DATA SPACE FOR
SMART AND SUSTAINABLE
CITIES AND COMMUNITIES

Deliverable D3.2

Architecture Model

WP 3 – Technical Blueprint

Authors: Clara Pezuela (FIWARE), Chandra Challagonda (FIWARE), Gaber Terseglav (CCIS), Ines Vlahovic (CCIS), Martin Traunmuller (AIT),Ghazal Etminan(AIT), Thimo Thoeye (OASC), Michael Mulquin (OASC)

Reviewers: Justine Gangneux (EuroCities), Sophie Meszaros (OASC), Ghazal Etminan (AIT), Martin Brynskov (OASC), Martine Delannoy (Digital Flanders)

Delivery date: 30/09/2023

Dissemination level: Public

Type: Report

Funded by
the European Unio

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 (FIWARE) Chandra Challagonda (FIWARE)	Table of Contents	05/05/2023
Clara Pezuela (FIWARE) Ines Vlahovic, Gaber Terseglav (CCIS)	Draft deliverable (sections 1 and 2)	16/05/2023
Chandra Challagonda (FIWARE)	Draft Architecture	04/06/2023
Chandra Challagonda (FIWARE) Gaber Terseglav (CCIS)	Components description	30/06/2023
Clara Pezuela (FIWARE)	Customise architecture templates	24/07/2023
Chandra Challagonda (FIWARE) Gaber Terseglav (CCIS)	Pending contributions	27/08/2023
Clara Pezuela (FIWARE) and all contributors	Full draft complete	31/08/2023
Clara Pezuela (FIWARE) Ines Vlahovic (CCIS) Martin Traunmüller (AIT) Thimo Thoeye (OASC)	Integration of use cases and stakeholders feedback	15/09/2023
Justine Gangneux (EuroCities)	Review	21/09/2023

Sophie Meszaros (OASC)	Review	22/09/2023
Martine Delannoy (Digital Flanders)	Review	26/09/2023
Ghazal Etiman (AIT)	Review	28/09/2023
Clara Pezuela (FIWARE)	Integration of comments from review and generation of final version for submission	28/09/2023
Sophie Meszaros (OASC)	Final Review	29/09/2023
Martin Brynskov (OASC)	Final Review	29/09/2023

Abbreviations

WP	*Work Package*	*SDG*	*Sustainable Development Goals*
BB	*Building Block*	*KPI*	*Key Performance Indicator*
MIM	*Minimal Interoperable Mechanism*	*DSSC*	*Data Spaces Support Center*
SSCC	*Smart and Sustainable Cities and Communities*	*EU*	*European Union*
DSBA	*Data Spaces Business Alliance*	*FAQ*	*Frequently Asked Question*
GDPR	*General Data Protection Regulation*		

Table of Contents

Executive Summary

This document provides the Reference Architecture Model and the Cookbook for the deployment of the sustainable and smart cities and communities data space.

While the previous deliverable of this WP, the D3.1, provided a complete Catalogue of Specifications for the Building Blocks (BBs) to be used in the design and deployment of a data space in cities and communities domain, this document defines a reference architecture which explains how to use those BBs in a coherent and consistent manner to develop the functionality which is expected to be provided by a data space.

After the analysis of several architecture methods, none of them has been fully applied. Instead, a more pragmatic approach has been opted based on use cases driven, but without missing any important element in the description of the architecture. However, we have analysed the most relevant reference architectures for data spaces and smart cities to gather from each of them the most relevant insights for the DS4SSCC architecture. Additionally, existing European regulation and legislative frameworks have been analysed to determine their impact on the DS4SSCC architecture, like for example the Green Deal Objectives.

The designed architecture takes into account the basic design principles for a data space (interoperability, sovereignty, ecosystem, security and decentralisation), the specific considerations of the smart cities domain and the existing Minimal Interoperability Mechanisms Plus (MIMs Plus) already adopted in the field. The proposed architecture is fully aligned with the recommendations given by the Data Spaces Support Centre (DSSC), so it follows the spirit, concepts and taxonomy of building blocks provided by the DSSC as common technical grounds for all the data spaces.

The current architecture is considering three typical scenarios, actually deployed in the cities and communities nowadays:

- existing data platform which integrates different verticals or some digitised vertical services in isolation (called brownfield),
- no digital infrastructure at all (referred as greenfield)
- and virtual representation of the real world for monitoring and simulation (usually a Digital Twin).

This document presents how these scenarios should evolve into a data space by defining a high level architecture that would fit with them. This architecture proposes the incorporation of three main components to the existing data platforms to naturally become data spaces. These components are: Universal Trust Data Registry, Authorization Policies Store and Federated Layer.

Four use cases (Amsterdam, Helsinki, Flanders and Valencia) were selected to customise the high level architecture and thereby demonstrate the blueprint's use in

different situations and share the steps that are required for adopting it. Each use case is accompanied by a Data Cooperation Canvas, which lists a set of technical requirements, the customised architecture and the list of implementation rules. The Data Cooperation Canvas is a detailed description of data cooperation at use-case level including business models adopted, governance structures, technical infrastructure and data used and implementation components.

Finally, the document also provides a set of short guidelines, recipes per type of scenario and the validation of the architecture by the use cases and the stakeholders. The main take away from the preparatory action is that while the proposed architecture is a good starting point, further development and more detailed analysis and implementation are required in order to make a real deployment of a data space.

This architecture will be further developed and deployed in the upcoming deployment phase of the data space by the project ‘Data Space for Smart Cities and Communities Deployment’ (DS4SSCC-DEP), starting on 1st October 2023.

The present document is addressed for data space architects and software designers who will use this document as a reference source for designing and building their data spaces.

1 Introduction

This deliverable presents the Reference Architecture Model for the Smart and Sustainable Cities and Communities Data Space (DS4SSCC) (in section 3). The model is complemented with an extensive state of the art about existing reference architectures, regulations and initiatives which can be used as baseline work (section 2), an example of application of the architecture in a concrete use case (section 4) and a set of guidelines and recipes (CookBook) to facilitate the use of the model (section 5). Section 6 summarises the main conclusions of the work and establishes the next steps in the deployment of the data space. Section 1 helps in framing the scope of the work and explains the method for building the architecture.

This document describes the work carried out by the WP3 Technical Blueprint during the second half of the project (M6-M12) after the delivery of the D3.1 Catalogue of Specifications delivered at M6. The D3.2 required inputs from WP4 Data Space Establishment and from WP2 Development of a multi-stakeholder data governance scheme. WP4 provided the description of the selected use cases as representative scenarios to generate the adapted architectures from the generic high-level architecture and specific recipes. One of these use cases has been also used for the validation and exemplification of the architecture model. WP2 provided the Data Cooperation Canvas (D2.2 Multi Stakeholders Governance Schema) in which they have provided the elements from the governance, legal and business which affect the technical decisions to be taken into the architecture design.

1.1 DS4SSCC Blueprint and WP3

Following the definition of the blueprint provided by the Data Spaces Support Center (DSSC) in their glossary[1], a *Data Space Blueprint is a consistent, coherent and comprehensive set of guidelines to support the implementation, deployment and maintenance of data spaces*. Thus, the DS4SSCC blueprint defines the guidelines and mechanisms required for the upcoming deployment of the data space. Although this document is aiming at describing the DS4SSCC blueprint, the required elements are spread across other work packages and deliverables in the project (D2.2 and D4.2). The picture below shows all the elements which form part of the blueprint.

[1] https://dssc.eu/wp-content/uploads/2023/03/DSSC-Data-Spaces-Glossary-v1.0.pdf

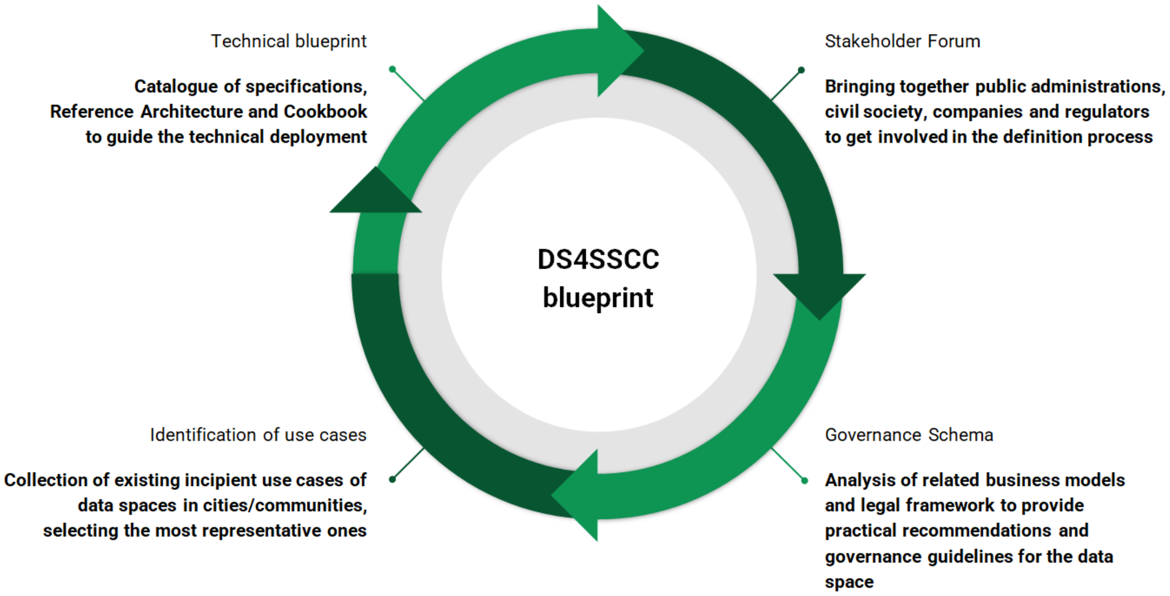

Technical blueprint
Catalogue of specifications, Reference Architecture and Cookbook to guide the technical deployment
Stakeholder Forum
Bringing together public administrations, civil society, companies and regulators to get involved in the definition process
DS4SSCC blueprint
Identification of use cases
Collection of existing incipient use cases of data spaces in cities/communities, selecting the most representative ones
Governance Schema
Analysis of related business models and legal framework to provide practical recommendations and governance guidelines for the data space

Figure 1. DS4SSCC Blueprint

Furthermore, the stakeholder forum is an important part of the data space ecosystem because it brings together relevant players in data spaces. This allows for collaboration and coordination between stakeholders, which can help ensure the success of the data space deployment. The *Stakeholder Forum* is the emerging **data space ecosystem** which is being built and coordinated by WP5 and WP1. It relies on the Living-in.eu movement[2] expanded with relevant players in data spaces.

The *Governance Schema* built under the coordination of WP2 is establishing the **rules** to govern the data space ecosystem, aiming at the mutual benefit of participants.

The *Technical Blueprint* developed by the WP3 contributes to the overall blueprint with the **Catalogue of Specifications of Building Blocks**, **Reference Architecture** and **CookBook** to deploy the technical infrastructure for the data space.

WP4 has identified the relevant and representative *Use Cases* in Europe that may bring tangible examples of incipient data spaces. The selected use cases are showing their **priority data sets** and commonly used technologies to bring to the data space.

Therefore, the DS4SSCC blueprint is formed by all these elements above mentioned (the ecosystem, the data, the governance, the technology) and all need to be used and followed to deploy the data space for smart and sustainable cities.

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

1.2 Scope of the document

While the previous deliverable of this WP, the D3.1, provided a complete Catalogue of Specifications for the Building Blocks (BBs) to be used in the design and deployment of a data space in cities and communities domain, this document intends to define a reference architecture which explains how to use those BBs in a coherent and consistent manner to develop the functionality which is expected to be provided by a data space. This exercise represents a great challenge due to several reasons:

- Many cities and communities have already started their processes of digitalization and they already have in place Urban Data Platforms[3] to collect and manage their data and data coming from their suppliers (for example, the four use cases selected in this document). Thus, existing architectures are already defined for their systems. The proposed architecture must be defined considering the existing platforms, and as evolution of them.
- There are already some existing proposals of architecture for data spaces (DSBA Technical Convergence, IDS-RAM…) to take as reference (see Section 2.1). Thus, they need to be analysed and adapted to the cities and communities context.
- European regulation is providing a legal framework for data sharing and interoperability. Thus, the proposed architecture must be compliant with all of this, and even mandatorily fulfil some of the recommendations, especially if they are addressing public entities.
- Data spaces in the context of cities and communities have some specificities which need to be considered to define the best architecture for them, for example the large amount of open data and the existing data platforms they already have. The data space ecosystem is a mix of different types of entities with diverse interests and a special type of business model, not only focused on data monetization. Besides, cities and communities should be the main beneficiaries and that data sharing in the context of DS4SSCC should align with the Green Deal Objectives. Additionally, the architecture needs to fulfil the wide adopted conceptual framework provided by the Minimal Interoperability Mechanisms (MIMs).
- The built architecture model has to be easy to use and to understand. Thus, several methodologies have been explored aiming at selecting the best from each one in favour of simplicity but also of the minimal formality.

The following section details the above mentioned challenges and how the project team has decided to approach each of them in the present document which describes the architecture for the SSCC data space.

[3] https://www.datavaults.eu/wp-content/uploads/2021/03/2019-Study-on-Urban-Data-Platforms-key-findings-6-3-2020.pdf

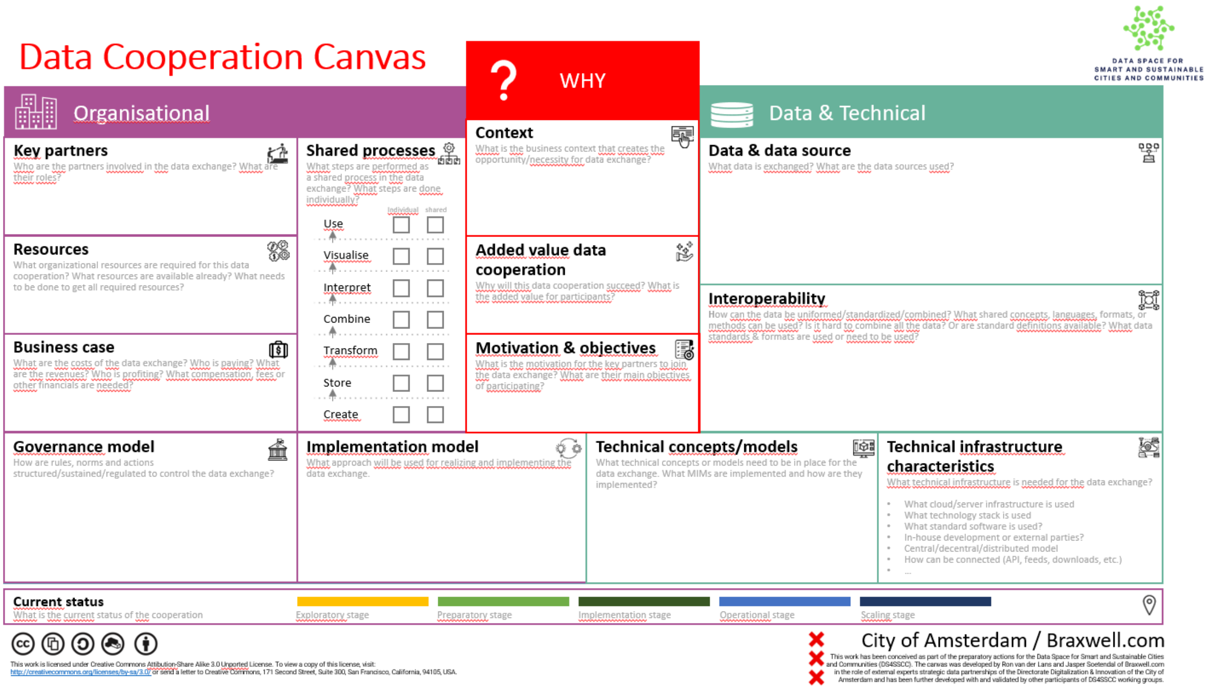
The present document is addressed for data space architects and software designers who will use this document as a reference source for designing and building their data spaces. However, the introductory sections and the high level architecture overall description can be also of interest for policy makers and business developers to understand how the data space is built and which functionally is able to provide.

While the high level architecture has been customised into several instances, the document cannot address   all possible scenarios and data space particularities. So, every data space deployment will have to take this document as a reference and starting point and follow the steps described here to create their own instance of the architecture.

It is expected that the deployment project of the SSCC data space will leverage this document to define the overall architecture for the actual deployment phase. Similarly, the testing experimental facilities project (Citcom.ai) will use this document to plan the required infrastructure requirements for the deployment and testing of the SSCC data space.

1.3 Data Cooperation Canvas

A common and transversal instrument was developed in the project to condense into ‘at a glance’ the motivation, the governance and the technical dimensions of a data space. By reusing the concept of the Business Model Canvas from Osterwalder, DS4SSCC provides the template below developed and validated by WP2 (for more information, please refer to D.2.2 Multi-stakeholder governance scheme).



Data Cooperation Canvas

? WHY

DATA SPACE FOR SMART AND SUSTAINABLE CITIES AND COMMUNITIES

Organisational

Key partners
Who are the partners involved in the data exchange? What are their roles?

Resources
What organizational resources are required for this data cooperation? What resources are available already? What needs to be done to get all required resources?

Business case
What are the costs of the data exchange? Who is paying? What are the revenues? Who is profiting? What compensation, fees or other financials are needed?

Shared processes
What steps are performed as a shared process in the data exchange? What steps are done individually?

	Individual	shared
Use	□	□
Visualise	□	□
Interpret	□	□
Combine	□	□
Transform	□	□
Store	□	□
Create	□	□

Context
What is the business context that creates the opportunity/necessity for data exchange?

Added value data cooperation
Why will this data cooperation succeed? What is the added value for participants?

Motivation & objectives
What is the motivation for the key partners to join the data exchange? What are their main objectives of participating?

Data & Technical

Data & data source
What data is exchanged? What are the data sources used?

Interoperability
How can the data be uniformed/standardized/combined? What shared concepts, languages, formats, or methods can be used? Is it hard to combine all the data? Or are standard definitions available? What data standards & formats are used or need to be used?

Governance model
How are rules, norms and actions structured/sustained/regulated to control the data exchange?

Implementation model
What approach will be used for realizing and implementing the data exchange.

Technical concepts/models
What technical concepts or models need to be in place for the data exchange. What MIMs are implemented and how are they implemented?

Technical infrastructure characteristics
What technical infrastructure is needed for the data exchange?

- What cloud/server infrastructure is used
- What technology stack is used
- What standard software is used?
- In-house development or external parties?
- Central/decentral/distributed model
- How can be connected (API, feeds, downloads, etc.)
- ...

Current status
What is the current status of the cooperation

🟨 Exploratory stage | 🟩 Preparatory stage | 🟩 Implementation stage | 🟦 Operational stage | 🟦 Scaling stage

City of Amsterdam / Braxwell.com

This work has been conceived as part of the preparatory actions for the Data Space for Smart and Sustainable Cities and Communities (DS4SSCC). The canvas was developed by Ron van der Lans and Jasper Soetendal of Braxwell.com in the role of external experts strategic data partnerships of the Directorate Digitalization & Innovation of the City of Amsterdam and has been further developed with and validated by other participants of DS4SSCC working groups.

Figure 2. Data Cooperation Canvas template

DS4SSCC, in collaboration with WP2, WP3 and WP4, filled out this template for the selected use cases to validate the architecture and they can be found in section 3.5. WP3 completed the Data and Technical part. It consists of 4 sections:

- Data & data sources, where the city indicates which are the datasets collected digitally in the city and in which formats. In the context of DS4SSCC, this information has been collected by WP4 through the use cases analysis.
- Interoperability, where the city indicates the standards and other mechanisms put in place to make its data and services interoperable. In the context of DS4SSCC, the cities and communities have provided these inputs for the Catalogue of Specifications.
- Technical concepts/models, where the city indicates the MIMs that are considered in its digital systems. In the context of DS4SSCC, this information has been valuable to customise the architecture and define the steps to follow for evolving towards a data space.
- Technical infrastructure characteristics, where the city indicates the technologies and architectures used in its data platform. In the context of DS4SSCC, these initial architectures are the baseline to evolve them according to the proposed high level architecture.

1.4 Building process

The building process of the architecture was a collective activity across all the partners with a significant cooperation of the 4 use cases involved in the customization (Helsinki, Amsterdam, Flanders and Valencia). This section describes the following approach, the set of steps and all the considerations that have been taken into account in the depiction of the architecture.

1.4.1 Method to build the architecture

We analysed Togaf, Archimate and Data Sharing Coalition (DSC) use case development. The summary of findings is:

- TOGAF:
  - A very detailed and structured way to get an enterprise architecture.
  - Offers a step-by-step approach to developing Enterprise Architecture.
  - Reference: used in many cases including smart cities and communities.
  - Complex framework, training needed.
  - There is no mapping with Open DEI (despite Open DEI's use in the DSSC Blueprint).
  - Licence needed if used externally.
  - Suitable for an enterprise architecture, however adjustments needed for a guiding tool for a reference architecture.
- ArchiMate:
  - A tool to describe enterprise architecture.
  - As it is a tool, some time is needed for installing and training.
  - Complementary to TOGAF (a methodology still needed).
- DSC:
  - No special training needed. No licence needed.
  - Simple tool that guides to create a use case. Lots of details needed from use cases, most likely through workshops.
  - There is no mapping with Open DEI which will be used in DSSC Blueprint.
  - Focus on data sharing between data spaces.

Given the novelty of the data spaces concept, the listed methodologies are not suitable to define architecture points for evolving (data sharing) initiatives into data spaces.

Our method for building the reference architecture was the following:

1) Use case agnostic reference architecture:
   - Prepare the list of most used architectures in cities (current state).

- Prepare the list of the most important architecture topics when a city/community wants to move to data space (this information is gathered/confirmed also during the stakeholder forum).
- Based on the above listed topics and relevant documents (iShare, i4Trust, DSBA Technical convergence, Odala, FIWARE, Gaia-X, Open DEI…) we will prepare the architecture with the most important aspects.

2) Use case specific reference architecture:
   - For each selected use case, we will first fill out the Data Cooperation Canvas and then we will also get technical details of its current state from the stakeholder. Technical details will be provided either in a workshop or through filling in the form. We prepared a template for a workshop/survey in chapter 3.5.1. Questionnaire.
   - Using current state and agnostic reference architecture, we will prepare use case specific reference architecture, where we will:
     - o Describe targeted use case reference architecture.
     - o propose technical guidance, rules, policies, protocols, standards (catalogue of specifications).
     - o Define what will be done (implementation guidelines).
     - o describe potential design decisions for each topic including rationale behind.
     - o Name and list all possible challenges and standard solving it including rationale behind.

1.4.2 Data spaces design principles

The DS4SSCC endorses the principles developed by OpenDEI Project (Design Principles of Data Spaces[4]) and International Data Spaces (IDS - International Data Spaces in a Nutshell) to be considered prior to embarking on the data space journey. **Data Sovereignty** - Both IDS and OpenDEI consider “Data Sovereignty” as a fundamental aspect. It can be defined as a natural person’s or corporate entity’s capability of being entirely self-determined with regard to its data. This means that a data owner can define usage restrictions to their data, before sharing it with data consumers. Data consumers must accept the usage restrictions.

1. **Data Ecosystem & Data level playing field** - Should enable new business models and nurture innovation. It is rarely possible to build solutions with one source of the data and often solutions are built from the different sources of data, which complement the other data. Therefore, a data ecosystem needs a data space to enable these new innovative services. It by default implies that new entrants face no barriers to access or use the data. Data should be fairly shared and made available equally to all the players and at any cost the

[4] https://design-principles-for-data-spaces.org/

monopolistic situations should be avoided. The Data Level playing field is a pivotal condition to create a fair data sharing ecosystem.

2. **Security & Trust** - A participant has to trust that other participants in the data space get valuable data, which should only be used with regard to the usage policies, defined by the data provider. Security is strongly coupled with the Trust and should have state-of-the-art security, to also guarantee trust and data sovereignty. The key thing to consider is the Authentication & Authorisation component which would enable decentralised Identity Management.
3. **Interoperability** - In an ideal Data Space different ecosystems will exchange the different kinds of data in different formats and protocols. So it is important to have a standardised interoperability with APIs and Data models
4. **Decentralised Soft Infrastructure** - A soft infrastructure for data space interoperability and data sovereignty of users is the way to prevent that the current mode of operation, which is characterised by a limited number of providers and concentration of 'data power' in a few hands, will prevail. The soft infrastructure will lead to decentralisation and a level playing field for data sharing and exchange. The Data Space should be decentralised and federated. This is key to get more data ecosystems joined together.
5. **Governance** - EU Data Governance Act, confirms the notion of a governance structure constituted by multiple entities. For European data spaces, it is recommended to have a (domain) governance authority for each data space and a central governance authority overseeing all aspects in connection with interoperability of data spaces, i.e. the de-facto 'soft infrastructure'. This central authority will interact with all data space specific authorities. The architecture should support and enable the Governance and should give technical challenges to implement governance body decisions. Check D2.2 Multi stakeholders governance schema deliverable for more information.

As stated in the Design Principles of Data Spaces, it is important to consider the Architecture Requirements for Data Spaces listed below:

- Data-sharing empowerment
- Data-sharing trustworthiness
- Data-sharing publication
- Data-sharing economy
- Data-sharing interoperability
- Data space engineering flexibility
- Data space community

1.4.3 Domain specific considerations

When it comes to cities in the context of the Data Spaces, it is important to consider the two scenarios mentioned below.

