organisation acquire data held by other stakeholders? Select the three most relevant of the followings:
    - Open data
    - Data marketplaces
    - Personal Data Stores
    - Data brokers or trusted third parties
    - Public administration partnership/agreements
    - Business to Government (B2G) partnership/agreements
    - Business to Business (B2B) partnership/agreements
    - None of the above
61. What types of contractual agreements does your organisation use the most to exchange data with other stakeholders?
    - Multilateral data sharing agreements
    - Bilateral data sharing agreements
    - Service Level Agreements
    - Public procurement of data
    - Data purchase agreements
62. Which protocols for data management, if any, does your organisation use (e.g. DAMA DMBoK)?
63. Which tools, processes and/or practices related to data quality assurance has your organisation implemented?
64. What are the main barriers faced by your organisation when sharing data with external partners?
65. Where do you see the data sharing and access opportunities for your organisation in terms of a European data space for smart and sustainable cities and communities?

9.5.2 Guidelines for interviews

9.5.2.1 Interview guide - Governance Experts

Introduction project/ aim of interviews:

Brief background of the interview participant:

- Affiliation
- Position
- Years of experience

Questions	Notes
From your research/ experience, what key lessons have you learnt in terms of sharing data between different stakeholders?	
What key roles have you identified for data sharing/data stewardship?	

What are the main enablers/barriers in terms of data sharing between the public sector and private sector? With civil society?	
How do you ensure trust in local data ecosystems?	
How do we ensure interoperability, not only at the technical level?	
Is there a specific business model that you would recommend?	
What types of contractual agreements enable data flows between different stakeholders? What are the advantages/challenges there?	
Are you aware of specific mechanisms (legal/technical/organisational) used by local stakeholders to support different levels of access rights and identity management?	
Are you aware of any frameworks/good practices in terms of data governance/data sharing between different types of stakeholders that the DS4SS could draw on?	

9.5.2.2 Interview guide - Technical managers at supply side

Introduction project/ aim of interviews:

Brief background of the interview participant:

- Affiliation
- Position
- Years of experience

Questions Notes

Which type of data (formats, models...) are you using/collecting in your city?	
Do you know the standards are you using for data modelling?	
Which standardized APIs are you using to exchange data within the city or with external entities to the city?	
Are you following any process for ensuring the traceability and provenance of your data? Could you describe a bit?	
Which IAM standards are you relying on?	
Which discovery and publication of data standards/mechanisms are you using?	
Are you following any standard for accounting the access of the users to the data? Which one?	
Which mechanism are you following to provide access to the data? Have you implemented any marketplace? Did you use an open implementation?	
Do you know what are the MIMs? Which MIMs is your city/solution implementing?	
Any concrete reference implementation are you using for any of the functional blocks commented above?	

9.5.2.3 Interview guide - Local data ecosystem stakeholders

Introduction project/ aim of interviews:

Brief background of the interview participant:

- Affiliation
- Position
- Years of experience

Questions	Notes
How did the ecosystem/data partnership start? / What was the initial use-case?	
Who are the key stakeholders? / What are their roles in the ecosystem? *Who is coordinating/orchestrating it? How does it work in terms of decision-making?*	
What is the business model of the ecosystem/partnership? *Does the data have licensing fees?*	
What are the benefits of the ecosystem?	
What are the ambitions (short/long term)?	
What were the main challenges? What issues/questions were raised during the set-up of the data ecosystem?	
What are the incentives for participation?	
How are newcomers/third parties managed?	
What type of data is involved?	

How is data accessed/shared? (Different data access rights?)	
What are the data standards used?	
How is data quality ensured? *How is the monitoring or auditing of data use managed?*	
What are the operating costs? How are these covered?	
What types of contractual frameworks are used? SLA/data sharing agreements, etc?	
How is the sustainability of the ecosystem ensured?	

About Data Space for Smart and Sustainable Cities and Communities (DS4SSCC)

Data is a central aspect of the twin green and digital transformation, and European cities, regions, towns, and rural areas play a vital role in safely leveraging its potential. This preparatory action for a Data Space for Sustainable and Smart Cities and Communities (DS4SCC) provides a coordinated starting point for public, private, and individual stakeholders to contribute and use data, aligned with European values and policies.

This preparatory action emphasises the sustainability aspect – green, social, and economic – and the diversity of communities, and aims to:

- Develop a multi-stakeholder data governance scheme by bringing together European cities and their local stakeholders ('quadruple helix') to collaborate on use cases relevant to Green Deal objectives through operational local data governance core group".
- Create a blueprint for the European DS4SSCC by co-creating with stakeholders a methodology for setting it up, from the vision of a full-fledged pan-EU DS4SSCC, not only from a technical perspective but also giving operational guidance e.g., for procurement.
- Bring an agreed set of priority datasets into conformity with the new blueprint by delivering a catalogue of domains, use cases and related data sets for DS4SSCC.
- Develop a roadmap and action plan towards a mature, connected pan-EUDS4SSCC.
- Shape and implement the data space on the local, regional, national and EU levels, taking into account their different levels of maturity, will be an exercise in co-creation with the stakeholder forum.

Documentation will include recommended actions for standardisation, business models and strategies for running data spaces, and a vision for the federation of platforms. Building on core European networks of cities and communities that have championed the Living-in.EU movement, DS4SSCC is a timely, ambitious, and essential contribution towards the sustainability goals of European citizens.

Our consortium:

OPEN & AGILE SMART CITIES | EUROCITIES | European Network of Living Labs


FIWARE
FOUNDATION


AUSTRIAN INSTITUTE
OF TECHNOLOGY


Gospodarska
zbornica
Slovenije
Chamber of Commerce
and Industry of Slovenia