1. Greenfield cities or communities where there is nothing incorporated from Smart Cities point of view or does not have any digital platform providing any data or services.
2. Brownfield cities which already have a well defined and implemented smart city platform and are very mature in collecting and using the data.

Considering the above mentioned scenarios, the whole idea of this architecture is to enable cities to create Local Data Spaces and join into the National and the EU Level Data Space including domain specific Data Spaces (e.g. Mobility, skills, Green deal, etc).

Data platform owners inside the city might be owned and managed by private partners and might be entirely proprietary. So, it is important to consider the interoperability aspect to allow such platforms to join a Data Space. However, the existing data platform's main objective should not be compromised in order to join the Data Space.

Data of following types should be considered:

a. City administration data
b. Citizen data (including personal and health data, but excluding social media and crowdsourcing etc)
c. GIS, BIM etc
d. Transportation and utilities data
e. Data from private service providers
f. Other research data (economical, environmental, social…)
g. Data on media files like scanned archives and CSV files
h. Contextual data

Many communities work with digital platforms to collect and process data about air quality, traffic management, parking spots, affluence of people and many more scenarios. Data spaces evolve this centralised approach into a decentralised one where data sharing happens in a trusted ecosystem of participants with well defined governance and business rules. Hence, the proposed architecture should provide the required mechanisms to engage and reuse the existing data platforms in cities. Section 3.3 describes how to adapt the existing architectures and platforms to the proposed one for data spaces.

Data spaces will also be subjected to certain European regulations like Data, Data Governance and Interoperability Acts, ePrivacy and Open Data Directive, eIDAS or GDPR, among others. The architecture must be designed bearing in mind all these regulations, and others may come, to be compliant with law and with EU values.

Section 2.2 provides an analysis of this related regulation and how relevant it is for the DS4SSCC architecture design.

Specifically, the European Interoperability Framework has defined a reference architecture to ensure interoperability across Europe. The impact on the DS4SSCC architecture is evaluated in section 2.4.

1.4.4 MIMs role in Data Space architecture

The MIMs (or Minimal Interoperability Mechanisms) were originally designed as a set of tools to help cities, communities and regions who were just starting on the road to utilising data from IoT devices to develop useful applications to support their citizens. The MIMs provided them with a minimal but sufficient way of handling that data that also enabled good-enough interoperability with other applications complying with the MIMs. This meant that MIMs-compliant applications developed in one community could be comparatively easily ported to be used in other communities. It also meant that the data coming from one MIMs-compliant IoT based application within a community could be combined with data coming from another MIMs-compliant IoT application to provide added insights and value.

The MIMs have also proved their worth in data spaces, which are designed to enable data from many different organisations and platforms within a city to be shared and re-used. The challenge here is that the different organisations and platforms are likely to use very different ways of collecting and handling data and therefore perfect interoperability is not practical in the short to medium term. Here the role of the MIMs is to provide a minimal but sufficient set of requirements that can be agreed on by all participants in a data space to enable data to be shared and reused with minimal effort

For instance, many organisations use the package of solutions represented by NGSI-LD to manage context information, while for others the geo-spatial standards developed by the Open Geospatial Consortium provide an adequate mechanism to do this. The Context Management MIM defines a minimal set of requirements that are needed to manage context information. The MIM then shows how both NGSI-LD and the OGC set of standards address those requirements and this enables commonalities between these two approaches to be easily identified.

One example is that NGSI-LD and the OGC approach differ in the way they treat geographical data. NGSI-LD uses GeoJSON notation, whereas in OGC standards, WFS services typically provide location information on Features in GML format. This is an XML format which defines Geometries such as a Point, Polygon, etc. that have coordinates. However, GeoJSON to GML conversion and vice versa is straightforward and can always be done in its entirety - without losing any geographical data. Many tools exist to do this. To achieve minimum interoperability between these two implementations simply requires setting up an automated transformation of geographic data.

There are ten MIMs that are under various stages of development covering issues such as context information management, data models, data security, and the handling of personal data. More will undoubtedly be needed. The existing set of MIMs are covered in detail within the DS4SSCC Catalogue of Specifications[5]. They are also mapped to the DSSC Building Blocks taxonomy. As the DSSC taxonomy has evolved during last months, here below it is included again the mapping in the latest version of the taxonomy.

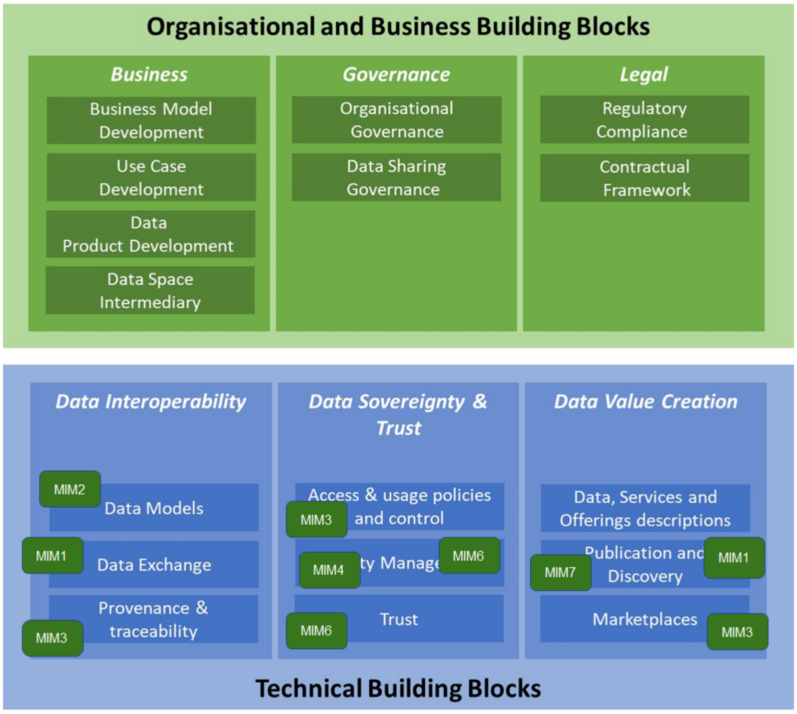


Figure 3. MIMs mapping into DSSC Building Blocks taxonomy (Sep 2023)

The architecture model could identify all areas in data management within a data space where it would be helpful to require participants to comply with a particular MIM to enable easy sharing and re-use of data.

It would be useful if the architecture would not only show where the existing MIMs fit in, but also indicate gaps where new MIMs might be needed and linkages between the MIMs. Potentially it could also help with scoping out some of the MIMs that are still under development.

[5] https://inventory.ds4sscc.eu/

At a more foundational level, the MIMs are built using a simple and logical process that might prove some value within the architecture model itself.

A MIM is based on a clear objective related to data sharing. Several capabilities or business requirements that will enable a good-enough way of fulfilling that objective are then listed, and these are then translated into a set of functional and quality requirements that form the core of the MIM. For all participants to comply with those requirements would form a good basis for interoperability within any data space.

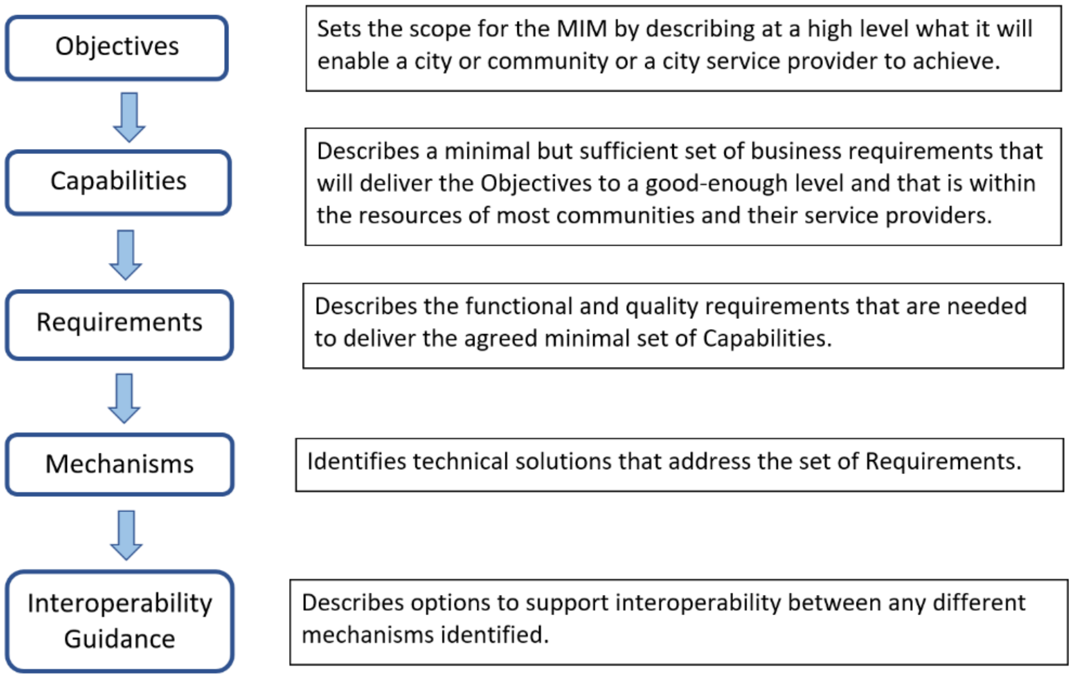
The MIM then identifies technical solutions that are already used by city and community stakeholders to meet the objective of the MIM – the Mechanisms – and describes how they address each of the requirements.

For instance, one of the requirements of the MIM on data management is that . “Unique and persistent identifiers should be used to identify particular instances of any entity used in data sets, and the type of identifier used should be made explicit”. One technical solution that addresses this requirement is to use the W3C Uniform Resource Identifier (URI) standard and another technical solution is to use the ISO defined Digital Object Identifier (DOI).

This step of identifying how different mechanisms and technical solutions address the set of requirements will identify common components used by each of these mechanisms and common interfaces for which open APIs can be developed, to support interoperability between those different mechanisms. In this way, it will be easier to align, share and reuse data coming from stakeholders and stakeholder platforms in a community that may use different mechanisms to handle data.

In the example given above, the use of DOIs and URIs can be used to help integrate between mechanisms using these different technical solutions because DOIs can be turned into URIs by setting up a “resolver” service, which can generate URIs for each DOIs.

Below Figure 4 outlines the structure of a MIM. This structure was first developed in the context of the Living-in.EU Technical Subgroup under the coordination of OASC together with Living-in.EU signatories as a preliminary effort to standardise MIMs. The updates to the MIMs Plus built on the work OASC, along with DG Connect, is doing to standardise the MIMs concept and format through the ITU, one of the three WTO recognised global Standards Development Organisations. This will be published as Recommendation Y.MIM. The latest version of Y. MIM will be reviewed at the ITU Plenary meetings and awaits consent. Hence, from now on a MIM is required to consist of the following elements: Objective (providing the scope of the MIMs), Capabilities, (defining what they will enable the city or community to achieve), Requirements, (translating the capabilities into terms that can be addressed by technical solutions), Mechanisms, (identifying the different technical solutions that cities and communities are already using to meet those requirements), and Interoperability Guidance, (suggesting how to help align the results of the different Mechanisms that cities might use.



Sets the scope for the MIM by describing at a high level what it will enable a city or community or a city service provider to achieve.

Describes a minimal but sufficient set of business requirements that will deliver the Objectives to a good-enough level and that is within the resources of most communities and their service providers.

Describes the functional and quality requirements that are needed to deliver the agreed minimal set of Capabilities.

Identifies technical solutions that address the set of Requirements.

Describes options to support interoperability between any different mechanisms identified.

Figure 4. MIMs description template

Incorporating work on the MIMs within the architecture model would therefore help to:

1. Identify a basic set of foundational capabilities needed to address each issue related to data sharing in a data space identified by the model, along with the functional or quality requirements needed.
2. Point to any widely used alternative mechanisms to address those functional or quality requirements
3. Identify and define common standardised components within those mechanisms
4. Identify interfaces to see if the use of APIs might be viable

2 State of play

This section includes the enumeration and brief description of the existing legal, technical, and business aspects which can be taken into account at the designing of the DS4SSCC architecture.

2.1 Existing reference architectures

Several reference architectures provided by diverse initiatives were analysed to identify possible full or partial reusability of them. Some of them are well-proven reference architectures for smart cities solutions while others are more incipient and specific for data spaces.

2.1.1 DSBA Technical convergence

Big Data Value Association[6] (BDVA), FIWARE Foundation[7], Gaia-X[8] and the International Data Spaces Association[9] (IDSA) decided to join forces and formed the Data Spaces Business Alliance[10] (DSBA) aiming at driving the adoption of data spaces across Europe and beyond. As part of this plan, members of the DSBA agreed to work towards defining a common reference technology framework, based on the technical convergence of existing architectures and models, leveraging each other's efforts on specifications and implementations. The goal was to achieve interoperability and portability of solutions across data spaces, by harmonising technology components and other elements. As a result of this effort, the DSBA Technical Convergence document was released, which continues to evolve. At the time of this deliverable, the 2nd iteration of the document was released and a 3rd iteration expected to be released during Autumn 2023.

A Minimum Viable Framework (MVF) created in the convergence document contains 3 main technology pillars of the Data Spaces.

- Data Interoperability
- Data Sovereignty & Trust
- Data Value Creation

The below picture illustrates the main actors involved in a data space and the systems they have to instantiate and operate.

[6] https://www.bdva.eu/
[7] https://www.fiware.org/
[8] https://gaia-x.eu/
[9] https://internationaldataspaces.org/
[10] https://data-spaces-business-alliance.eu/

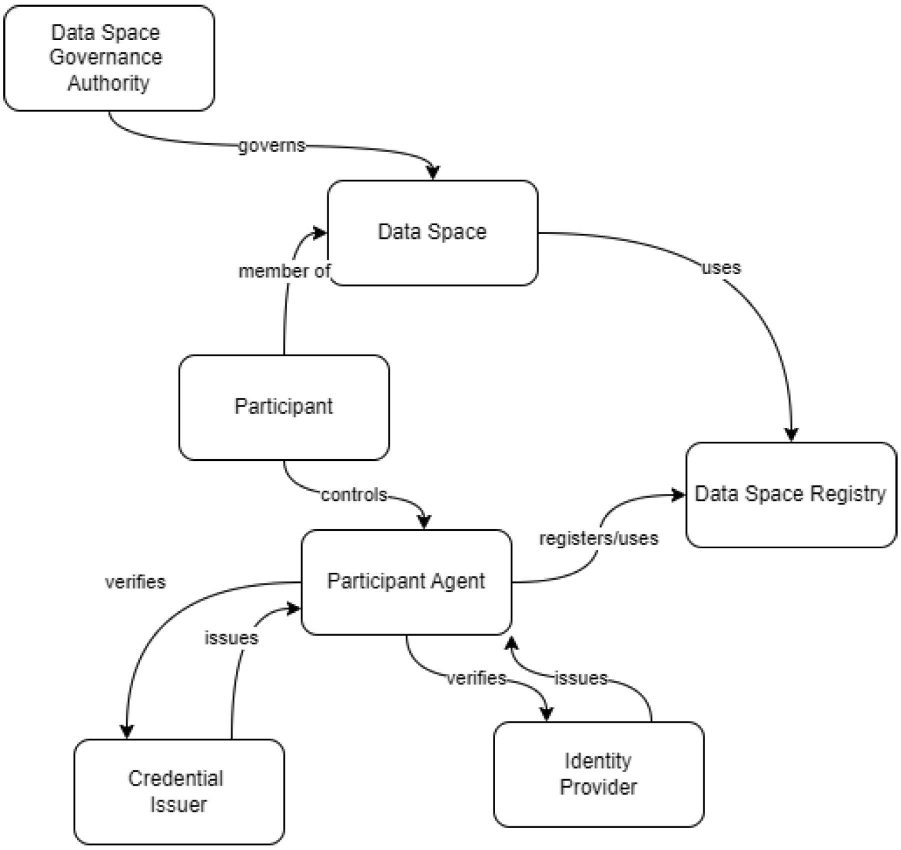


Figure 5. Data space roles according to DSBA

The convergence document describes the framework with which the existing data platforms can participate in a Data Space and satisfy all the requirements of the Technology building blocks. It also guarantees technical interoperability by following the IDS protocol standard. The framework enables usage of Gaia-X Trust Framework. The Framework is adapted in i4Trust framework and this framework is the only one currently available which has an implementation. Framework in action can be understood in the I4Trust architecture description section.

The convergence framework is very relevant to DS4SSCCC and some illustrations and components can be used in the DS4SSCC common architecture as it provides the framework and addresses all the blocks of Data Spaces adapted by DSSC.

- ***Relevance for DS4SSCC Architecture Model***

Relevant concepts, standards and technologies from this reference architecture that could be used in the Architecture Model are the following:

- open DEI building blocks for data spaces,
- multi marketplace concept for data spaces,

- mapping between system components (data space connectors, federated services, data space registry and Open DEI building blocks);

2.1.2 FIWARE Reference architecture for Smart Cities

FIWARE Smart City Reference Architecture (FIWARE-SCRA) represents one of the cornerstones for future smart city reference architectures. It introduces information context management (using CEF Context Broker Building Block), enables data exchange interoperability using interoperable data models (using Smart Darta Models), supports seamless integration with the IoT devices (using IoT management component), represents system-of-system digital twin concept (with connecting smart solutions) and shows possible integration of the business layer via data marketplace (using the BAE marketplace framework).

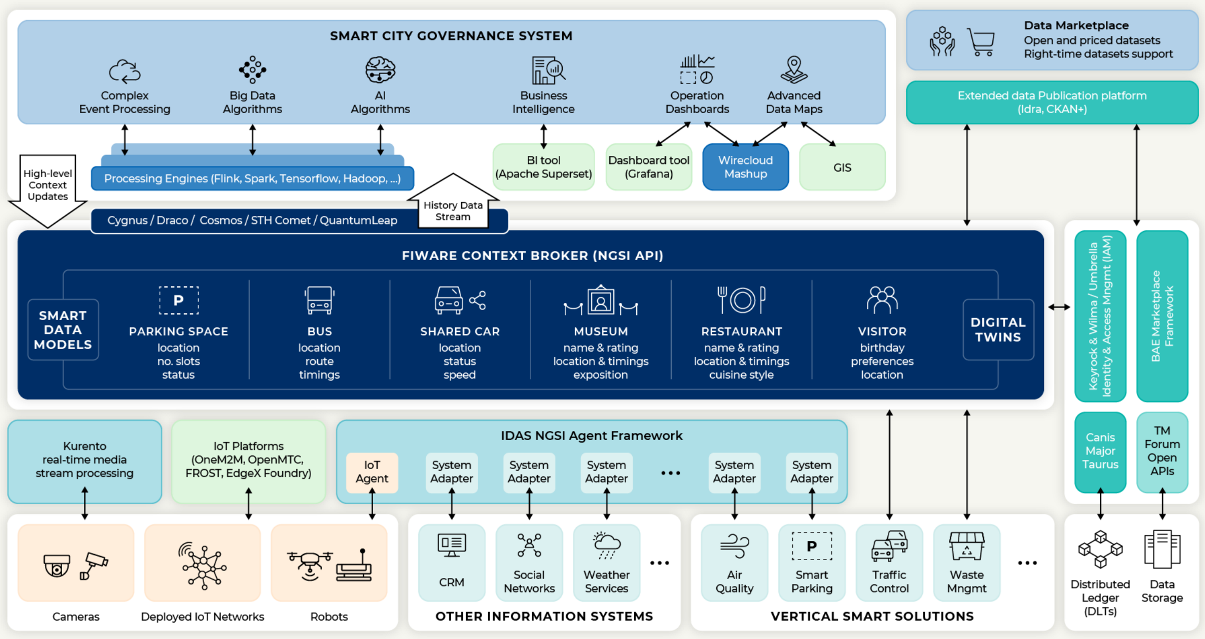


Figure 6. FIWARE Smart City Reference Architecture

- ***Interoperability***
  - Data Models

The FIWARE-SCRA uses data model interoperability in different areas such as:

- Interoperable Data Exchange: the FIWARE-SCRA incorporates the use of Smart Data Models (OASC MIM L2 standard) to achieve data interoperability and standardisation across different domains and applications.

Programming API

The FIWARE-SCRA uses programming API interoperability in different areas such as:

- Context Data Management: the FIWARE-SCRA adopts the NGSI-LD (Next Generation Service Interface Linked Data) data model (OASC MIM L1 standard) as a core component for context data management.
- Data Publication and Discovery: the FIWARE Context Broker component, based on NGSI-LD, enables publishing and discovery of context information,
- Data Marketplace: The FIWARE-SCRA can integrate with the TM Forum Open APIs to enhance its capabilities and support interoperability within the smart city ecosystem in different areas: catalogue management, ordering and billing, resource management, party management...

Identity management and access control

The FIWARE-SCRA provides identity management capabilities through its Identity Management component, which is responsible for managing user identities, authentication, and authorization within the FIWARE-SCRA. Different components such as Keyrock or Keycloak can be used for identity management which support OIDC, SAML 2.0 or OAUTH2 standards for authorization.

The Platform also incorporates advanced access control mechanisms to ensure secure and granular authorization within the platform. These mechanisms include Policy Administration Point (PAP), Policy Decision Point (PDP), Policy Enforcement Point (PEP), and Policy Management Point (PMP).

Marketplace

The FIWARE-SCRA uses FIWARE Business Application Ecosystem (BAE) for business integration support. The BAE serves as a platform for enabling the exchange, monetization, and collaboration of data and services. BAE components implement different TM Form Open APIs for this purpose.

Relevance for DS4SSCC Architecture Model

Relevant concepts, standards and technologies from this reference architecture that could be used in the Architecture Model are the following:

- introduction of system-of-systems architecture model,
- CEF Context Broker building block for interoperable data exchange,
- OASC MIMs for data and API interoperability,
- OIDC, SAML 2.0 or OAUTH2 compatible identity management,
- PEP (Policy Enforcement Point) for enforcing access control policies, and a PDP (Policy Decision Point) for making access control decisions,

- data marketplace support for business layer integration using FIWARE Business Application Ecosystem (BAE),
- open governance model, defining the lifecycle of data models comprising incubation of new data models and curation of the existing data models;

2.1.3 IDS-RAM

The International Data Spaces Reference Architecture Model (IDS-RAM) is a virtual data space leveraging existing standards and technologies, as well as governance models well-accepted in the data economy, to facilitate secure and standardised data exchange and data linkage in a trusted business ecosystem. The IDS-RAM aims at meeting the following strategic requirements: trust, security and data sovereignty, ecosystem of data, standardised interoperability, value adding apps and data markets. IDS-RAM enables secure and sovereign data sharing while maintaining data ownership and privacy.

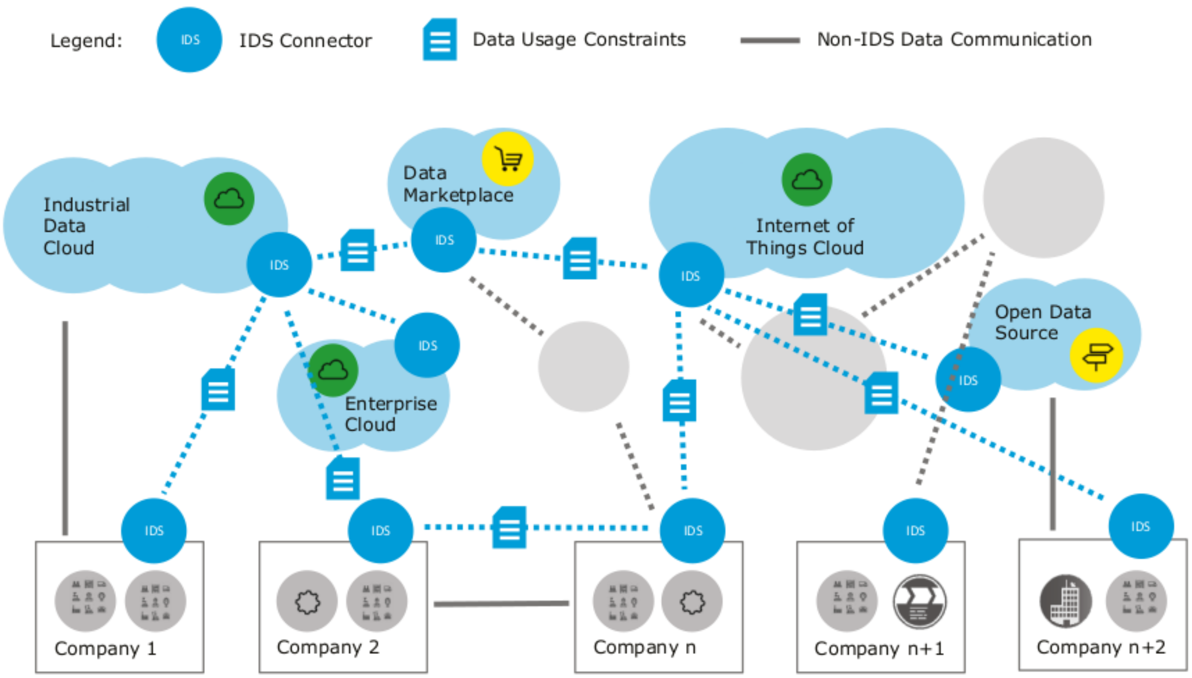


Figure 7. IDSA Reference Architecture Model

- ***Interoperability***
    - Data Models

Interoperability of data models is not addressed in IDS-RAM.

- Programming API

Interoperability of programming APIs is not addressed in IDS-RAM.

- *Identity management and access control*

Each Connector must have a valid X.509 certificate. The Certificate Authority (CA) issues certificates for all entities. With the help of this certificate, each participant in the IDS that operates an endpoint is able to verify the identity of any other participant. PKI (Public Key Infrastructure) can have several layers to achieve separation of duties (i.e., every Sub-CA is responsible for a specific topic).

IDS-RAM focuses on technical enforcement to address data usage control restrictions. To enforce data usage restrictions IDS-RAM uses Policy Enforcement Point (PEP) and Policy Decision Point (PDP). The interaction between the PEP and PDP follows a request-response model. The PEP sends access requests to the PDP, which evaluates the requests against the policies and provides a decision. The PEP then enforces the decision by either granting or denying access to the data based on the response from the PDP.

- *Marketplace*

The IDS-RAM enables the creation of novel, data-driven services that make use of data apps. It also fosters new business models for these services by providing clearing mechanisms and billing capabilities, and by creating domain-specific broker solutions and marketplaces.

- *Relevance for DS4SSCC Architecture Model*

Relevant concepts, standards and technologies from IDS-RAM that could be used in the Architecture Model are the following:

- IDS connector driven decentralised architecture,
- X.509 digital certificate based authentication & authorization for IDS connectors,
- modular design with standalone Connector, App Store, and Broker components,
- advanced certification and governance perspectives of IDS reference architecture model;

2.1.4 i4Trust Reference Architecture

The main goal of i4Trust is to boost the development of innovative services around new data value chains. i4Trust helps to achieve this by providing the right tools, education, coaching and initial funding for the creation of Data Spaces enabling trustworthy and effective data sharing. Ecosystems of collaborating SMEs and supporting DIHs will emerge in a sustainable way around such Data Spaces. i4Trust integrates standard-based building blocks from the FIWARE and iSHARE frameworks. Together with common data models, the FIWARE Context Broker building block supports effective data exchange among parties by using the standard NGSI-LD API.

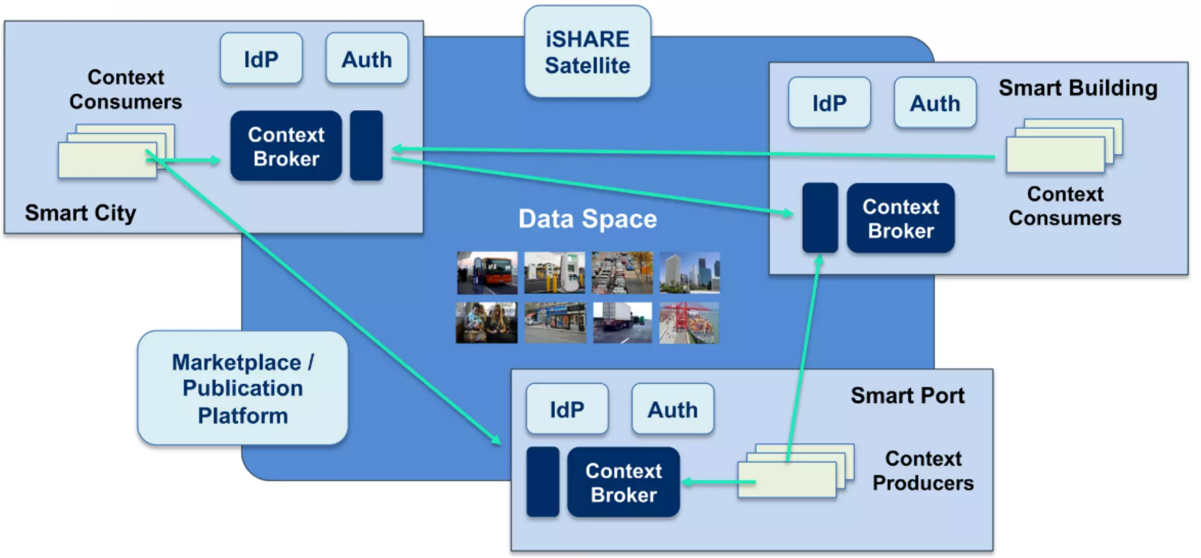


Figure 8. Data Exchange in an i4Trust data space

- ***Interoperability***
  - Data Models

Data model interoperability requires the adoption of common data models that are compatible with the published API. iTrust leverages Smart Data Models initiative to achieve this task.

It provides a library of Data Models described in JSON/JSON-LD format which are compatible respectively with the NGSIv2/NGSI-LD APIs as well as any other RESTful interfaces compliant with the Open API specification. It is also compatible with schema.org and complies with other existing de-facto sectoral standards when they exist.

Smart Data Models initiative supports an open governance model, defining the lifecycle of data models comprising incubation of brand new data models as well as curation of data models via harmonisation of different contributions.

  - Programming API

Participants in i4Trust data spaces exchange digital twin data using the NGSI-LD API. Systems which have not been architected using FIWARE can still use the NGSI-LD API to share data they produce and consume using NGSI-LD system adapters.

- ***Identity management and access control***

Identity Management (IM) building block is implemented in i4Trust through multiple distributed Identity Providers (IdP) allowing identification and authentication of organisations, individuals, machines and other actors participating in a data space.

Human identities are based on the adapted in the iSHARE specification OpenID Connect standard while organisational identities are based on Public Key Infrastructure (PKI) and OAuth2.0 standards. iSHARE uses eIDAS based digital certificates for digitally signing data and assertions.

iSHARE compatible FIWARE Keyrock component, supports OpenIdConnect, SAML 2.0 and OAuth2 standards, and is used for identity management. An alternative to IM and IdP based solutions which can coexist in iSHARE, is the usage of a Self-Issued OpenID Provider (SIOP).

Additionally users can use OpenID Connect for Verifiable Presentations (OIDC4VP) flow to authenticate and authorise in compliance with the European Digital Identity Wallet Architecture and Reference Framework.

- ***Marketplace***

FIWARE Business Application Ecosystem (BAE) 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.

- ***Relevance for DS4SSCC Architecture Model***

Relevant concepts, standards and technologies from i4Trust reference architecture model that could be used in the Architecture Model are the following:

- Open DEI building blocks for data spaces,
- OASC MIMs for data and API interoperability,
- mapping between Open DEI building blocks and FIWARE/iShare technical components,
- CEF Context Broker building block for interoperable data exchange,
- standard for secure and controlled exchange of data using iSHARE decentralised IAM,
- eIDAS based digital certificates for digitally signing data and assertions,
- data marketplace support for business layer integration using FIWARE Business Application Ecosystem (BAE),
- European Digital Identity Wallet Architecture and Reference Framework,
- recording of NGSI-LD transaction logs into different Distributed Ledgers / Blockchains,

- avoided vendor lock-in scenarios since all the tools are standard-based and supported by open source reference implementations,
- open governance model, defining the lifecycle of data models comprising incubation of new data models and curation of the existing data models;

2.1.5 ODALA Architecture

Launched in September 2020, the “Collaborative, Secure, and Replicable Open Source Data Lakes for Smart Cities” (ODALA) is a strategic project to improve data management in cities and regions. European cities and regions from four different countries together with a cluster of private companies and research institutes will leverage open source technologies and digital transformation – for the benefit of public administrations. ODALA will adopt the European Union Digital Service Infrastructure (DSI), also known as Connecting Europe Facility (CEF) Building Blocks. The building blocks support the creation of a digital single market where cities and companies can connect and share data. This environment is called ‘data lake’ and will allow cities to connect different data sources – static, historical, and real-time data – from diverse departments within the cities.

ODALA’s architecture of smart city and integration of the services is derived from FIWARE Smart City reference architecture. Below is the reference architecture (iteration 1) from ODALA publicly available pages

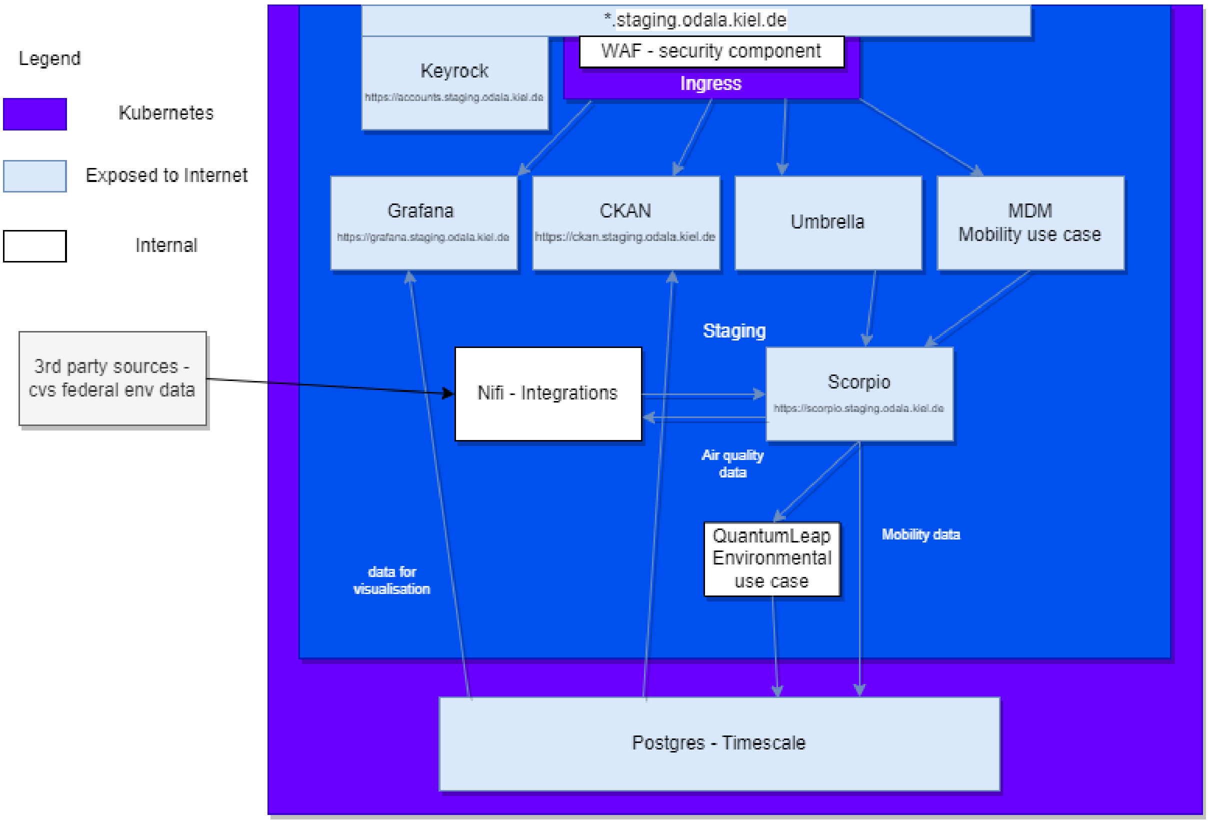


Figure 9. ODALA reference architecture (iteration 1)

Architecture highlights are:

- A south bound options for data collection
- NGSI-LD Broker for APIs and messaging
- PEP (Policy Enforcement Point) for security
- Keyrock for Authentication and Authorisation
- CKAN for Dataset publishing and Grafana for charts and visualisations

In the iteration 2 ODALA has evolved the Smart City architecture which is mentioned above to Data Space architecture to adapt i4Trust Framework in addition to Federation. Below is the reference architecture (iteration 2) from ODALA publicly available pages.

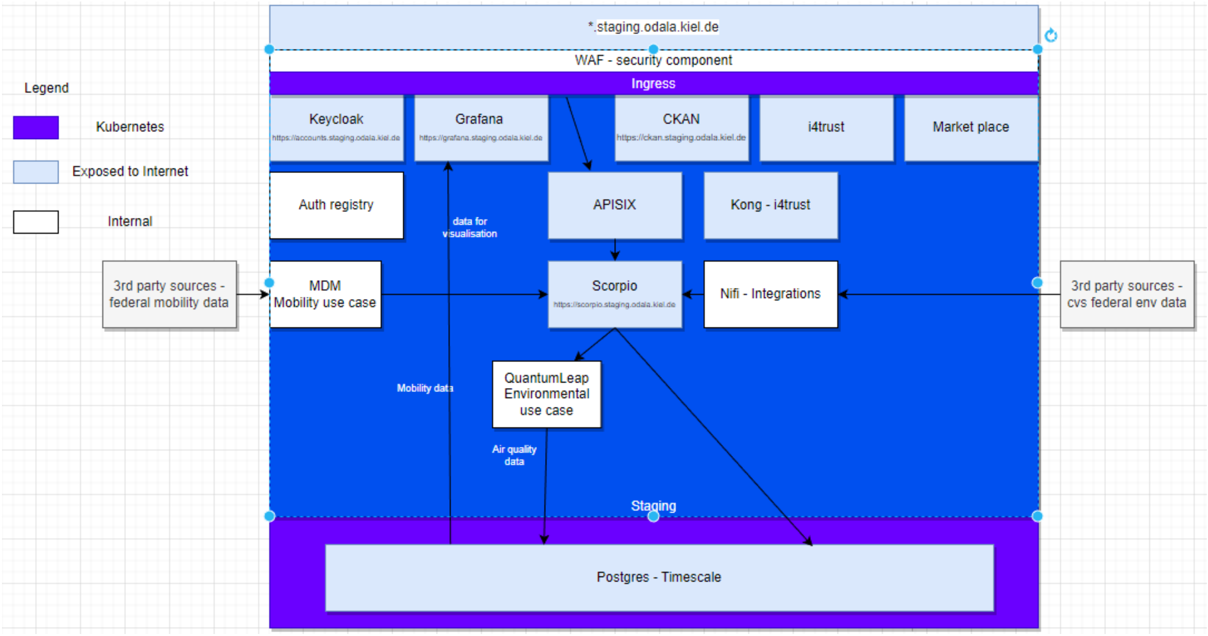


Figure 10. ODALA reference architecture (iteration 2)

- ***Relevance for DS4SSCC Architecture Model***

ODALA is a perfect example for DS4SSCC as it started as a Data Lake for a city and then evolved to be a Data Space. The project has adapted the evolution of the Data Spaces and created a PoC that Data platforms in the city can easily evolve to the Data Space by following the standardised architecture and open standards. Above all ODALA also uses the Federation architecture, which might be key in some implementations.

More information at https://gitlab.publiccode.solutions/odala-public/trusted-broker-federation/-/tree/main/Documentation.

2.1.6 Gaia-X Reference Architecture

Gaia-X provides the components to address compliance, federation and interoperable data-exchange: The specifications and the supporting Open Source Code are defined in the Gaia-X Framework. Gaia-X Framework is following the same principles as DSBA Technology convergence.

Gaia-X defines the framework in 3 planes:

1. Usage Plane
2. Management Plane
3. Trust Plane

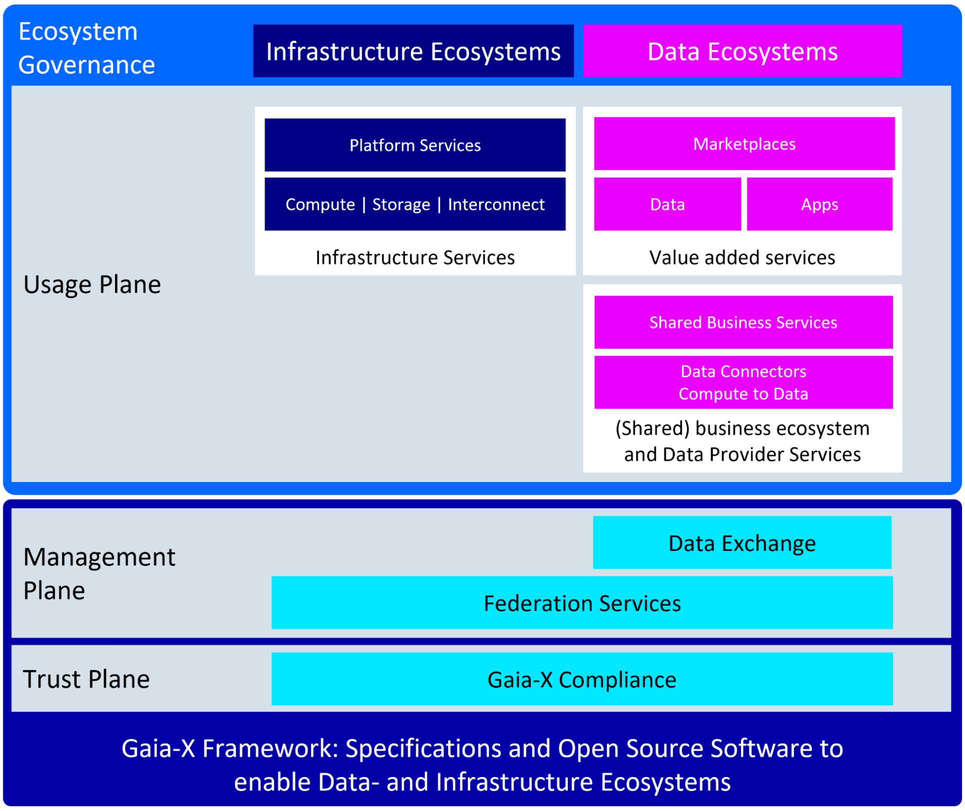


Figure 11. GAIA-X framework

Gaia-X is more about the ecosystem of the Data Spaces but does not define the Data Space architecture. Still, Data Spaces are going to be part of the Ecosystem which is following the Gaia-X Framework.

The data Space architecture cannot be designed and determined in isolation. The Gaia-X Ecosystem is the virtual set of Participants, Service Offerings, Resources fulfilling the requirements of the Gaia-X Trust Framework. Gaia-X enables Interoperability between independent autonomous ecosystems.

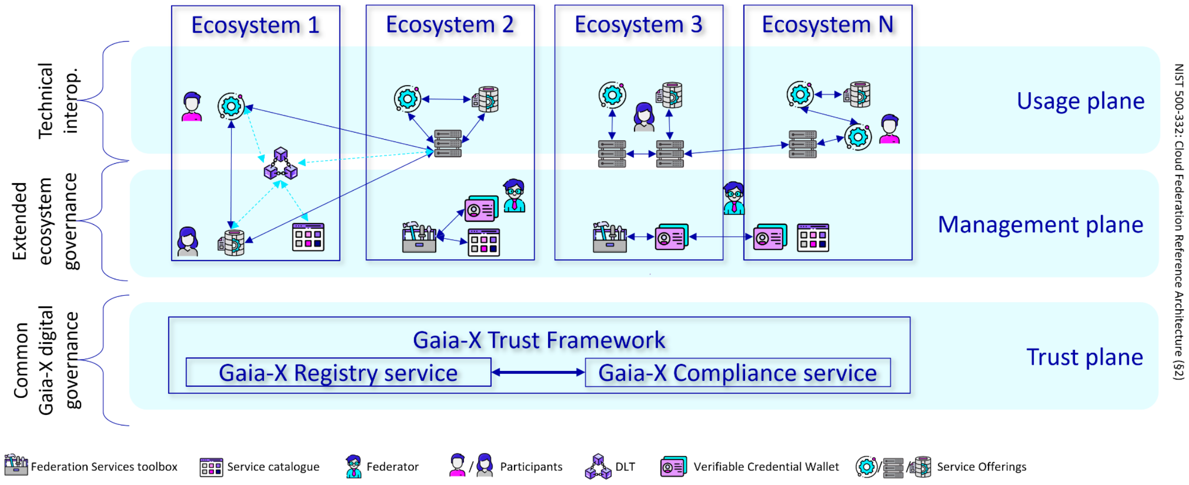


Figure 12. GAIA-X trust framework

It is important that the various building blocks of the Data Space architecture complies with Gaia-X architecture[11] and follow the interoperability framework defined in Gaia-X.

- ***Relevance for DS4SSCC Architecture Model***

Gaia-X defines the ecosystem FW and interoperability for the Data Spaces that they could potentially be part of, so it is important to design the Data Space architecture considering the bigger picture of Ecosystem which they could be part of.

2.2 Related EU regulation

Some EU regulation may affect the DS4SSCC architecture by introducing some requirements to be considered in the design or configuration of the technical solutions. D2.2 provides an extensive analysis (Data Act, Data Governance Act,...), so this section is not going to replicate here the same information. However the section illustrates which are those requirements that are of relevance for the technical design.

2.2.1 Identity and Access

The Single Digital Gateway Regulation and the e-IDAS2 Proposal form the primary pillars of identity and access management in Europe's digital transformation.

[11] https://docs.gaia-x.eu/technical-committee/architecture-document/22.10/

Relevance:

These regulations determine the implementation of the Trust building block of the dataspace. One provides EU citizens and businesses with borderless access to services and information via a centralised portal. The other builds on e-IDAS to advance the adoption of electronic identification, authentication and trust services across Europe by introducing digital identity wallets and a European Digital Identity.

2.2.2 Privacy and Security

The European Union addresses the question of privacy with the following legislations: General Data Protection Regulation (2016), ePrivacy Regulation Proposal (2017), Cybersecurity Act (2019), Network and Information Security 2 Directive (2022).

Relevance:

These regulations articulate the need for a harmonised approach to cybersecurity, whilst each Member state should define their respective national cybersecurity strategy. As such, when adopting a blueprint to a use case developers should align with the regional approach and consult with the national computer security incident response teams. In doing so, they are required to identify vulnerabilities and risks together with mitigation strategies.

2.2.3 Platforms

The Digital Services Act (2022) together with the Digital Markets Act (2023) were launched to create a competitive European digital market where competition and rights are regulated.

Relevance:

These legislations are in place to regulate gatekeepers and create a fair market place for all. In case of a public-private partnership, a data space is in the position to be exploited by businesses. As such a data space should be technically equipped to identify and/or disable malicious behaviour as defined in Article 5, 6, 7, 8 of the Digital Market Act with the exemption of public security (Article 10, Digital Market Act).

2.2.4 Interoperability

The European Interoperability Framework, the European Interoperability Framework for Smart Cities and Communities (EIF4SCC), and the European Interoperability Act proposal foster collaboration across borders, industries, and public sectors and mitigate fragmentation in Europe.

Relevance:

In line with these regulations, the blueprint of a data space for smart communities is required to technically align with common principles, models and recommendations to support the interaction between a) administration to administration b) administration to business c) administration to citizens by means of the exchange of data between their ICT systems. In the context of smart and sustainable cities and communities, the dataspace holds the responsibility of being a key enabler of interoperability. Specifically, it should focus on neutrality, data portability by including open standards, open technical specifications, and multiple access and assistance channels. Following the openness principle a technology should be developed with relevant stakeholders, the results should be available for everyone. In the case of a dataspace blueprint, it should list specifications with intellectual property rights licensed on Fair, Reasonable and Non-Discriminatory (FRAND) terms, in a way that allows implementation in both proprietary and open source software, and preferably on a royalty-free basis. Consequently, the development of the blueprint follows an open process, and adopts the Minimal Interoperability Mechanisms (MIMs Plus) that are developed in the environment of the Living-in.EU movement.

2.2.5 Data

The Data Governance Act (2023) and the Data Act Proposal (2022) together with the Open Data Directive (2019) serve to ensure ethical and safe data use and sharing across sectors in Europe. In the context of data spaces these regulations have the greatest impact.

Relevance:

The European Data Innovation Board (EDIB) - proposed by the Data Governance Act - will provide guidelines regarding the following: (i) cross-sectoral standards to be used and developed for data use and cross-sector data sharing; (ii) requirements to counter barriers to market entry and to avoid lock-in effects, for the purpose of ensuring fair competition and interoperability; (iii) adequate protection for legal data transfers outside the Union; (iv) adequate and non-discriminatory representation of relevant stakeholders in the governance of a common European data space; (v) adherence to cybersecurity requirements in line with Union law. Additionally, the data space for smart communities should consult the high-quality public sector dataset that will be made available due to the Act on High-Value Datasets (2023).

2.3 Data Spaces Support Center blueprint

The Data Spaces Support Center (DSSC) is in charge of coordinating and supporting the Common European Data Spaces in the design and deployment of their data spaces in different sectors. In relation to the DS4SSCC technical blueprint, the DSSC is providing a common blueprint for all data spaces as a guidance to generate specific blueprints across all data spaces. Thus, the

DS4SSCC blueprint leverages on the DSSC blueprint in terms of glossary, conceptual model and templates.

According to the DSSC, a data space blueprint is *a consistent, coherent and comprehensive set of guidelines to support the implementation, deployment and maintenance of data spaces*. DSSC blueprint includes:

- a glossary of terms in the scope of data spaces
- a conceptual model which defines the terms and relationships in a data space
- a set of Building Blocks descriptions and specifications according to the DSSC taxonomy of BBs
- a collection of candidate standards and technologies to be recommended by the DSSC (cross-domain standards) and the data space (domain-specific standards) in the implementation of the BBs
- an overall integration document explaining and providing the guidance to use all the elements in the deployment of a data space

Since the first full version of the DSSC Blueprint will be officially published at the same time as this document, the DSSC has been providing partial releases of the different elements for the sake of the data spaces.

- ***Relevance for DS4SSCC Architecture Model***

The DSSC Blueprint is crucial for the DS4SSCC architecture, as it states the basis for the proposed Catalogue of Specifications and Reference Architecture. DS4SSCC followed the same structure of BBs, the Conceptual Model as it is mostly relying on the DSBA Conceptual Model, and some of the standards proposed in the candidate collection of standards.

2.4 European Interoperability Framework (EIF)

The New European Interoperability Framework (EIF) is a set of guidelines for developing public services. Figure 13 depicts the interoperability levels of the EIF. They cover legal, organisational, semantic and technical interoperability. Each level deserves special attention when a new European public service is established.

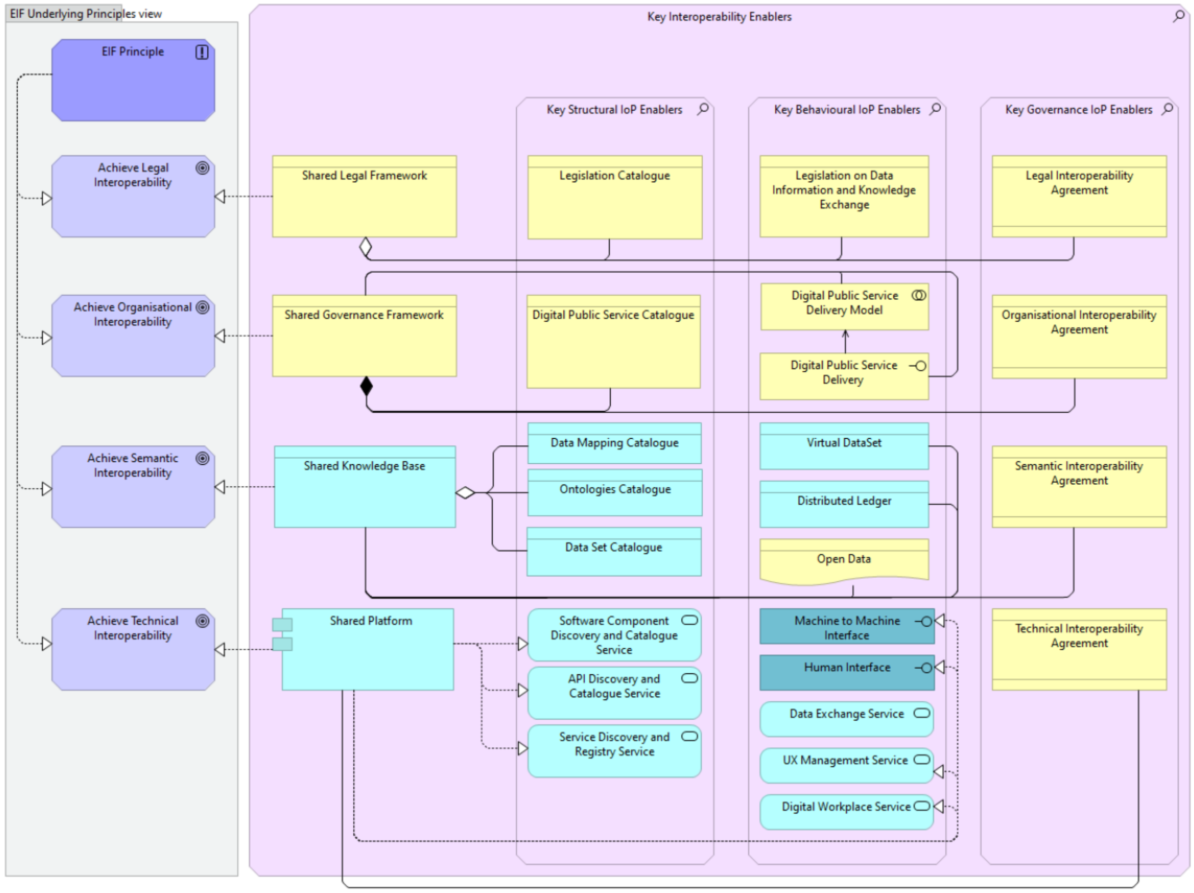

EIF Underlying Principles view
Key Interoperability Enablers
EIF Principle
Key Structural IoP Enablers
Key Behavioural IoP Enablers
Key Governance IoP Enablers
Achieve Legal Interoperability
Shared Legal Framework
Legislation Catalogue
Legislation on Data Information and Knowledge Exchange
Legal Interoperability Agreement
Achieve Organisational Interoperability
Shared Governance Framework
Digital Public Service Catalogue
Digital Public Service Delivery Model
Digital Public Service Delivery
Organisational Interoperability Agreement
Achieve Semantic Interoperability
Shared Knowledge Base
Data Mapping Catalogue
Ontologies Catalogue
Data Set Catalogue
Virtual DataSet
Distributed Ledger
Open Data
Semantic Interoperability Agreement
Achieve Technical Interoperability
Shared Platform
Software Component Discovery and Catalogue Service
API Discovery and Catalogue Service
Service Discovery and Registry Service
Machine to Machine Interface
Human Interface
Data Exchange Service
UX Management Service
Digital Workplace Service
Technical Interoperability Agreement

Figure 13. EIF framework

- ***Relevance for DS4SSCC Architecture Model***

EIF requires following aspects:

- **A common terminology to design, assess, and communicate about eGovernment solutions:** Public administrations can benefit largely from a common terminology to communicate efficiently and unambiguously – across language barriers and domain-specific jargon – when designing, assessing, documenting and discovering Solution Building Blocks (frameworks, tools, services) used to deliver interoperable digital public services;
- **Stable and standardised interfaces for digital public services**: IT architects and developers are tasked with defining stable interfaces between digital public services, according to open standards and interoperability specifications, so that partners can rely on them to build new, aggregated digital public services and avoid vendor lock-in;
- **An overview of already existing Solution Building Blocks (SBBs):** Decision makers, public procurers, ICT experts and architects in public administrations gain value from being able to find already existing (reusable) Solution Building

Blocks that have been developed in-house or by others, to unlock the potential of shared development effort and to be able to find best-in-class reusable components and services.

A **solution architecture template** can include additional interoperability specifications. It is usually applied within a community. Acting as a template for solutions (and their specific architectures), it guides the development of a certain kind of solutions (and their specific architectures). A solution architecture template can exist on different levels of details. For example, it can be used to describe a template for national portals offering e-services to its citizens. It can also be used to describe a template on how to securely exchange files among public administrations.

A solution architecture template consists of the following:

- A goal and a description of the particular supported business capabilities and the involved business information exchanges;
- A subset of EIRA© core Architecture Building Blocks covering all EIRA© views;
- A set of specific Architecture Building Blocks extending EIRA©'s views enabling specific functionalities to be provided by implementations derived from the SAT;
- A set of interoperability specifications for Architecture Building Blocks in the SAT;
- A narrative for each EIRA© view.

3 Architecture model for DS4SSCC

The architecture for the Data Space should follow the below mentioned core principles:

- In line with the technology building blocks defined in the Catalogue of specifications.
- In alignment with the Data Spaces Support Center blueprint.
- Should support the evolution from Data Platforms to Data Spaces

3.1 Existing scenarios

Creating a Data Space might happen with three types of scenarios in conformance with the three levels of digital maturity assessed in the Lordimas: Level 1 - No digital strategy; Level 2- A digital strategy but no technological roadmap and Level 3 - maturity level towards a digital twin (evolving towards the Citiverse).

DS4SSCC architecture is focusing on the most advanced scenarios (brownfield and digital twin) as the existence of a good level of digitalization in the city/community is required, otherwise aspiring to a data space is not realistic. However this document provides some recipes for the greenfield scenario as well, indicating the previous steps are required to get the minimal level of data infrastructures that are needed in a city to become or engage a data space.

- Scenario 1 - Data Spaces can be built from scratch without using any existing platform. But the architecture defined in this document may not cover the data collection for example collecting the data from IoT devices etc and the architecture will assume that there is data (in one way or another). The reference architecture for creating a Smart City can be found in the section in the Link. Further this scenario is called **greenfield.**
- Scenario 2 - Using an existing Data Platform/Data Lake/Smart City Platform - Most of the cities or communities have a Smart City platform or Data platform or Data Lake, if not common it is very much a platform for utilities or traffic etc. The existing platform can be extended by the given architecture in this document to a Data Space. They used to have a strategic plan for the digitalization of the city where the data platform is included. Further this scenario is called **brownfield.**
- Scenario 3 - A Smart Solution which is a **Digital Twin** representation. It is an advanced scenario of brownfield case, where the cities and communities may have some simulation and predictions features beyond the usual data platforms function A Digital Twin is defined as follows -

> *Digital Twin data representation is built based on information gathered from many different sources, including sensors, cameras, information systems, social networks, end users through mobile devices, etc. It is constantly maintained and accessible in near real-time ("right-time" is the term also often used, reflecting that the interval*

between the instants of time at which some data is gathered and made accessible is short enough to allow a proper reaction). Applications constantly process and analyse this data (not only current values but also history generated over time) in order to automate certain tasks or bring support to smart decisions by end users. The collection of all Digital Twins modelling the real world that is managed is also referred to as ***Context*** *and the data associated with attributes of Digital Twins is also referred to as* ***context information****.*

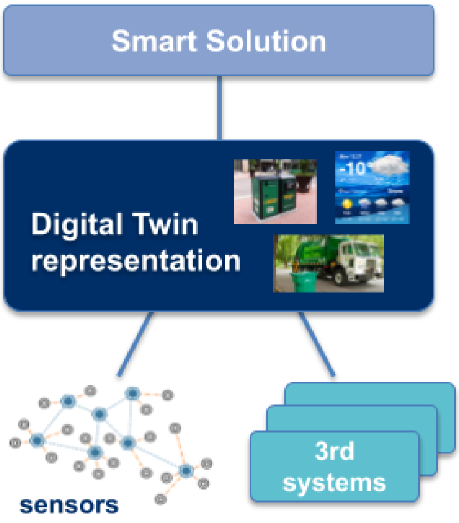


Figure 14. Digital Twin representation

There could be many Digital Twins in a single Smart Solution and can be called as System of Systems.

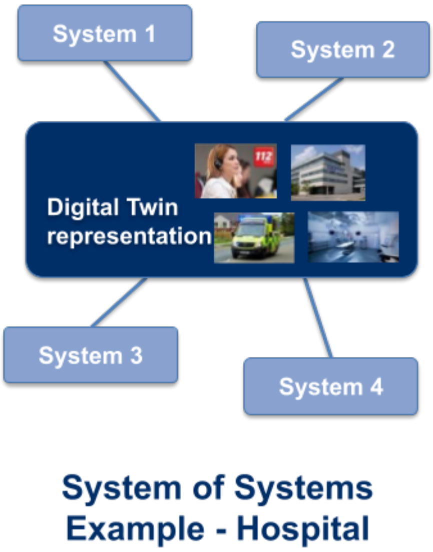


Figure 15. System of systems paradigm

Any platform either Smart City or Data Lake or Data Platform or a Smart Solution having a Digital Twin representation, has common components to handle Security, Data Access, APIs/Data Publication and Data Broker. Typically as shown below -

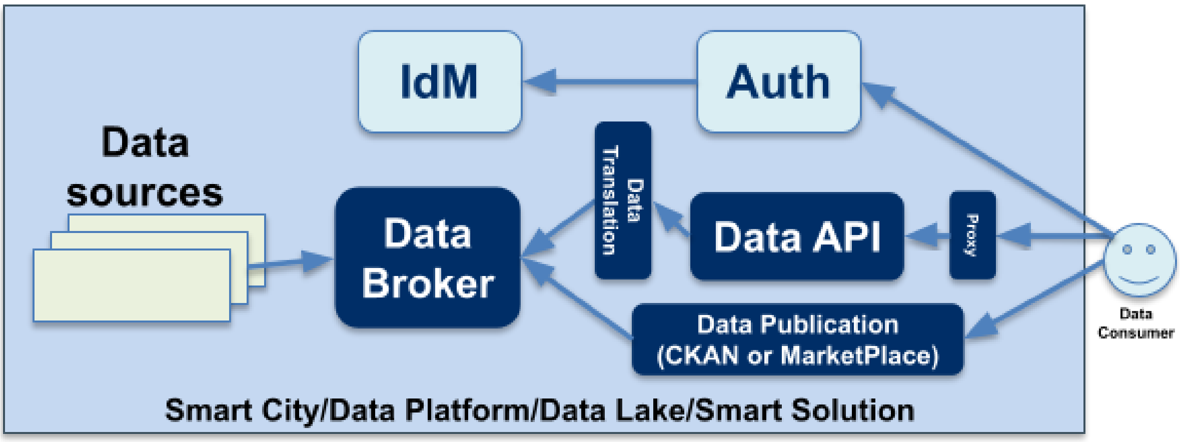

IdM
Auth
Data sources
Data Broker
Data Translation
Data API
Proxy
Data Consumer
Data Publication (CKAN or MarketPlace)
Smart City/Data Platform/Data Lake/Smart Solution

Figure 16. Smart City solution - main building blocks

- IdM - Identity Management for storing the identities of users for example email ids, user ids, passwords etc. IdM can also be used as the Data Access Policy Definition Point (PDP).
- Auth - Authentication for users to log in or register. For example Single Sign on (In some cases IdM and Authentication activities are handled by a single component)
- Data API is the API through which the Data consumers can access the data. Typically the APIs are exposed by the Data Broker and in Smart Solutions by Context Broker. Some platforms will also have a proxy in front of the Data Broker to protect and throttle the access to the Data API. The Proxy component can also behave as Policy Enforcement Point (PEP), which would be allowing the access to the Data API based on the access policies defined at PDP.
- Data Translation (optional) - This component is an optional component to be used when the data platform's data or APIs are not in a standardised format. This may have been added as a result of interoperability agreements between participants in the Data Space.

3.2 Matching with Data Spaces Building Blocks

The existing platforms, either in Scenario 1 or 2 mentioned in the Section 3.1, already have some blocks from Data Spaces Building Blocks.

- IdM (Identity Management) is same as Identity Management and Access & Usage Control under Data Sovereignty & Trust
- Data API is same as Data Exchange API
- Data publication (CKAN/Marketplace) is similar to Metadata & Discovery Services and Publication & Marketplace Services under Data Value Creation.

The literal matching can be seen in the picture below.

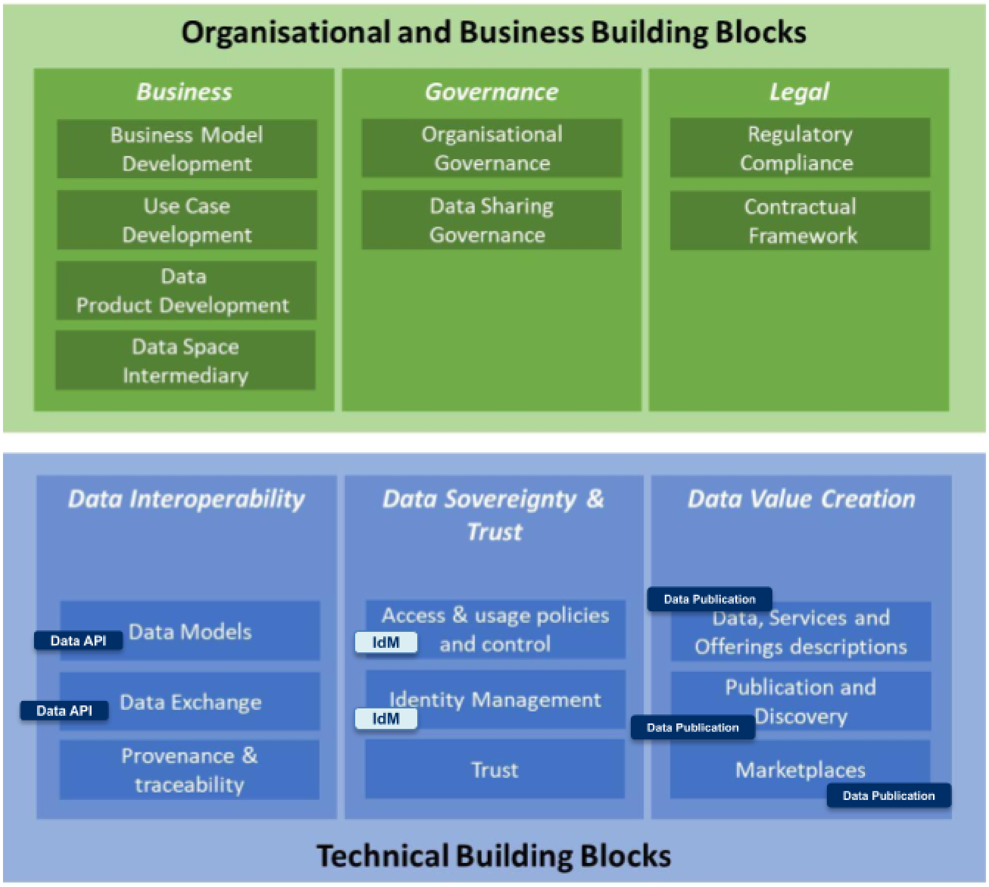


Figure 17. Mapping of typical city components with DSSC building blocks

3.3 High-level architecture

The architecture proposed in this document tries to make the evolution of Smart Solutions/Data Platform to Data Spaces easy, modular and incremental. The emphasis is on "Evolution" of the existing platforms rather than creation of the Data Spaces from scratch, but all three scenarios must be considered. Below picture gives a high level overview of the architecture.

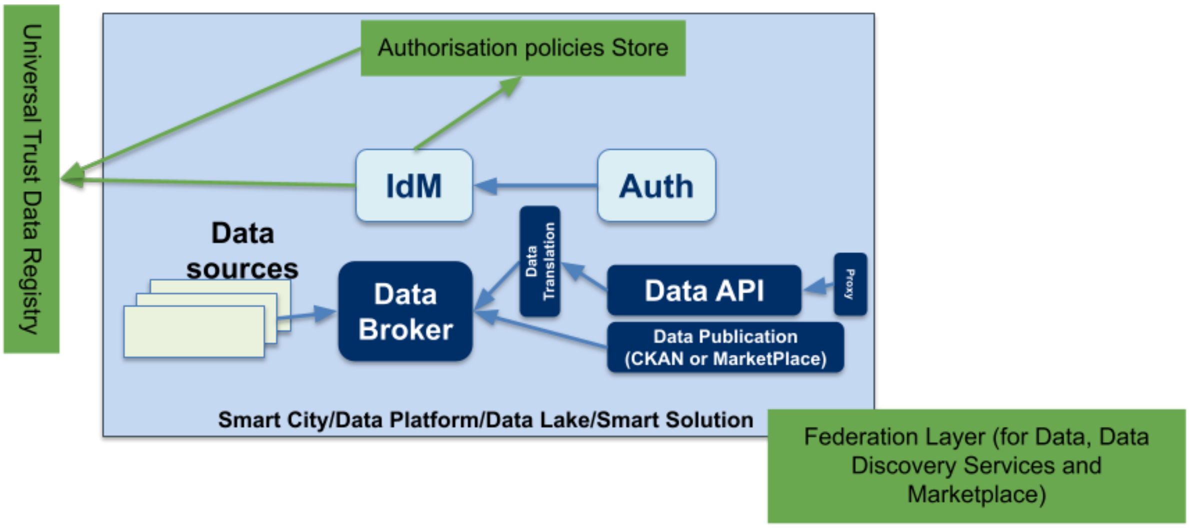

Universal Trust Data Registry
Authorisation policies Store
IdM
Auth
Data sources
Data Broker
Data Translation
Data API
Proxy
Data Publication (CKAN or MarketPlace)
Smart City/Data Platform/Data Lake/Smart Solution
Federation Layer (for Data, Data Discovery Services and Marketplace)

Figure 18. High level architecture view

The above high level architecture view is an evolution from the data platform which is described in Section (Existing Scenarios) to the Data Space. The Green parts are evolving the Platform to the Data Spaces.

Figure 19 shows how the proposed components in the high level architecture (in green) and the typical components deployed in smart city data platforms (in light and dark) are mapped into the DSSC taxonomy of building blocks. Thus, all essential building blocks for deploying a data space are in place.

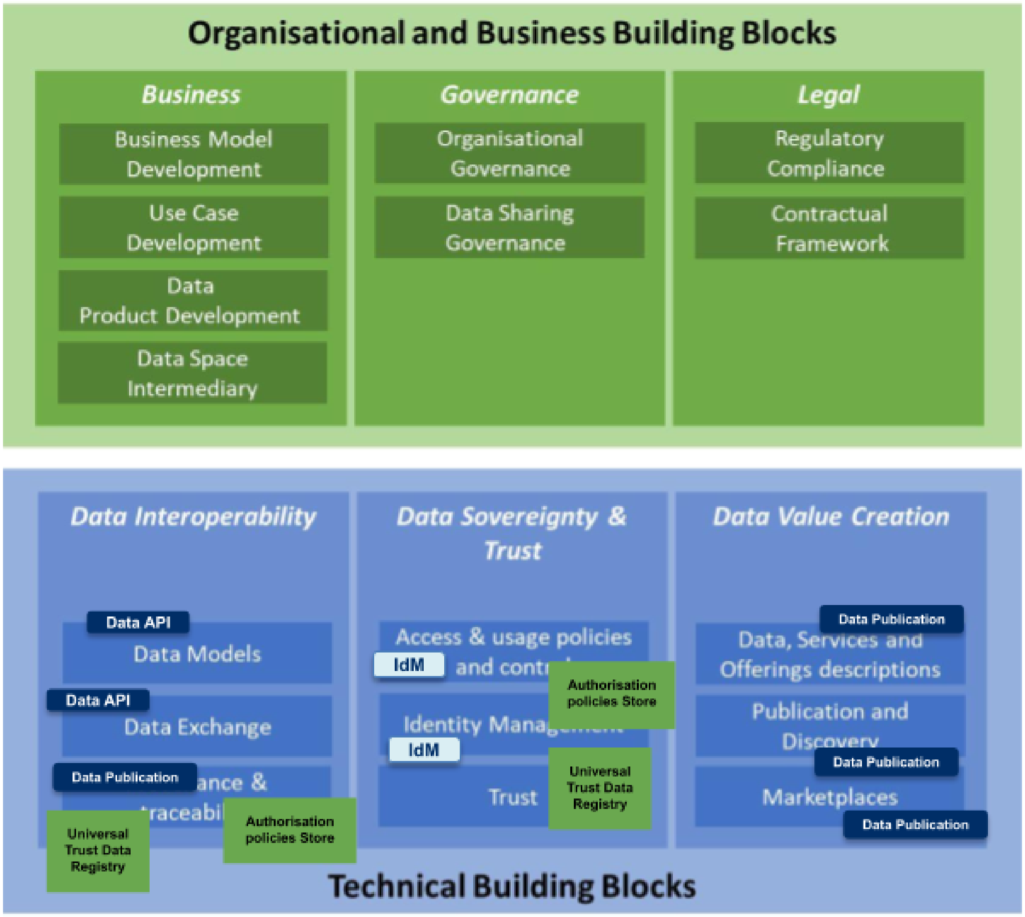


Figure 19. Mapping of data space components in the DSSC BBs taxonomy

3.3.1 Emulation of usage flow

In order to illustrate how the engagement process in a data space would work with the proposed architecture, Figure 20 shows the flow across the different components in the scenario where a data space participant orginsation's user (the Actor) is requesting for data.

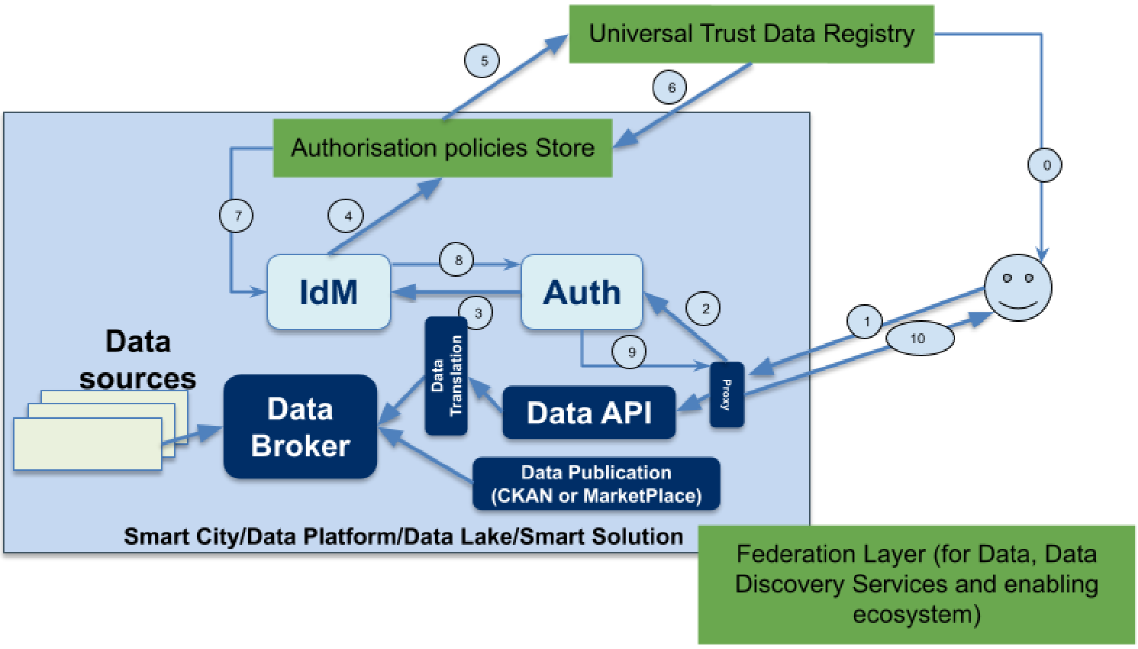

Universal Trust Data Registry
Authorisation policies Store
IdM
Auth
Data sources
Data Broker
Data Translation
Data API
Proxy
Data Publication (CKAN or MarketPlace)
Smart City/Data Platform/Data Lake/Smart Solution
Federation Layer (for Data, Data Discovery Services and enabling ecosystem)
0
1
2
3
4
5
6
7
8
9
10

Figure 20. Emulation flow in the use of the components

1. Starting a data request:
    a. Actor who is the user of the organisation which is participating in the Data Space gets its Verifiable Credentials (VCs) from the Universal Trust Registry through the IdM of the organisation, after registering with the Universal Trust Registry.
    b. Actor is requesting for data using his verifiable credentials obtained in step 1.a.
2. Proxy (PEP) requests Idm & Auth components for the verification of Actor's credentials and data access policy. These are in step 2 & 3
3. IdM requests Authorization policies store in step 4
4. Authorisation policies Store module sends a request to Universal Trust Data Registry to verify the credentials in step 5. Authorization policies Store has the Data access policies defined for the Data Space. In some scenarios step 5 can be also done by the IdM without going through the Authorization policy Store.
5. If the credentials are verifiable and are valid, then Universal Trust Data Registry returns approval and if not valid then rejection in step 6.
6. IdM gets the Actor's credentials verified and also the access policies of the Data Space for that Actor in step 7.
7. In step 8 and 9, the proxy determines if an actor's request can be allowed or not.

8. In step 10 the data is returned or Authorization denied based on the flow above.

In the above scenario, the Actor does not need to have a registration to the Data Space above, but can be part of another Data Space which is registered in the Universal Trust Data Registry. Universal Trust Data Registry will verify the credentials of the Actor based on the credentials provided by other Data Space.

3.3.2 Systems view

A Data Spaces System should have the following essential components.

1. **Data platform or digitalized services at the city infrastructure** - It is essential that the Data platform has standardised data (Interoperable with the other participants in the Data Space), an interface to the data or provides a service, which the data platform wishes to share with the other data space participants.
    a. Data platform should have an IdM (Identity Management) which is capable of integrating with the Authorisation policy store as it is described in section **Authorization policy store** section below.
    b. Data Space should have an interface and data standardised to be able to be interoperable with the other participants of the Data Space.
2. **Universal Trust Registry at the data space** - Every data space needs to have an Universal Trust Registry which validates the data platform which is participating in the data space. Universal Trust Registry validates the service/data provider and consumer.

Apart from the above mentioned technical components, it is essential to have a governance model and governance operation in place. The **Federation Layer** is not mandatory but it is described below in the next section.

This reference architecture uses concepts from the standard XACML (eXtensible Access Control Markup Language) architecture for authorization. XACML standard architecture comprises PEP, PDP, PIP and PAP components. Reference model should support attribute-based access control (ABAC) access control paradigm and static role-based access control (RBAC) permission model.

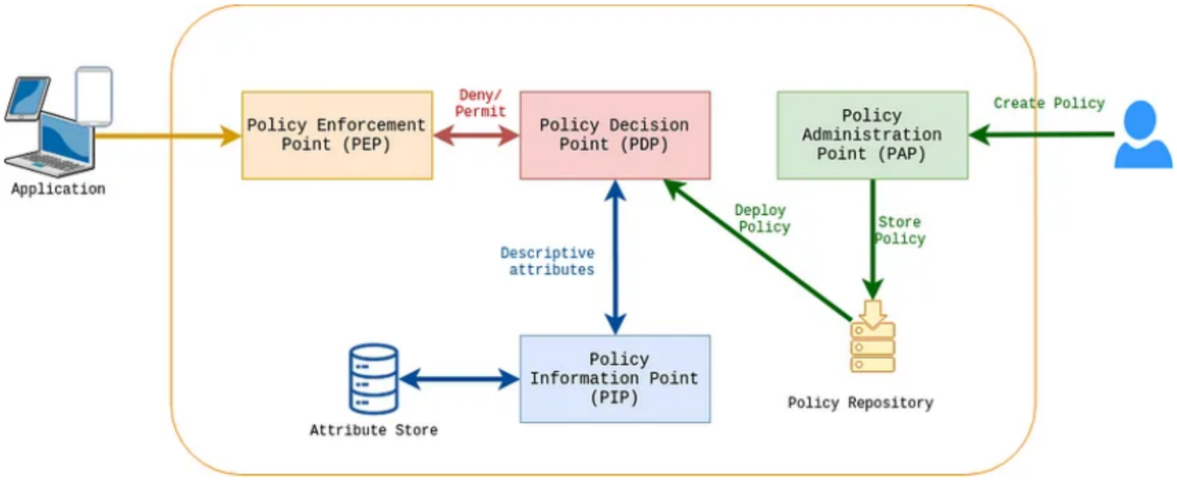

Application
Policy Enforcement Point (PEP)
Deny/ Permit
Policy Decision Point (PDP)
Policy Administration Point (PAP)
Create Policy
Deploy Policy
Store Policy
Descriptive attributes
Policy Information Point (PIP)
Attribute Store
Policy Repository

Figure 21. XACML architecture (source)

Thus, the Universal Trust Data Registry maps with PDP in the general sense and XACML components and Authorization policy store maps with PDP and PIP XACML components.

3.3.3 Components

Universal Trust Data Registry

Trusted exchange among Data Spaces provides certainty that participants involved in the data exchange are who they claim to be, and that they comply with defined rules/agreements. Trust refers to the fact that data providers and data consumers can rely on the identity of the members of the data ecosystem and beyond that, between different security domains.

Universal Trust Data Registry brings a scheme that enables organisations to give each other access to their data. It results in a set of agreements which improve circumstances for data exchange and focuses on the topics of Identification, Authentication and Authorisation.

Universal Trust Registry is an authentication & authorization protocol for both machine2machine (M2M) and human to machine (H2M) communication based on a JSON REST API architecture. Authentication is heavily based on Public Key Infrastructure (PKI) and therefore certificates and public / private key pairs. Keys are how the organisation is recognised and further the individuals are needed to use the organisation keys issued by any authority recognised or registered to Universal Trust Registry. The Universal Trust Registry Owner is playing the role of a trust authority providing a trusted framework which keeps the scheme, and its network of participants, operating properly. Every participant to the Registry must have a relation with the Owner, and can check at the Owner whether other parties participating in the Registry are trustable.

The Data Space governance authority defines the rules for a data space and therefore provides the governance framework of a data space. To do so, it makes use of a Data Space Registry, which manages the registration of participants in a data space based on the rules given. To enable cross data space interoperability as defined in ISO/IEC 21823-1:2019, a common governance and rules should be adopted by the Data Space governance authorities with the use of a common meta registry

Certification agencies that are registered in the Trusted Issuer List that is in the Dataspace Registry (another name for Universal Trust Registry in the context of Verifiable Credentials) of the Data Space. Other characteristics correspond to self-attested characteristics. In all the cases, each of these Product Specification Characteristics get mapped into a Verifiable Credential

The Data Space registry can be realised as a public or private registry and may make use of different measures to realise itself and the mechanisms for the identification of trusted participants. In the DSBA Technical Convergence document, such identification relies on the use of Verifiable Credentials (VCs) issued by Trusted Issuers registered in, or accredited via, the Data Space Registry. As per the IDSA RuleBook Data Space registry can have 3 different approaches which are Centralised, Decentralised and Federated approach.

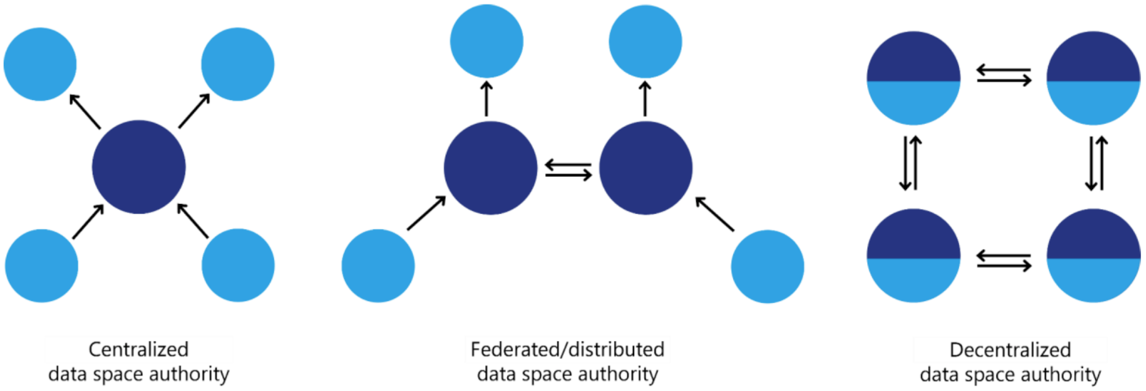


Figure 22. Data Space registry approaches

Universal Trust Registry should have the following components

- Public Key Storage
- Revocation List
- Trusted Participant List
- Trusted Issuers List
- REST API as an interface

On the side note DSBA proposes a decentralised Trust framework compatible with the EU DID Wallet Architecture and EBSI.

Decentralised Identity Management based on latest W3C and OIDC standards:

- W3C DID (Decentralised Identifiers), Verifiable Credentials (VC)
- Verifiable Credentials Issuance Protocols: OIDC4VCI
- Self-Issued OpenID Provider: SIOPv2
- Verifiable Credentials Exchange Protocols: OIDC4VP

Authorization framework following PEP-PDP-PIP and PRP/PAP architecture for ABAC (attributes ⇔ claims in VCs), and adopting ODRL as Policy Definition Language

Existing implementations

There are 2 alternatives (iSHARE Trust Framework, Gaia-X Trust Framework) available for a Universal Trust Registry component for the Data Spaces as described below.

1. **iShare Trust Framework** - iSHARE is the trust framework for data spaces, enabling trust and data sovereignty generically and cross sector. iSHARE's Satellites are autonomous data space coordinators, using iSHARE's ledger for participant registration and basic discoverability. Data Sovereignty is enabled by the Authorisation Registry providers in the iSHARE network, enabling third parties to company data throughout value chains. Hence The iSHARE trust framework is live in Logistics and Energy, and many more sectors are following. iSHARE trust framework is most used in the DataSpace and there are many existing deployments using iSHARE.

   iSHARE is built on principles of OpenDEI project definitions. iSHARE Trust Framework can be deployed on premise and also iSHARE organisation provides a node, which can be used as a Trust Framework.

   iSHARE brings together the following aspects to DataSpaces

   1. Federated Legal framework for all participants, optional data space specific legal
   2. Participant trust registration and administration
   3. Participant discovery and cross data space interoperability
   4. Standard and available service providers for Authorization Registries to save guard data sovereignty of the data owner
   5. Data space profile registration
   6. Trust Governance

Further more the details of deployment of iSHARE Trust Framework can be found at https://ishare.eu/about-ishare/the-foundation/governance/

i4Trust and ODALA projects have a reference deployment using iSHARE Trust Framework and reference implementations are available for everyone's use.

2. **Gaia-X Trust Framework** - In the Gaia-X Trust Framework, certificates that signal compliance can be awarded to any entity (participant, resource, service offering) in the form of verifiable credentials. These certificates of compliance are compulsory to be part of the Gaia-X ecosystem, and additional labels certifying compliance with specific rules (e.g., European Control, Art. 6 GDPR) are attainable for service offerings that have been audited and vetted. Gaia-X offers two main components for compliance: the Gaia-X Registry holding compliant services and participants, and the Compliance API, where certificates can be obtained and verified. Additionally, at the dataspace level, software components as well as operational environments can be certified, e.g. for the IDS standards using the IDS Certification Scheme. The Trust Framework foresees verifiable credentials and linked data representations as cornerstone of its future operations. Trusted information shall be retrieved in machine readable manners, and where such manners are missing, Gaia-X will define processes to translate trusted information in a machine readable format. This is a prerequisite of federating trusted statements within the Gaia-X Ecosystem and developing mechanisms to re-assess validity of claims within the Trust Framework.

   Gaia-X Trust Framework makes helps being compliant with Gaia-x compliance requirements:

   1. API of Gaia-X Trust Framework can be found at https://www.gaia-x.at/wp-content/uploads/2023/04/WhitepaperGaiaX.pdf
   2. Hosted Gaia-X nodes can be found in the Gaia-X Lab Registry Service that is designed to be used by the Gaia-X Lab Compliance Service. However, an API is exposed for the registry as well, to get the content of the registry as well as to verify the validity of signed claims (e.g., Self Descriptions) by checking the provided certificates against Gaia-X endorsed Trust Anchor certificates.
   3. This white paper gives a good overview and direction about Building DataSpace using the Gaia-X Trust Framework.

Usage in Data Spaces

Universal Trust Registry can be deployed for a Data Space or can be deployed as a common component for multiple Data Spaces or the supplier's node can be used. The deployment is specific to the Trust Frameworks. Some general steps include:

- Registering the Data Space with the Universal Trust Registry. This step is previous to any engagement of a participant in a data space.
- Data Space administrator can register participants in a specific step depending on the Trust Framework used
- To allow for discovery and cross domain interoperability to find where data services from a participant are reachable and what standards are used.
- Authorization end-point, to find where a participant has their authorizations available for querying.

The detailed steps are available for iSHARE at https://ishare.eu/about-ishare/benefits/for-developers/.

Authorization policy store

Authorization policy store is closely related with identity management component. Since the architectural requirement of a data space solution is that each participant can use its own identity management and authorization module, reference architecture must support decentralised identity management and authorization.

The two most important questions that are addressed during the authentication process are 1) is the party who is trying to access the resource really who it says it is, 2) does this party have granted access to the resource according to active access policies. Because of the decentralised nature of the data space solution these questions are raised multiple times per request when a subcomponent of one system is accessing a subcomponent of another system. Then an authorization request is either confirmed or denied access to the requested resource.

There are different options individual data space participants can use to participate in the authorization flows of the data space solution with their systems. It is still not decided which permutations of the possible identity management and authorization options between participating systems in the data space solution will be supported in the proposed reference architecture.

Supported identity management and authorization options

Each participant system needs to include a local identity management component (for authentication support) and an authorization policy store (for authorization support). Additionally, a data space solution needs a universal trust data registry (for trust support), which establishes trust between included participant systems. It is used to verify identities by any data space participant. Ideally, there should be many different instances of the universal trust data registry operated by different entities in the data space, because having just one entity/instance increases the risk of centralisation.

Decentralised IAM based on OIDC presents a straightforward data space authorization and authentication architecture. The Authorization Policy Store maps with PDP XACML component and uses PIP XACML component to retrieve additional information needed. Authorization Policy Store should support the following access controls: ABAC and RBAC. Identity management can be provided using existing IAM solutions (e.g. Keyrock, Keycloak or similar). Authorization policy store component can already be embedded in identity management solution or it can be introduced as a standalone component (e. g. iSHARE Authorization Registry). The universal trust data registry component should be centralised (e.g. iSHARE Satellite).

Decentralised IAM based on DID and VC/VP presents an upgraded data space authorization and authentication architecture, which is still under development. Central universal trust data registry component (e.g. Universal Resolver from the Decentralised Identity Foundation Identifiers & Discovery Working Group) will be used to verify identities by any data space participant. W3C Verifiable Credentials with DIDs as identifiers will be used for the authentication and authorisation. Bidirectional mechanism to derive DIDs from the eIDAS digital certificate should be used, so there is no need to invent new identifiers or have a central entity in a data space assigned identifiers to participants.

Detailed descriptions of authorization flows exceed the scope of this document and can be observed in DSBA Technical Convergence[12] and i4Trust Building Blocks[13] documents.

Evolution from existing local data platform to data space

The following section describes the key prerequisites and required activities that enable transition from the existing data platform to data space. The document will only focus on transition to decentralised IAM based on OIDC option, since decentralised IAM based on DID and VC/VP option is still work in progress.

Each participant system that would like to become compatible with the decentralised IAM based on OIDC option, should have at least OIDC based identity management with added user and organisation identities. Key steps, needed for the transition are the following:

- deploy and activate universal trust data registry component if none exists (e.g. e.g. iSHARE Satellite),
- pre-register at the universal trust data registry component (e.g. iSHARE Satellite),
- deploy and activate local authorization policy store component (e.g. iSHARE Authorization Registry),
- integrate existing IAM solution to be compliant with the new decentralised IAM based on OIDC authorization flows,

[12] DSBA Technical Convergence Discussion Document, Version 2.0, 2023-04-21

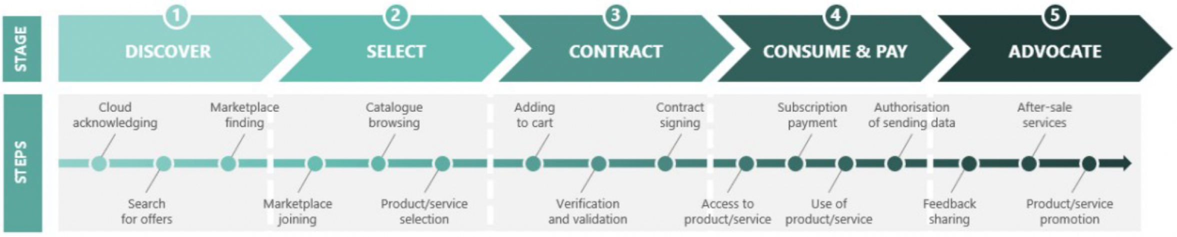
[13] i4Trust Building Blocks (version 4.0 - under review)

- connect data sources with a Federation Layer via Data Translation Component that will be responsible for exchanging data within the data space (e.g. existing solution might not use standardised data exchange and translation for communication within the data space is needed),
- add Policy Enforcement Point (PEP) and Policy Decision Point (PDP) capabilities to the domain application architecture,
- configure ABAC and RBAC access policies at the PEP proxy and PDP,
- upgrade authorization logic of the domain application (e. g. marketplace) to be compliant with the new decentralised IAM based on OIDC authorization flows,
- add logic to the administration console of the domain application to support Policy Administration Point (PAP) and Policy Information Point (PIP) capabilities of the new decentralised IAM based on OIDC authorization flows;

Federation layer

The Federation Layer (or federated services and components) in the architecture is formed by the components that users need to access the services and data offered by the participants of another Data Space, without compromising the individual data sovereignty. These services are mainly about Catalogue Services, Marketplace Services and Metadata Broker Services. It can be implemented through several mechanisms, for example through Data or Services Publication Platforms. They can be offered by data space intermediaries or fully decentralised based on P2P mechanisms.

A typical customer service or data is as shown in the picture below for acquiring data or services in the Data Space.



As per the IDSA Rule book, The Dataspace Protocol is a set of specifications designed to facilitate interoperable data and services sharing between entities governed by usage control and based on web technologies developed under the umbrella of IDSA. These specifications define the schemas and protocols required for entities to publish data and offer services from the participating organisations in the Dataspace. This component is used to negotiate usage agreements, and access data or services as part of a federation of technical systems termed a dataspace.

The Federation Layer should support the offering of data resources and services under defined terms and conditions, including applicable pricing models, and marketplaces for services and the data must be established. This component supports the 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. The component should also support data discovery services. The following features are desired:

1. Publication and Query
2. Standard information model and supporting APIs for the implementation of data marketplace services
3. Backend components implementing marketplace services
4. Data Catalogue / Publication functions to publish data resources which can be found via metadata and are connected with marketplace
5. (Portal) Public marketplace human readable information and marketing things, landing page
6. Metadata and data sets Publication and Discovery
7. Data Services Marketplaces
8. Data usage accounting

Some of the cloud or edge data services registered in the data space may bring access to static data or near real-time data resources available through RESTful APIs (e.g., IoT data). The data space will integrate data publication functions enabling the exposure of such data resources in compliance with DCAT specifications defined by W3C and DCAT-AP recommendation by the EC. This way, data resources linked to data services offered through Data Space can be harvested through external Data Publication platforms (e.g., the European Data Portal) . This will enable interoperability with Data Publication Platforms.

An example of marketplace services can be found at FIWARE, built on top of the FIWARE BAE (Business Application Ecosystem) component, a combination of the FIWARE Business Framework and a set of APIs provided by the TMForum. It allows the monetization of different kinds of assets during the whole service life cycle, from offering creation to its charging, accounting and revenue settlement required for billing and payment to involved participants.

An evolutionary concept of marketplace is under development in the Decentralised Open Marketplace Ecosystem (DOME) project. It is a very promising evolution of Data or Service discovery and publication platforms. Further it is more explained in the following section.

3.4 Evolution of Smart Solution/Data Platform to Data Spaces

The architecture defined in this document gives an evolution of smart platforms to data spaces. The core purpose of the platforms should not be compromised, but should enable the data space functionalities like Data Sovereignty and Trust, Data

Interoperability and Data Value Creation across the data space. One of the example use cases is that the users of the platforms should be able to access the services and data of the other platforms in the data space with their own credentials of the platform which they have already been the users of.

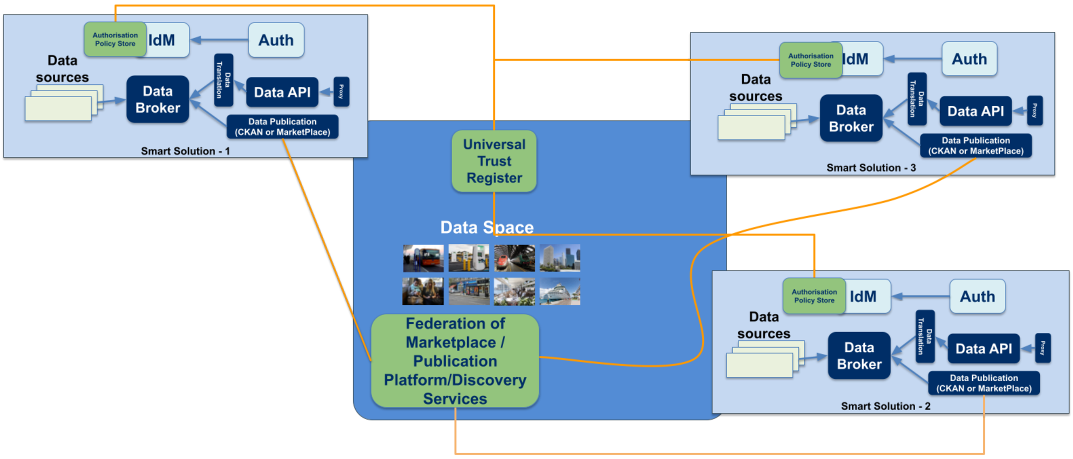


Figure 23. Evolution of smart city platforms towards data spaces approach

Each participant in the data space solution (e.g. marketplace) can use an independent identity management and authorization model as shown in the picture above. Due to this architecture requirement data space solution needs to support decentralised identity management and authorization model.

3.5 Future evolutions

This section describes the evolving technology and related components, which has a strong potential to influence the data spaces and the way the data platforms can choose to participate in a data space.

3.5.1 Data Space Connector

The concept of Data Space Connector has evolved to match the idea of an integrated suite of components every city/community participating in a data space should deploy to "connect" to the data space. Connector gives a one component functionality. It simplifies the participation of a platform to a data space easily with a streamlined way of deployment of the components needed and configuring. The functionalities are implemented in individual micro services or as a single comprehensive software block. In addition, the services do not have to be deployed in the same infrastructure.

The architecture proposed is compatible with the Data Space Connector defined and specified in Data Space Connector in IDS RAM 4.0. The components which needed to make a platform participate in the Data Space are put together as a Data Space Connector.

Compliance of Data Space Connectors with DSBA recommendations

Aligning with DSBA recommendations would have several implications:

1. Interface with Trust Services should align with EBSI specifications (DID-Registry, Trusted-Issuers-Registry APIs but extended to support authentication based on Verified Credentials)
2. Authentication should be based on W3C DID + VC/VP standards and SIOPv2/OIDC4VP protocols and implement the connection to trust services
3. Authorization should implement a PxP architecture implementing ABAC using ODRL as policy language
4. Compatibility with NGSI-LD as data exchange API
5. Contract Management is under analysis since there are two approaches to reconcile, TM Forum APIs would be a good candidate for Contract Management API and there has been some initial work in IDS RAM 4.0 regarding specification of a Contract Management protocol

This is still an evolving concept when this document is being written and if necessary a new revision on this document shall be made to accommodate Data Space Connector in the architecture as follows below. This concept will be reviewed under the deployment phase of the data space (DS4SSCC-DEP project).

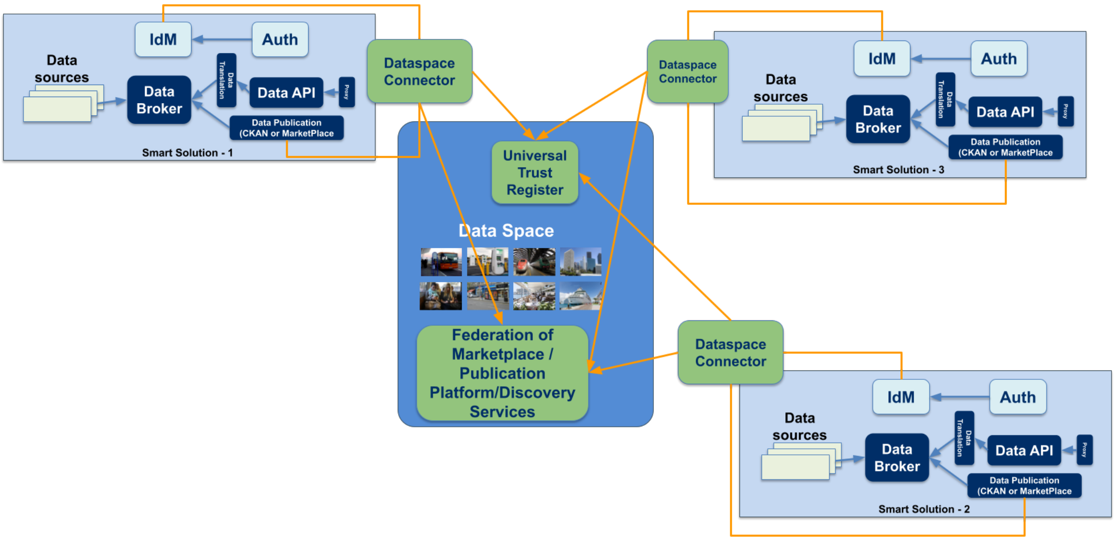


Figure 24. Proposed insertion of Data Space Connector into DS4SSCC architecture

An example of a Data Space Connector can be found in FIWARE Foundation at https://github.com/FIWARE-Ops/data-space-connector.

3.5.2 Evolution of Verified Credentials (VC)

VerifiableCredentials provide a mechanism to represent information in a tamper-evident and therefore trustworthy way. The term "verifiable" refers to the characteristic of a credential being able to be verified by a 3rd party(e.g. a verifier). Verification in that regard means, that it can be proven, that the claims made in the credential are as they were provided by the issuer of that credential. These characteristics make VerifiableCredentials a good option to be used for authentication and authorization, as a replacement of other credentials types, like the traditional username/password. The SIOP-2/OIDC4VP standards define a flow to request and present such credentials as an extension to the well-established OpenID Connect.

In connection to the VCs, DSBA has been working on a new Decentralised Authorisation Framework, which is compatible with EU DID Wallet Architecture and EBSI. This is inline with the eIDAS regulation which is expected to be adopted on 23rd of July, 2024. The proposed Decentralised Identity Management based on latest W3C and OIDC standards:

1. W3C DID (Decentralised Identifiers), Verifiable Credentials (VC)
2. Verifiable Credentials Issuance Protocols: OIDC4VCI
3. Self-Issued OpenID Provider: SIOPv2
4. Verifiable Credentials Exchange Protocols: OIDC4VP

The authorization framework would be following following PEP-PDP-PIP and PRP/PAP architecture for ABAC (attributes ⇔ claims in VCs), and adopting ODRL as Policy Definition Language

The VCVerifier component (under evolution in the i4Trust project) provides the necessary endpoints required for a Relying Party(as used in the SIOP-2 spec) to participate in the authentication flows. It verifies the credentials by using WaltID SSIkit as a downstream component to provide Verifiable Credentials specific functionality and return a signed JWT, containing the credential as a claim, to be used for further interaction by the participant.

The development of this component can be followed at https://github.com/FIWARE/VCVerifier

3.5.3 Evolution of Federated Marketplace from DOME Project

Data spaces should provide support for the creation of multi-sided markets where participants can generate value out of sharing data. 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 supporting the lifecycle of contracts that are established when a given participant acquires the rights to use a service, according to certain service offering.

The proposed approach will take the form of a Decentralised Open Marketplace Ecosystem (DOME) based on the federation of marketplaces, all of them connected to a commonly shared digital catalogue of cloud and edge services and service offering descriptions.

Each of the federated marketplaces in the referred DOME will be a marketplace provided by an independent marketplace provider or a marketplace connected to the offering of a given cloud / edge infrastructure service provider (IaaS or platform provider). Besides these marketplaces, A DOME global portal would implement functions through which cloud/edge service providers may register their product offerings and end customers can discover offered products.

DOME will rely on the adoption of common open standards for the description of cloud and edge services and service offerings as well as their access through a shared catalogue.

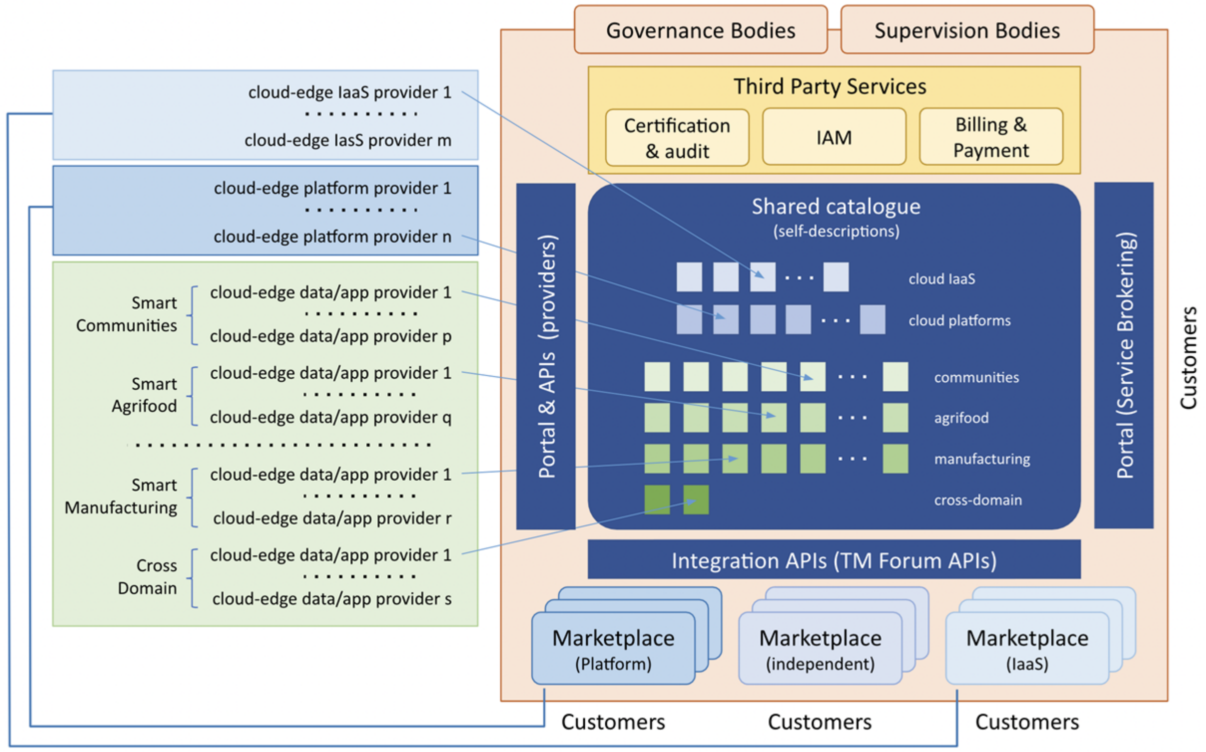

Governance Bodies
Supervision Bodies
Third Party Services
Certification & audit
IAM
Billing & Payment
Shared catalogue
(self-descriptions)
cloud IaaS
cloud platforms
communities
agrifood
manufacturing
cross-domain
Portal & APIs (providers)
Portal (Service Brokering)
Customers
Integration APIs (TM Forum APIs)
Marketplace (Platform)
Marketplace (independent)
Marketplace (IaaS)
Customers
Customers
Customers
cloud-edge IaaS provider 1
cloud-edge IasS provider m
cloud-edge platform provider 1
cloud-edge platform provider n
Smart Communities
cloud-edge data/app provider 1
cloud-edge data/app provider p
Smart Agrifood
cloud-edge data/app provider 1
cloud-edge data/app provider q
Smart Manufacturing
cloud-edge data/app provider 1
cloud-edge data/app provider r
Cross Domain
cloud-edge data/app provider 1
cloud-edge data/app provider s

Figure 25. DOME marketplace

The more progress about DOME can be followed at https://dome-marketplace.eu/

3.6 Customised architectures

To facilitate the adoption of the proposed High Level Architecture in section 3.4, we have selected 4 representative use cases of the typical scenarios mentioned in section 3.1. They are all brownfield scenarios since they already have a city data platform or at least a set of digitalized services. However, Flanders and Helsinki are starting to go further and quite close to the Digital Twin scenario.

For these selected use cases, we have customised the high level architecture and provided options on how to move from current state to a data space approach. Therefore, we needed more technical details on each specific use case, so we prepared a technical questionnaire and requested some documentation about their existing technical deployment in the city or the region. The requested questionnaire can be consulted in the Annex I.

The proposed architectures have been validated with the use cases and their feedback is reported in the section 4.4.

3.6.1 Customised Architecture (Helsinki Real-time Data Space)

3.6.1.1 Scenario description

In previous tasks we were working with Helsinki on Heat and Climate Atlas. During the workshop in this work package, we realised that the use case is not appropriate for a data space, as it includes only open data, there is no trust data registry nor authorisation needed. Additionally, this use case is not a high priority for Helsinki and it is in maintenance mode. Therefore, Helsinki proposed a new use case that is setting up the data space as an infrastructure to support different use cases, primarily for real-time data usage in several use cases from different domains.

Although Helsinki has already collected and used lots of data in many scenarios,they are still tackling aspects of urban data collection and usage:

- **Urban space management:** Planning and managing urban space efficiently is becoming increasingly important for the city. This involves considerations for amenities, environmental impact, accessibility, and the conditions for businesses. To address these challenges, Helsinki needs more precise and real-time data.
- **Real-time data utilisation**: While there is an abundance of real-time data available, its usability for decision-making and innovation in services remains low. This is often due to the lack of guidance, harmonisation, regulation, governance, and technical standards. Establishing a well-structured data space can help improve this situation.
- **Opportunities with data act:** The Data Act opens up opportunities for Helsinki to leverage new sensor data sources and work towards creating a

fair and harmonised urban data space. This can enable better data utilisation and decision-making.

- **Advanced digital twins and 3D models:** Helsinki's digital twins and 3D city models are among the most advanced globally. These not only provide virtual representations of the city but also serve as a sophisticated open platform for integrating real-time sensor data with static environmental features.

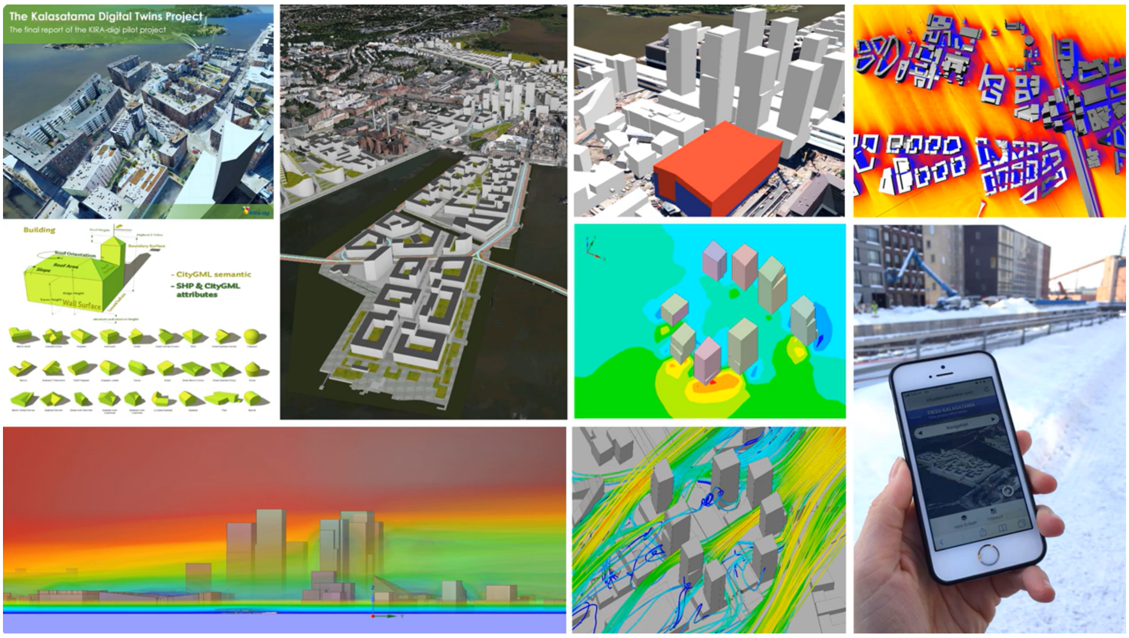

The Kalasatama Digital Twins Project
The final report of the KIRA-digi pilot project
Building
- CityGML semantic
- SHP & CityGML attributes

Figure 26. Helsinki Digital twin and 3D city model presented on youtube channel *https://www.youtube.com/watch?v=Qg6R7hrRVv0*

- **High trust in public institutions:** The high level of trust that Finns have in public institutions creates a conducive environment for data sharing and collaboration, promoting data altruism among citizens.
- **Cross-domain use cases:** Helsinki, along with Forum Virium Helsinki (FVH), has experience in various cross-domain use cases related to real-time data. These include initiatives such as last-mile logistics automation pilots, city logistics hubs, and clean air routing, demonstrating the city's commitment to utilising data for innovative solutions.

Helsinki's real-time data space holds great potential for optimising the city's physical space and improving various aspects of urban life. Here are some specific use cases and benefits:

- **Cross-domain data harmonisation**: Harmonising data across domains and enabling easy access and control of data for various purposes.
- **Micro-positioning / user-centric positioning:** Enhancing location-based services and navigation for residents and visitors.
- **Logistics:**

    - Automating last-mile logistics for efficient delivery services.
    - Planning and optimising urban logistics hubs to reduce congestion and improve delivery efficiency.
    - Context-aware routing for in-city logistics to minimise traffic and environmental impact.
- **Situational awareness:**
    - Automated traffic rerouting for momentary exceptions, like a ferry unloading at port.
    - Real-time updates for rescue routes.
    - Improved routing for individuals with disabilities based on accessibility and current conditions.
    - Visualising indoor and outdoor environmental conditions, energy consumption, and production.
- **Built environment optimisation:**
    - Optimising various parameters within the build environment such as temperature, warming, lighting and access control.
    - Personalising and utilising digital signage.
- **City planning:**
    - Data-driven downtown revitalization efforts to enhance the comfort and attractiveness of the city centre, benefiting local businesses.

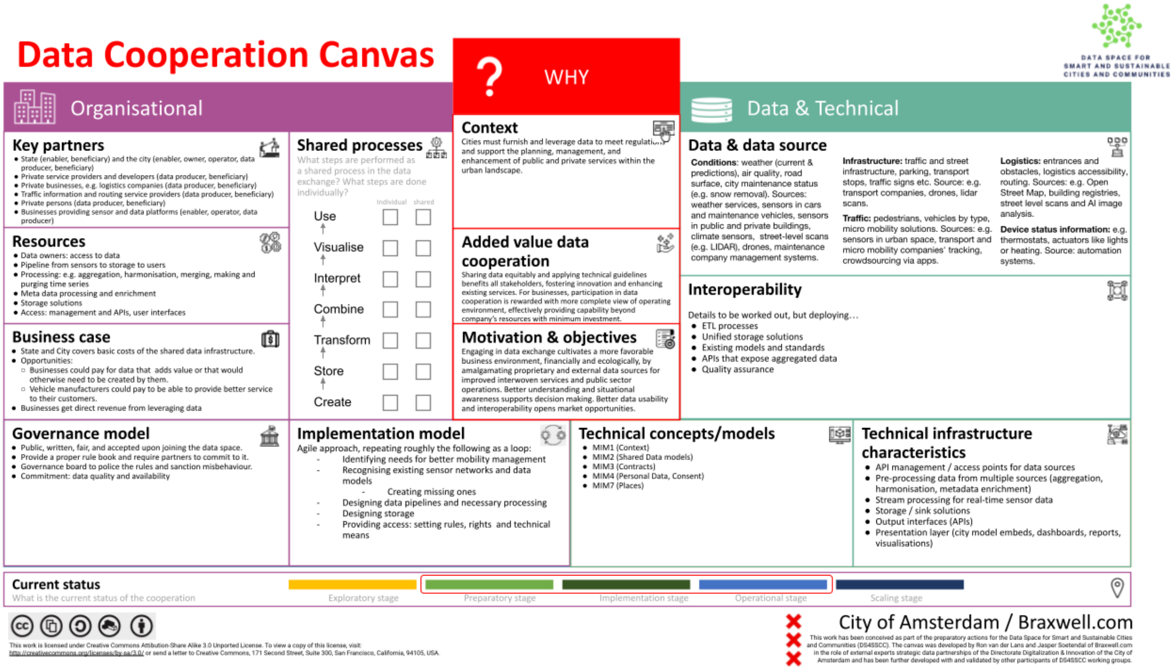
These use cases offer a wide range of benefits to citizens and communities, from improved mobility and safety to a cleaner environment and enhanced quality of life. They also support economic growth and sustainable urban development, making the city a better place to live, work, and visit for everyone.

The benefits of these use cases extend to various stakeholders, including bicycle traffic planners, traffic managers, port and traffic light controllers, clean air route guidance providers, the City of Helsinki, and private companies operating in the city centre. These benefits include increased efficiency, reduced environmental impact, improved urban planning, and enhanced quality of life for residents and visitors.

Looking to the future, the vision of enabling two-way communication and control of the physical environment across assets managed by diverse stakeholders, both public and private, represents a significant step toward creating a more connected, responsive, and sustainable urban environment in Helsinki.

3.6.1.2 Data Cooperation Canvas

As explained in section 1.3, an instrument referred as Data Cooperation Canvas has been created by the DS4SSCC to facilitate the cities and communities to consolidate in one page chart all the aspects of their data space. In the case of Helsinki, the result is presented below.



DATA SPACE FOR SMART AND SUSTAINABLE CITIES AND COMMUNITIES

Data Cooperation Canvas

Organisational

Key partners
- State (enabler, beneficiary) and the city (enabler, owner, operator, data producer, beneficiary)
- Private service providers and developers (data producer, beneficiary)
- Private businesses, e.g. logistics companies (data producer, beneficiary)
- Traffic information and routing service providers (data producer, beneficiary)
- Private persons (data producer, beneficiary)
- Businesses providing sensor and data platforms (enabler, operator, data producer)

Resources
- Data owners: access to data
- Pipeline from sensors to storage to users
- Processing: e.g. aggregation, harmonisation, merging, making and purging time series
- Meta data processing and enrichment
- Storage solutions
- Access: management and APIs, user interfaces

Business case
- State and City covers basic costs of the shared data infrastructure.
- Opportunities:
  - Businesses could pay for data that adds value or that would otherwise need to be created by them.
  - Vehicle manufacturers could pay to be able to provide better service to their customers.
- Businesses get direct revenue from leveraging data

Shared processes

What steps are performed as a shared process in the data exchange? What steps are done individually?

	Individual	shared
Use	□	□
Visualise	□	□
Interpret	□	□
Combine	□	□
Transform	□	□
Store	□	□
Create	□	□

? WHY

Context

Cities must furnish and leverage data to meet regulations and support the planning, management, and enhancement of public and private services within the urban landscape.

Added value data cooperation

Sharing data equitably and applying technical guidelines benefits all stakeholders, fostering innovation and enhancing existing services. For businesses, participation in data cooperation is rewarded with more complete view of operating environment, effectively providing capability beyond company's resources with minimum investment.

Motivation & objectives

Engaging in data exchange cultivates a more favorable business environment, financially and ecologically, by amalgamating proprietary and external data sources for improved interwoven services and public sector operations. Better understanding and situational awareness supports decision making. Better data usability and interoperability opens market opportunities.

Data & Technical

Data & data source

Conditions: weather (current & predictions), air quality, road surface, city maintenance status (e.g. snow removal). Sources: weather services, sensors in cars and maintenance vehicles, sensors in public and private buildings, climate sensors, street-level scans (e.g. LIDAR), drones, maintenance company management systems.

Infrastructure: traffic and street infrastructure, parking, transport stops, traffic signs etc. Source: e.g. transport companies, drones, lidar scans.

Traffic: pedestrians, vehicles by type, micro mobility solutions. Sources: e.g. sensors in urban space, transport and micro mobility companies' tracking, crowdsourcing via apps.

Logistics: entrances and obstacles, logistics accessibility, routing. Sources: e.g. Open Street Map, building registries, street level scans and AI image analysis.

Device status information: e.g. thermostats, actuators like lights or heating. Source: automation systems.

Interoperability

Details to be worked out, but deploying…
- ETL processes
- Unified storage solutions
- Existing models and standards
- APIs that expose aggregated data
- Quality assurance

Governance model
- Public, written, fair, and accepted upon joining the data space.
- Provide a proper rule book and require partners to commit to it.
- Governance board to police the rules and sanction misbehaviour.
- Commitment: data quality and availability

Implementation model

Agile approach, repeating roughly the following as a loop:
- Identifying needs for better mobility management
- Recognising existing sensor networks and data models
  - Creating missing ones
- Designing data pipelines and necessary processing
- Designing storage
- Providing access: setting rules, rights and technical means

Technical concepts/models
- MIM1 (Context)
- MIM2 (Shared Data models)
- MIM3 (Contracts)
- MIM4 (Personal Data, Consent)
- MIM7 (Places)

Technical infrastructure characteristics
- API management / access points for data sources
- Pre-processing data from multiple sources (aggregation, harmonisation, metadata enrichment)
- Stream processing for real-time sensor data
- Storage / sink solutions
- Output interfaces (APIs)
- Presentation layer (city model embeds, dashboards, reports, visualisations)

Current status

What is the current status of the cooperation

City of Amsterdam / Braxwell.com

This work has been conceived as part of the preparatory actions for the Data Space for Smart and Sustainable Cities and Communities (DS4SSCC). The canvas was developed by Ron van der Lans and Jasper Soetendal of Braxwell.com in the role of external experts strategic data partnerships of the Directorate Digitalization & Innovation of the City of Amsterdam and has been further developed with and validated by other participants of DS4SSCC working groups.

Figure 27: Helsinki Data Cooperation Canvas

3.6.1.3 Technical requirements

Existing Helsinki solution users are:

- state (enabler, beneficiary) and the city (enabler, owner, operator, data producer, beneficiary),
- private service providers and developers (data producer, beneficiary),
- private businesses, e.g., logistics companies (data producer, beneficiary),
- traffic information and routing service providers (data producer, beneficiary),
- private persons (data producer, beneficiary),
- businesses providing sensor and data platforms (enabler, operator, data producer).

The Helsinki solution is cross-domain with the following domains using data exchange:

- transportation and mobility (include vehicle positioning, estimation of timetables and routing, car parking, EV charging, shared cars or bikes),
- environment and weather (include sound, air pollution and general weather sensor data),
- energy (3D map data using CityGML),
- urban planning (include also data from events).

The existing IAM solution is not standardised. It is centralised and only used internally as most of the existing data exchange is available as open data.

There are currently no interoperable data models in use, with the exception of CityGML, which is used to store 3D digital models of the city and its landscapes. In addition, not all data is exposed with APIs.

3.6.1.4 Scenario architecture

Existing Helsinki solution architecture is built around the data streaming concept. Data flows from sensors and other systems on the left-hand side, transforms its format for compatibility reasons and stores the data in different data stores with the help of streaming components. On the right-hand side GeoServer component consumes the data and displays spatial information to the world.

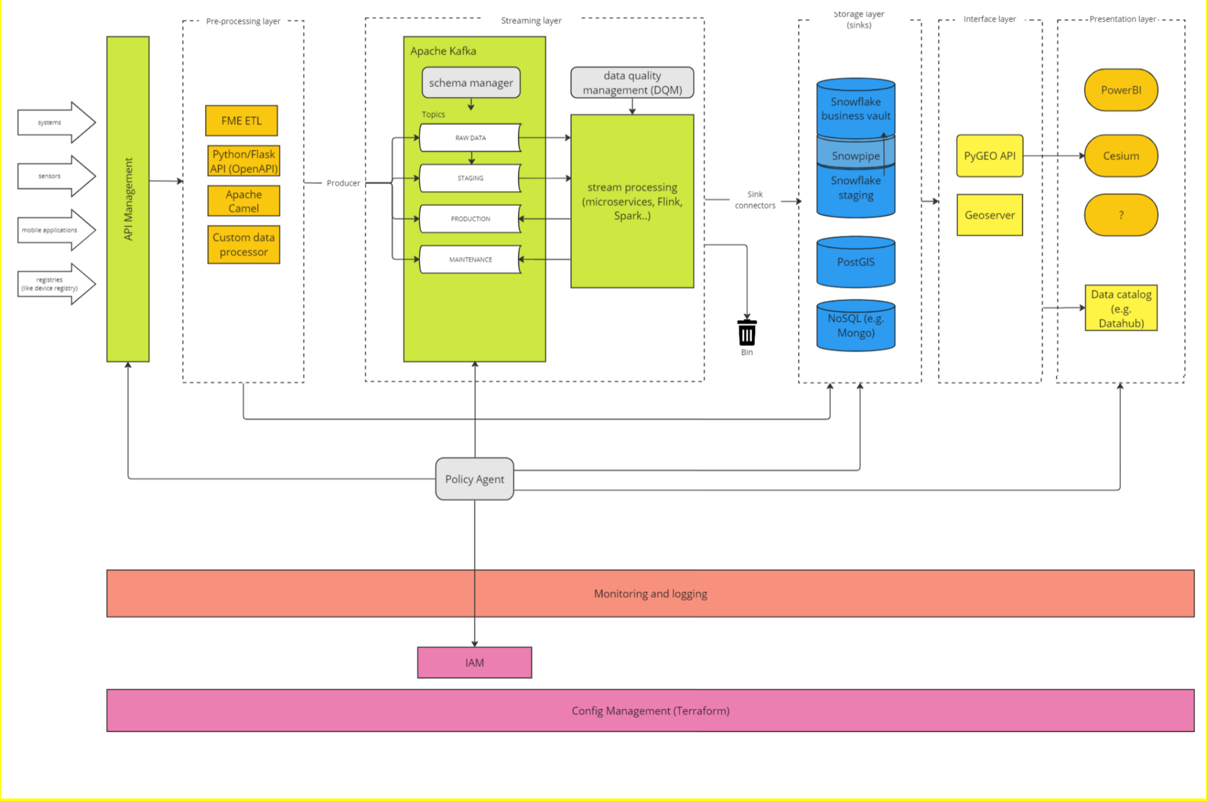


Figure 28. Existing high-level architecture of Helsinki city solution

To make existing architecture compliant with the Architecture Model, several components need to be added/integrated: migration to the compatible IAM solution, integration with Universal Trust Data Registry component and Federation Layer and implementation and configuration of Authorization Policy Store. In order to support interoperable data exchange between data space ecosystem solution and other dataspace participants, a custom Data Translation Component had to be introduced. Data Translation Component is not a part of the Architecture Model, since it needs to be implemented by every stakeholder that uses proprietary data models and would like to join the data space.

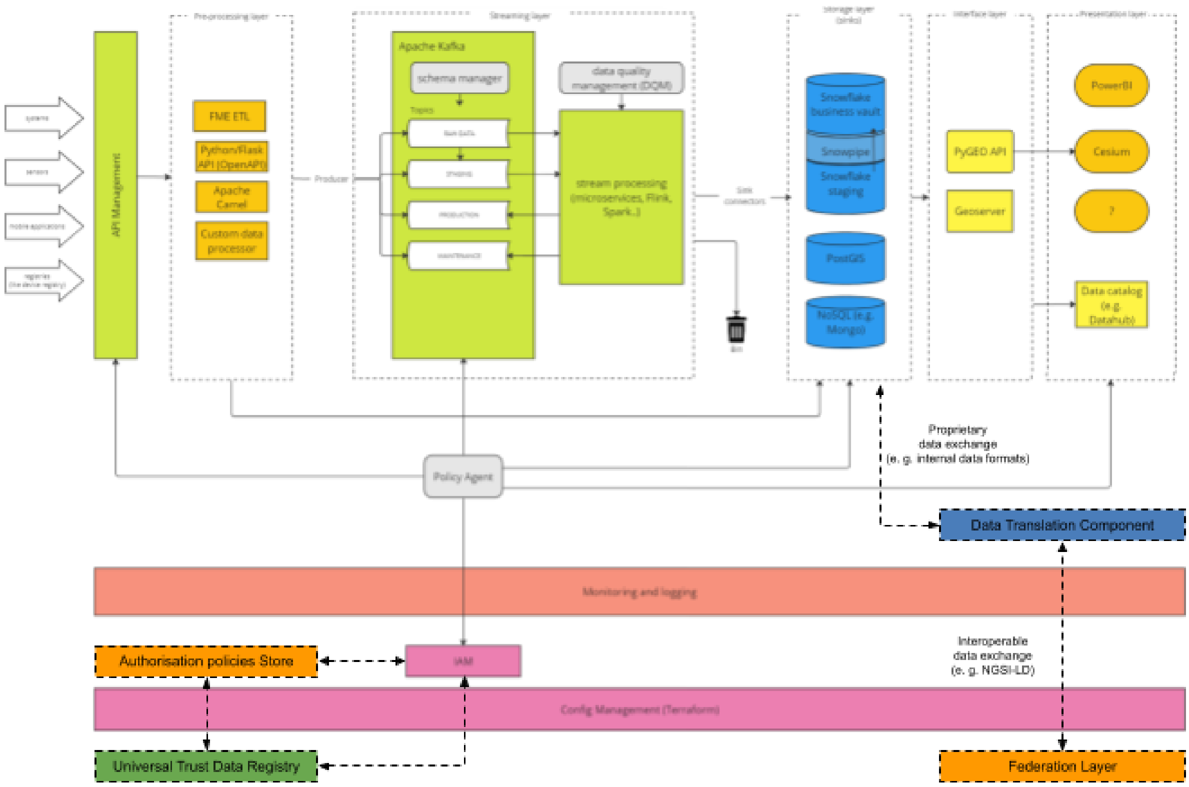


Figure 29. Customised high-level architecture of Helsinki data space

Proposed customised architecture will enable Helsinki solution to participate in the larger dataspace ecosystem. Added components, that enable integration into larger data spaces ecosystem, are enclosed with dashed line. Type of added components is differentiated by colour:

- blue dashed components: should be implemented inside each individual data space ecosystem solution,
- orange dashed components: provided by data spaces Architecture Model reference implementation and deployed on the individual data space ecosystem solution level,
- green dashed components: provided by data spaces Architecture Model reference implementation and deployed on the data space ecosystem level.

3.6.1.5 Implementation steps

In order to make the existing Helsinki solution compatible with the Architecture Model, the following steps need to be taken:

- deploy and activate universal trust data registry component (e.g. iSHARE Satellite),

- pre-register at the universal trust data registry component (e.g., iSHARE Satellite),
- deploy and activate local authorization policy store component (e.g., iSHARE Authorization Registry),
- implement Data Translation Component that will transform solution data to the format, compatible with data space (e.g., NGSI-LD),
- introduce Federation Layer for exchanging data with other data space participants,
- introduce IAM solution, compatible with the new decentralised IAM based on OIDC authorization flows,
- add Policy Enforcement Point (PEP) and Policy Decision Point (PDP) capabilities to the domain application architecture,
- configure ABAC and RBAC access policies at the PEP proxy and PDP,
- upgrade authorization logic of the domain application (e. g. marketplace) to be compliant with the new decentralised IAM based on OIDC authorization flows,
- add logic to the administration console of the domain application to support Policy Administration Point (PAP) and Policy Information Point (PIP) capabilities of the new decentralised IAM based on OIDC authorization flows.

3.6.2 Customised Architecture 2 (Flanders Smart Data Space)

3.6.2.1 Scenario description

In Flanders' case, they already have an emerging data space which is integrating diverse data platforms and data sources in the region.

The Flanders Smart Data Space is using the concepts of event streaming and linked data to enable the concept.

- Based on the need of more actual and real time data the Flanders Smart Data space uses the event streaming technology to make real-time data and the context information available for the ecosystem.
- The Flanders Smart Data Space is currently under development and is focussing on two pilot projects in the mobility and water domain as well as supportive work in the emerging domains of e-gov, energy, geographical data amongst others, which will help shape the eventual architecture and building blocks..
- Furthermore several cross domain data sharing will be possible due to the interoperability concepts in the core so the data sharing standard and building blocks are domain agnostic and can be deployed to enable inter data space interoperability.
- Flanders uses a novel data publishing approach called Linked Data Event Streams, allowing data consumers to replicate a dataset and staying up to date with the latest changes. The aim is to end up with a stable ecosystem of

partners and contributors in different domains while reusing the underlying technical architecture and governance models. The two pilot projects that are currently running can be summarised as follows:

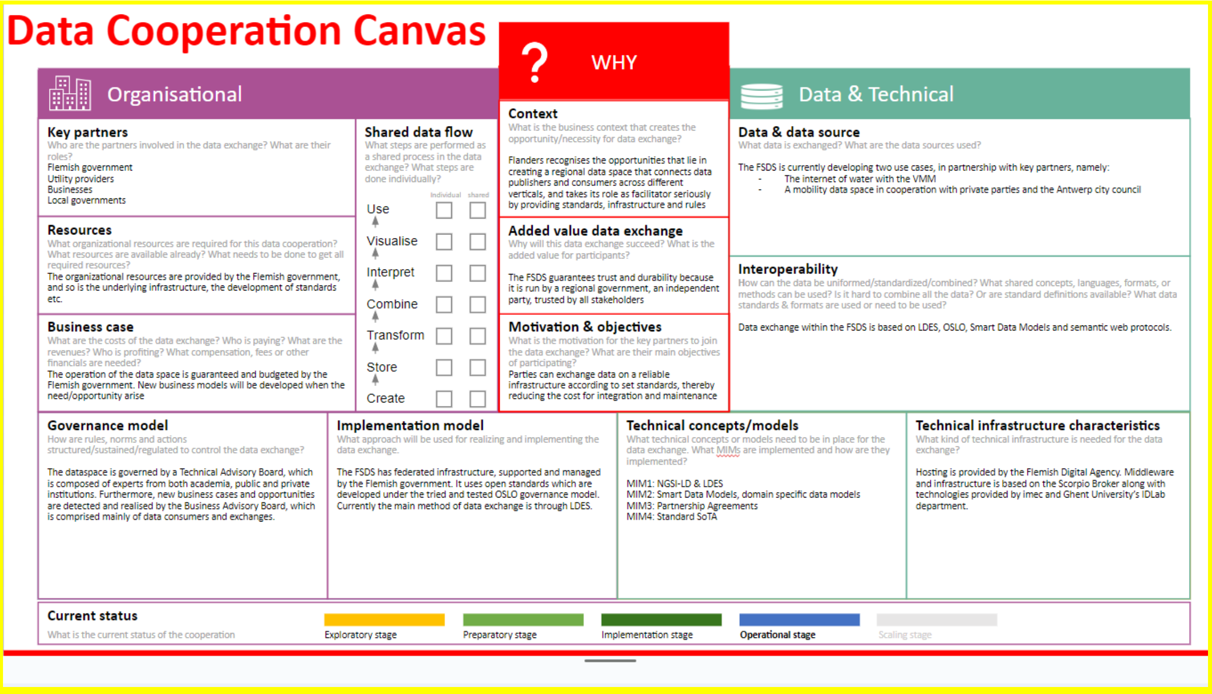
Mobility
Mobility and ITS data are very fragmented across different institutions in Belgium. In the Flemish region alone we count 5 traffic management centres, 2 departments of mobility, a handful of institutions that govern road management, vehicle registrations, public works oversight, etc. on top of that, many mobility related organisations that are vital to the proper management of mobility in Flanders are completely private. The Flanders Smart Data Space is tackling this fragmented landscape and wants to prove that collaboration in this field will strengthen all players. FSDS is facilitating the integration of various players (both public and private parties) within the mobility ecosystem, specifically emphasising traffic counts. To conduct these traffic counts, these stakeholders employ diverse methodologies, ranging from sensors to physical and manual data collection. All players will publish their data as a Linked Data Event Stream using a Flemish (semantic) data standard (only available in Dutch) about traffic counts.
To facilitate the smooth flow of this mobility data within the boundaries of Europe, the Flemish Mobility Data Space aligns with the European Mobility Data Space. Through connectors, data from Flanders will be able to seamlessly integrate with mobility data from other European countries, enabling the development of scalable applications.

Water
The Flemish government is building an “Internet of Water” together with the Flemish Environmental Agency. That project is establishing a sensor network in Flanders, where sensors at specifically chosen locations measure real-time and high-frequency water quality parameters such as temperature and conductivity. The connector developed within the framework of Digital Flanders to link existing data platforms to the Flemish Smart Data Space has been installed on the data platform of the Internet of Water platform.

The data streams available in the water data space can be used for smart applications. As a proof of concept, a predictive model was created that is always kept up-to-date with the latest measurements. The model is continuously fed with the newest data values through the provided Linked Data Event Streams (LDES). In contrast to traditional methods, where the model is calibrated using a predefined training dataset, in this model, the prediction is updated each time new data becomes available. This way, the model remains as currently. The model has the capability to incorporate new data while retaining its previous knowledge. This enables real-time analysis of water quality, allowing emerging trends and breaches of environmental standards to be detected and predicted in a timely manner.

Due to the fact that Flanders has already a data space under development, their validation to the DS4SSCC architecture has consisted of providing their lessons learned to the technical team about their process and results.

3.6.2.2 Data Cooperation Canvas



Data Cooperation Canvas

Organisational

Key partners
Who are the partners involved in the data exchange? What are their roles?
Flemish government
Utility providers
Businesses
Local governments

Resources
What organizational resources are required for this data cooperation? What resources are available already? What needs to be done to get all required resources?
The organizational resources are provided by the Flemish government, and so is the underlying infrastructure, the development of standards etc.

Business case
What are the costs of the data exchange? Who is paying? What are the revenues? Who is profiting? What compensation, fees or other financials are needed?
The operation of the data space is guaranteed and budgeted by the Flemish government. New business models will be developed when the need/opportunity arise

Shared data flow
What steps are performed as a shared process in the data exchange? What steps are done individually?

	Individual	shared
Use	□	□
Visualise	□	□
Interpret	□	□
Combine	□	□
Transform	□	□
Store	□	□
Create	□	□

WHY

Context
What is the business context that creates the opportunity/necessity for data exchange?
Flanders recognises the opportunities that lie in creating a regional data space that connects data publishers and consumers across different verticals, and takes its role as facilitator seriously by providing standards, infrastructure and rules

Added value data exchange
Why will this data exchange succeed? What is the added value for participants?
The FSDS guarantees trust and durability because it is run by a regional government, an independent party, trusted by all stakeholders

Motivation & objectives
What is the motivation for the key partners to join the data exchange? What are their main objectives of participating?
Parties can exchange data on a reliable infrastructure according to set standards, thereby reducing the cost for integration and maintenance

Data & Technical

Data & data source
What data is exchanged? What are the data sources used?
The FSDS is currently developing two use cases, in partnership with key partners, namely:
- The internet of water with the VMM
- A mobility data space in cooperation with private parties and the Antwerp city council

Interoperability
How can the data be uniformed/standardized/combined? What shared concepts, languages, formats, or methods can be used? Is it hard to combine all the data? Or are standard definitions available? What data standards & formats are used or need to be used?
Data exchange within the FSDS is based on LDES, OSLO, Smart Data Models and semantic web protocols.

Governance model
How are rules, norms and actions structured/sustained/regulated to control the data exchange?
The dataspace is governed by a Technical Advisory Board, which is composed of experts from both academia, public and private institutions. Furthermore, new business cases and opportunities are detected and realised by the Business Advisory Board, which is comprised mainly of data consumers and exchanges.

Implementation model
What approach will be used for realizing and implementing the data exchange.
The FSDS has federated infrastructure, supported and managed by the Flemish government. It uses open standards which are developed under the tried and tested OSLO governance model. Currently the main method of data exchange is through LDES.

Technical concepts/models
What technical concepts or models need to be in place for the data exchange. What MIMs are implemented and how are they implemented?
MIM1: NGSI-LD & LDES
MIM2: Smart Data Models, domain specific data models
MIM3: Partnership Agreements
MIM4: Standard SoTA

Technical infrastructure characteristics
What kind of technical infrastructure is needed for the data exchange?
Hosting is provided by the Flemish Digital Agency. Middleware and infrastructure is based on the Scorpio Broker along with technologies provided by imec and Ghent University's IDLab department.

Current status
What is the current status of the cooperation
Exploratory stage
Preparatory stage
Implementation stage
Operational stage
Scaling stage

Figure 30. Flanders Data Cooperation Canvas

3.6.2.3 Technical requirements

Operation of the Flemish data space realisation is underpinned by these key activities:

- Governance model
- Data actors ecosystem approach
- Fully in line with linked data best practices
- Supported by open standards and open source building blocks
- Data is available, discoverable and easy to consume
- Data is real-time and up-to-date
- Data exchanges are set up in a decentralised way and supported by a wide array of partners.

Some of the key technologies used in the Flanders Smart Data Space (FSDS) are Linked Data, Linked Data Event Streams (LDES), Linked Data Fragments amongst others. All technical information about the open-source building blocks developed by FSDS can be found on the FSDS Tech Docs.

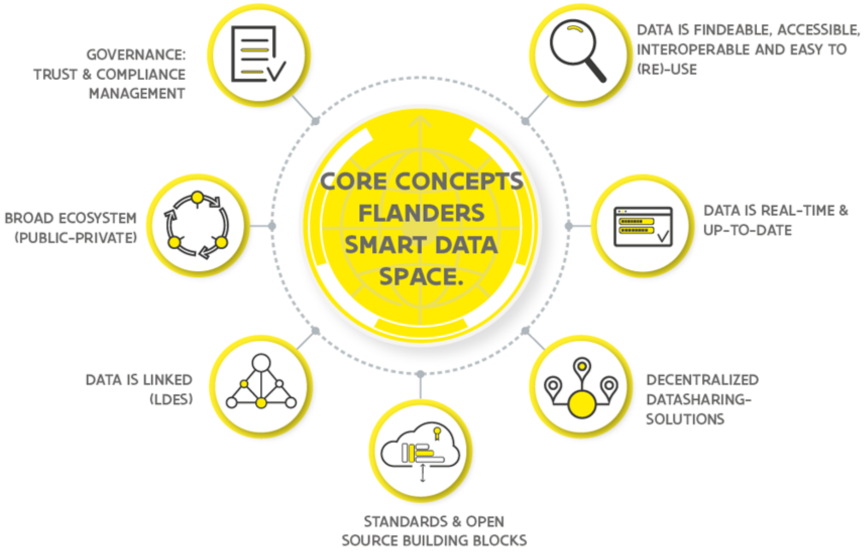

CORE CONCEPTS FLANDERS SMART DATA SPACE.
GOVERNANCE: TRUST & COMPLIANCE MANAGEMENT
DATA IS FINDEABLE, ACCESSIBLE, INTEROPERABLE AND EASY TO (RE)-USE
BROAD ECOSYSTEM (PUBLIC-PRIVATE)
DATA IS REAL-TIME & UP-TO-DATE
DATA IS LINKED (LDES)
DECENTRALIZED DATASHARING-SOLUTIONS
STANDARDS & OPEN SOURCE BUILDING BLOCKS

Figure 31: overview of the core concepts of the Flanders Smart Data Space assets and principles, including the standardised linked data approach in a decentral manner

3.6.2.4 Scenario architecture

The architecture of the Flanders Smart Data Space is generalised in 5 elements:

1. Data Standards. The Flanders Smart Data Space actively develops and maintains linked data standards building on the most prevalent vocabularies and taxonomies in different domains. It uses the governance framework developed by OSLO (Open Standards For Linked Organisations, which in turn is a regional descendant of the Interoperable Europe initiative (ISA^2 core vocabularies).
2. A data exchange standard (LDES). Event streams provide the perfect model for exchanging fast changing data and are based on the principle that only changes are logged in the most minimal way. Flanders provides building blocks to both publish (LDES Server) and consume (LDES Client) LDESes.
3. The approach is defined by different data “streets” that enable different parties to connect on a single data publication-subscription bus. These data streets consist of various building blocks available through Apache Nifi or Linked Data Interactions, a data pipeline framework developed within FSDS.
4. The FSDS is inherently dependent on its ecosystem approach, and actively facilitates interconnection between participating partners.

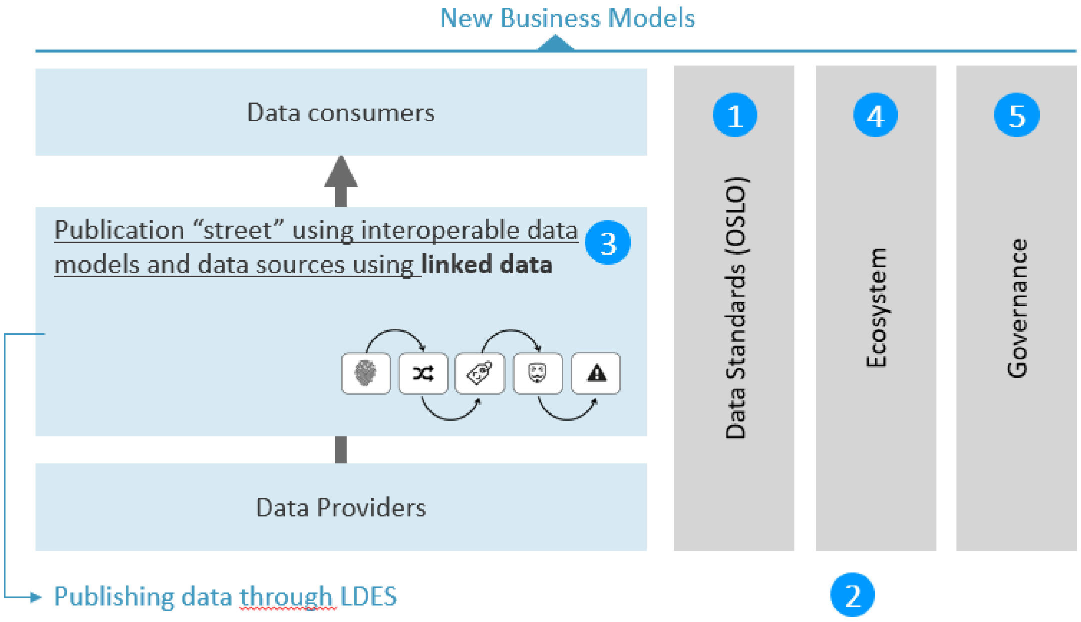

New Business Models
Data consumers
Publication "street" using interoperable data models and data sources using linked data
3
Data Providers
Publishing data through LDES
1
Data Standards (OSLO)
4
Ecosystem
5
Governance
2

Figure 32. Existing high level architecture of Flanders Smart Data Space decentral solution

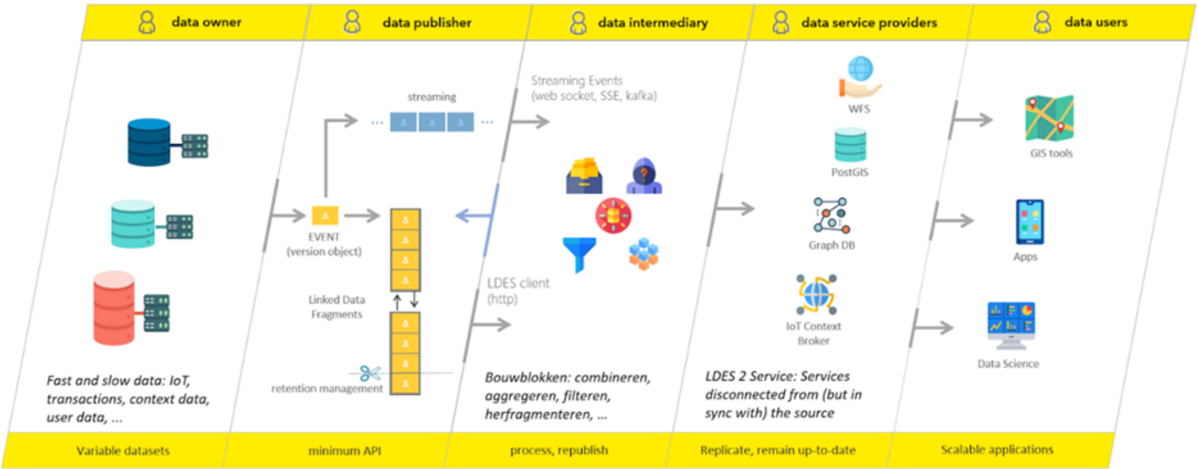


Figure 33. Decentral Architecture of the Flanders Smart Data Space building block for the enablement of the different roles in the ecosystem

As can be appreciated, the proposed architecture by Flanders Smart Data Space covers many of the aspects recommended in the DS4SSCC architecture in one or another way. It should be expanded with the registration of the data space users credentials in the Universal Trusted Data Registry, and incorporated into their Identity Management module, the Authorization Policies Store component. The Federation Layer should be only included if there would be a need to interconnect with other data spaces.

3.6.2.5 Implementation steps

Data Models

To achieve data interoperability and standardisation across different domains and applications, the Flanders Smart Data Space uses the technical specification Linked Data Event Streams and the semantic standard “Open Standards for Linked Organizations (OSLO).

Programming API

The Flanders Smart Data Space uses programming API interoperability in different areas such as:

- Context Data Management: Flanders Smart Data Space adopts the LDES (Linked Data Event Stream) and the Open Standards for Linked Organizations (OSLO) as core components for context data management.
- Data Publication and Discovery: the datasets in the different data spaces are published decentral and subsequently findable through their DCAT description. LDES has been aligned with the DCAT specification. The LDES Server exposes a DCAT endpoint with the LDES metadata (described using DCAT) which can be harvested by Open Data portals.

Identity Management

LDES Server provides a security option through an API gateway, which protects LDES Collections and Views from unauthorised access, recognizing the importance of data security. The API gateway serves as a security layer, managing access and applying authentication methods, such as ACM/IDM, reducing the chance of exposing sensitive data to unwanted parties. ACM/IDM stands for Access Control Management/Identity Management, a system that verifies the identity and permissions of users and devices. This improved security feature increases the trust and dependability of LDES Server for organisations working in security-sensitive environments.

3.6.3 Customised Architecture 3 (Amsterdam-IDEA)

3.6.3.1 Scenario description

Road authorities (local and national) have open data on road works. This data about the planned road works may differ from the actual road works due to f.e. subcontractors. Service providers and road authorities want to have data on actual road works.By validating the planned road works, using live data (from floating car data (FCD)), IDEA generates an high quality, real-time data feed for road works.Providing high quality, real time data on road works. Service providers can provide better information to road users. Road authorities have insight into their roadworks' actual impact. For example to check on subcontractors.

Key partners

- NDW (National Data Access Point on road traffic)
- City of Amsterdam, Traffic Department
- City of The Hague, Traffic Department
- Province of North Holland
- RWS (National Road Authority)

Resources

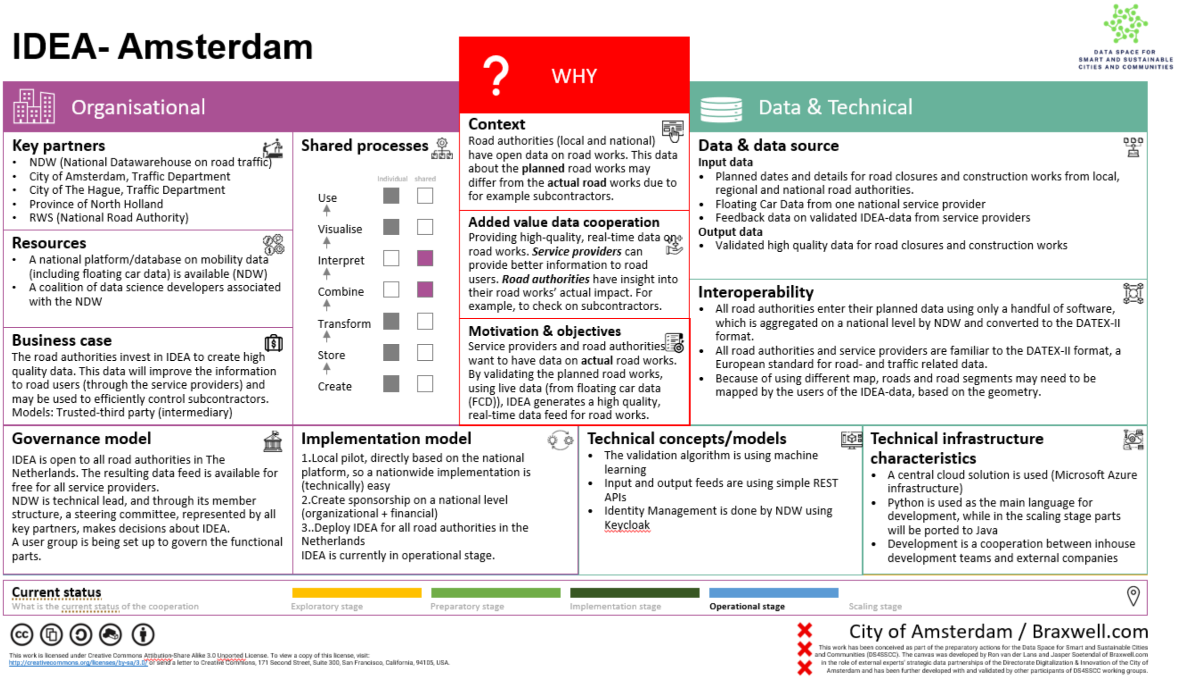
- A national platform/database on mobility data (including floating car data) is available (NDW)
- A coalition of data science developers associated with the NDW

Business case

The road authorities invest in IDEA to create high quality data. This data will improve the information to road users (through the service providers) and may be used to efficiently control subcontractors.

Models: Trusted-third party intermediary

3.6.3.2 Data Cooperation Canvas



IDEA- Amsterdam

? WHY

Organisational

Data & Technical

Key partners
- NDW (National Datawarehouse on road traffic)
- City of Amsterdam, Traffic Department
- City of The Hague, Traffic Department
- Province of North Holland
- RWS (National Road Authority)

Resources
- A national platform/database on mobility data (including floating car data) is available (NDW)
- A coalition of data science developers associated with the NDW

Business case
The road authorities invest in IDEA to create high quality data. This data will improve the information to road users (through the service providers) and may be used to efficiently control subcontractors. Models: Trusted-third party (intermediary)

Shared processes

	Individual	shared
Use	■	□
Visualise	■	□
Interpret	□	■
Combine	□	■
Transform	■	□
Store	■	□
Create	■	□

Context
Road authorities (local and national) have open data on road works. This data about the **planned** road works may differ from the **actual road** works due to for example subcontractors.

Added value data cooperation
Providing high-quality, real-time data on road works. ***Service providers*** can provide better information to road users. ***Road authorities*** have insight into their road works' actual impact. For example, to check on subcontractors.

Motivation & objectives
Service providers and road authorities want to have data on **actual** road works. By validating the planned road works, using live data (from floating car data (FCD)), IDEA generates a high quality, real-time data feed for road works.

Data & data source
Input data
- Planned dates and details for road closures and construction works from local, regional and national road authorities.
- Floating Car Data from one national service provider
- Feedback data on validated IDEA-data from service providers

Output data
- Validated high quality data for road closures and construction works

Interoperability
- All road authorities enter their planned data using only a handful of software, which is aggregated on a national level by NDW and converted to the DATEX-II format.
- All road authorities and service providers are familiar to the DATEX-II format, a European standard for road- and traffic related data.
- Because of using different map, roads and road segments may need to be mapped by the users of the IDEA-data, based on the geometry.

Governance model
IDEA is open to all road authorities in The Netherlands. The resulting data feed is available for free for all service providers.
NDW is technical lead, and through its member structure, a steering committee, represented by all key partners, makes decisions about IDEA.
A user group is being set up to govern the functional parts.

Implementation model
1.Local pilot, directly based on the national platform, so a nationwide implementation is (technically) easy
2.Create sponsorship on a national level (organizational + financial)
3..Deploy IDEA for all road authorities in the Netherlands
IDEA is currently in operational stage.

Technical concepts/models
- The validation algorithm is using machine learning
- Input and output feeds are using simple REST APIs
- Identity Management is done by NDW using Keycloak

Technical infrastructure characteristics
- A central cloud solution is used (Microsoft Azure infrastructure)
- Python is used as the main language for development, while in the scaling stage parts will be ported to Java
- Development is a cooperation between inhouse development teams and external companies

Current status
What is the current status of the cooperation

Exploratory stage | Preparatory stage | Implementation stage | **Operational stage** | Scaling stage

City of Amsterdam / Braxwell.com

This work has been conceived as part of the preparatory actions for the Data Space for Smart and Sustainable Cities and Communities (DS4SSCC). The canvas was developed by Ron van der Lans and Jasper Soetendal of Braxwell.com in the role of external experts' strategic data partnerships of the Directorate Digitalization & Innovation of the City of Amsterdam and has been further developed with and validated by other participants of DS4SSCC working groups.

Figure 34. Amsterdam Data Cooperation Canvas

3.6.3.3 Technical requirements

IDEA requires following input to generate validated high quality data for road closures and construction works:

- Planned dates and details for road closures and construction works from local, regional and national road authorities.
- Floating Car Data from one national service provider
- Feedback data on validated IDEA-data from service providers

An external party (consortium of traffic, data science and software partners) develops the application as a central model sourced at NDW with connections by REST API. As technical infrastructure Microsoft Azure is used.

All road authorities enter their planned data using only a handful of software, which is aggregated on a national level by NDW and converted to the DATEX-II format. All road authorities and service providers are familiar with the DATEX-II format, an European standard for road- and traffic related data. Because of using different maps, roads and road segments may need to be mapped by the users of the IDEA-data, based on the geometry.

3.6.3.4 Scenario architecture

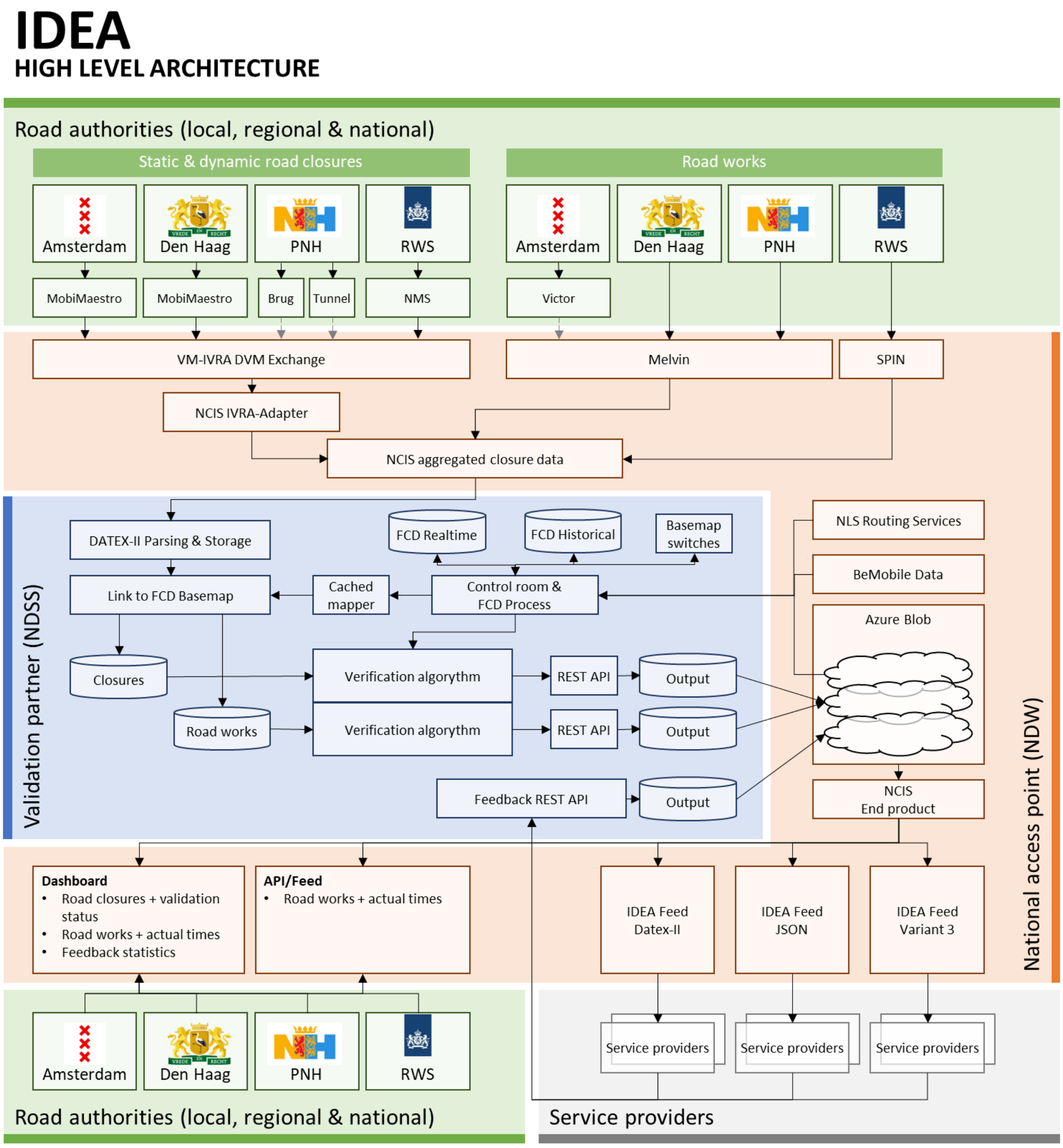


Figure 35. Existing high level architecture of Amsterdam city solution

Below presented diagram outlines the customised high level architecture for IDEA. The project is based on the National Data Warehouse for Traffic Information (NDW) data platform (orange), a database of both real-time and historic traffic data in the Netherlands. Official data on road construction (RWS), tunnel and bridge closures (PNH) is provided for the Cities of Amsterdam and Den Haag (MobiMaestro) (green).

DATEX2 standards are used to connect input data to developed data scientific machine learning (ML) algorithmic procedures for verification purposes by including floating car data (BeMobile) and data from NLS Routing Services (orange). Verified, hence trusted, results are accessed via Rest API to navigation providers, such as

Google or TomTom.

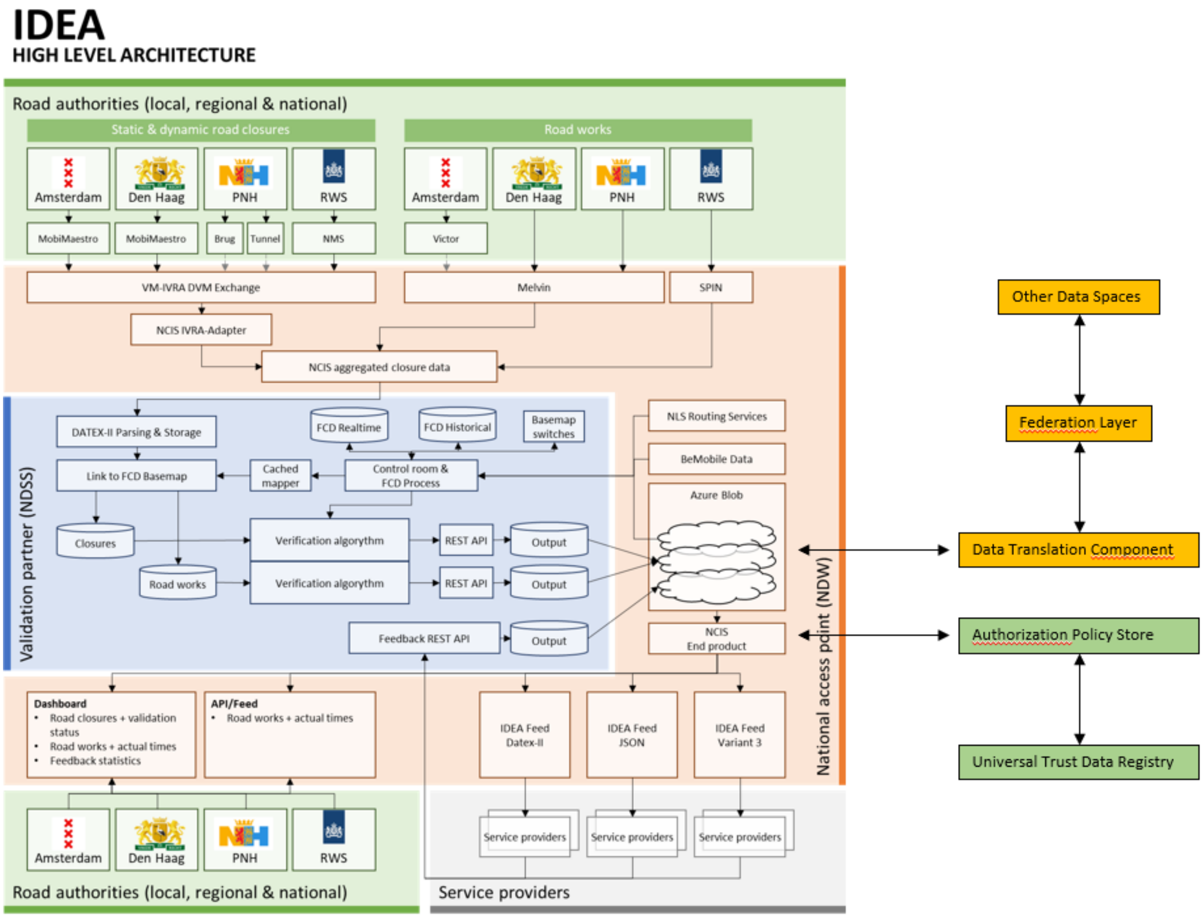


Figure 36. Customised high level architecture of Amsterdam data space

To make the existing architecture of the IDEA solution compliant with the Architecture Model, multiple components need to be integrated. The model will be extended with an integration with Universal Trust Data Registry Component and a configuration of Authorization Policy Store. In order to support interoperable data exchange between data space ecosystem solution and other dataspace participants, a custom Data Translation Component had to be introduced.

3.6.3.5 Implementation steps

Exploratory & preparatory stages:

- IDEA is initiated and initially financed by the City of Amsterdam.
- IDEA is developed in cooperation with NDW, the national shared service center for data from national, regional and local road authorities.
- By working together with the NDW, a local solution is built on top of the national framework, using only existing data sources. By doing this, from a technical viewpoint, nationwide scaling up to other road authorities would be very easy.

- A pilot was started with 2 local, a regional and a national road authority.

Operational stage:

- IDEA is now ready to be implemented for all road authorities in The Netherlands, using only existing systems and data sources.

3.6.4 Customised Architecture 4 (Valencia)

3.6.4.1 Scenario description

Valencia city has been a recognized smart city since many years ago, being referred as one of the first three cities in the world measuring the fulfilment of the SDGs. For Valencia city, being smart means also being sustainable. Due to this commitment with sustainability and the technology, a dedicated Smart City Office was created in 2018 to support the development of the smart city strategy and technical platform.

Looking at the history of Smart City concept in Valencia, already in 2014 the city set up the first platform (called VLCi platform) and the Open Data portal. In 2015, the city included a KPI dashboard and integrated the mobility service. In 2017, lighting and parking services were additionally integrated. In 2018, the gardening service was added. And since then, advanced features have been added to the platform. The figure below shows the evolution of Valencia smart city plan towards the Agenda 2030.

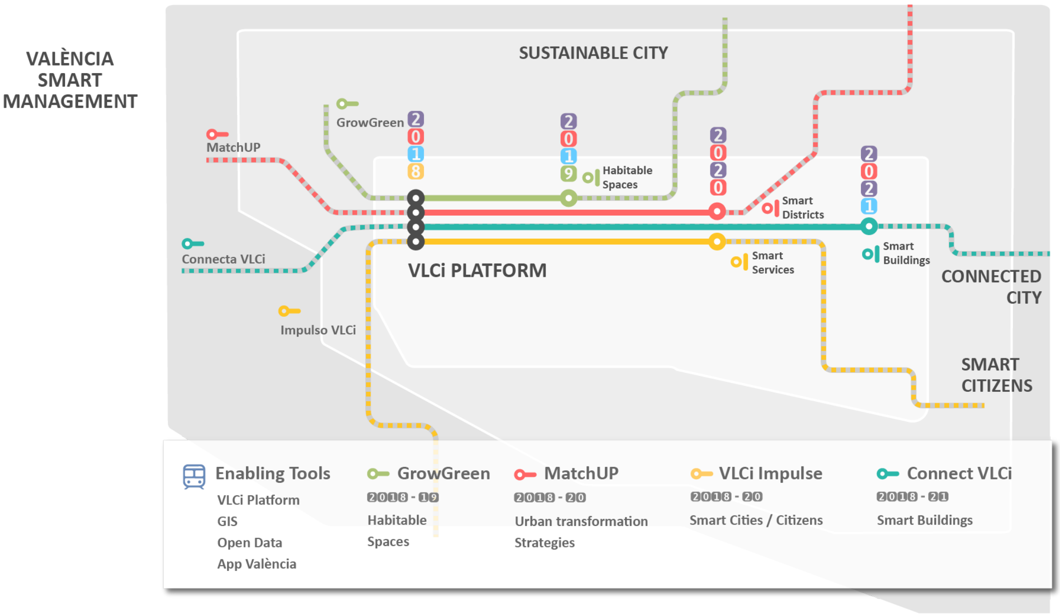


Figure 37. Valencia strategic roadmap for city platform

In the current scenario of Valencia, they have some vertical services from third-party providers integrated in the city platform. This integration consists of getting the data from the sensors in the city that these services are gathering (parking, lights, gardens…). Normally there is a contract between the city and the provider who has won a public tender to offer the service. The city obtains the collected data and via API is able to send some commands to the third-party apps over the sensors.

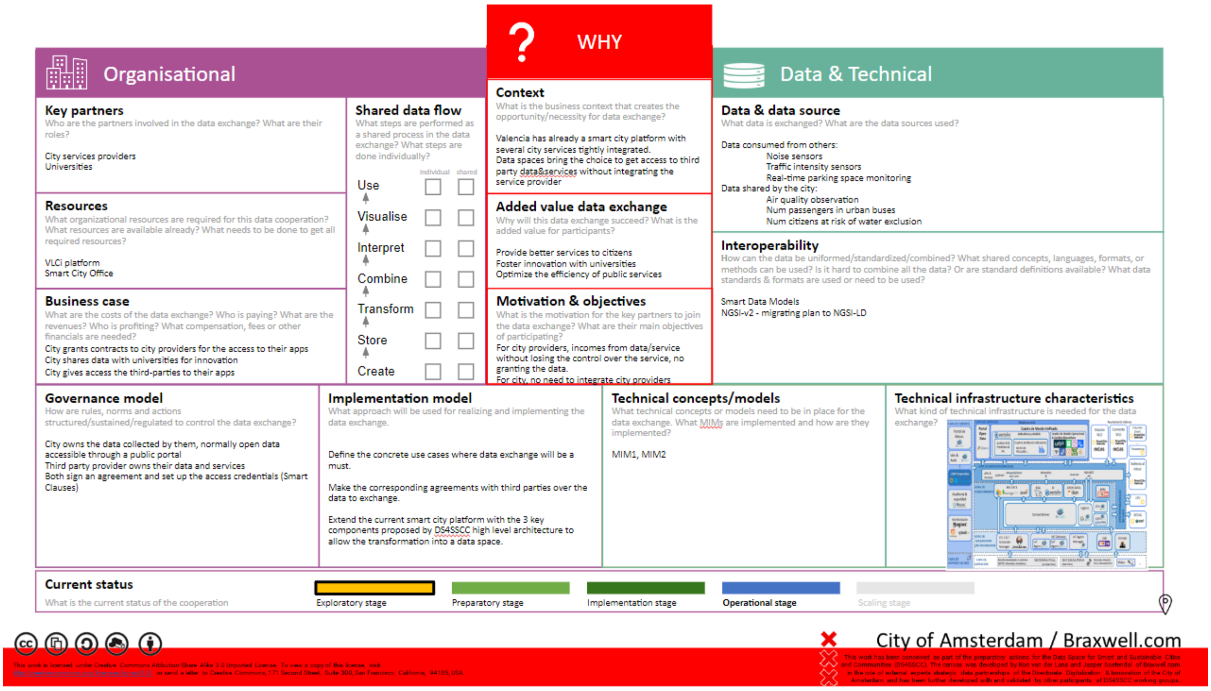
By expanding the current city platform to the data spaces approach, several scenarios are feasible for Valencia:

- **Get access to third-party applications** with the same identity credentials as in the smart city platform. Thus, city employees and citizens (users of the smart city platform) could use these applications to visualise, interact or actuate (depending on your access rights). They could use the allowed functionality in the application for them.
- **Provide access to their city applications** without the need to register or integrate new users/participants. The external users would have their universal credentials and they just need to define the type of access. If wished, the city could monetize these accesses and charge costs for the access to the data (in case the data must not be open data).
- **Get access to federated data**. They could access data from other cities or data spaces which are not directly in their data space through the federation services provided by their data space. This scenario could be useful to carry out benchmarking or comparison analysis with other cities, i.e. number of trees per city.

In summary, the integration of data, applications and users is simplified significantly, making the management of the platform more efficient and saving time and costs in integration processes.

3.6.4.2 Data Cooperation Canvas

As explained in section 1.3, an instrument referred as Data Cooperation Canvas has been created by the DS4SSCC to facilitate the cities and communities to consolidate in one page chart all the aspects of their data space. In the case of Valencia, the result is presented below. At the current status of data space development, they are in the Exploratory phase.



? WHY

Organisational

Key partners
Who are the partners involved in the data exchange? What are their roles?

City services providers
Universities

Resources
What organizational resources are required for this data cooperation? What resources are available already? What needs to be done to get all required resources?

VLCi platform
Smart City Office

Business case
What are the costs of the data exchange? Who is paying? What are the revenues? Who is profiting? What compensation, fees or other financials are needed?
City grants contracts to city providers for the access to their apps
City shares data with universities for innovation
City gives access the third-parties to their apps

Shared data flow
What steps are performed as a shared process in the data exchange? What steps are done individually?

	individual	shared
Use	□	□
Visualise	□	□
Interpret	□	□
Combine	□	□
Transform	□	□
Store	□	□
Create	□	□

Context
What is the business context that creates the opportunity/necessity for data exchange?

Valencia has already a smart city platform with several city services tightly integrated.
Data spaces bring the choice to get access to third party data&services without integrating the service provider

Added value data exchange
Why will this data exchange succeed? What is the added value for participants?

Provide better services to citizens
Foster innovation with universities
Optimize the efficiency of public services

Motivation & objectives
What is the motivation for the key partners to join the data exchange? What are their main objectives of participating?
For city providers, incomes from data/service without losing the control over the service, no granting the data.
For city, no need to integrate city providers

Data & Technical

Data & data source
What data is exchanged? What are the data sources used?

Data consumed from others:
Noise sensors
Traffic intensity sensors
Real-time parking space monitoring
Data shared by the city:
Air quality observation
Num passengers in urban buses
Num citizens at risk of water exclusion

Interoperability
How can the data be uniformed/standardized/combined? What shared concepts, languages, formats, or methods can be used? Is it hard to combine all the data? Or are standard definitions available? What data standards & formats are used or need to be used?

Smart Data Models
NGSI-v2 - migrating plan to NGSI-LD

Governance model
How are rules, norms and actions structured/sustained/regulated to control the data exchange?

City owns the data collected by them, normally open data accessible through a public portal
Third party provider owns their data and services
Both sign an agreement and set up the access credentials (Smart Clauses)

Implementation model
What approach will be used for realizing and implementing the data exchange.

Define the concrete use cases where data exchange will be a must.

Make the corresponding agreements with third parties over the data to exchange.

Extend the current smart city platform with the 3 key components proposed by DS4SSCC high level architecture to allow the transformation into a data space.

Technical concepts/models
What technical concepts or models need to be in place for the data exchange. What MIMs are implemented and how are they implemented?

MIM1, MIM2

Technical infrastructure characteristics
What kind of technical infrastructure is needed for the data exchange?

Current status
What is the current status of the cooperation

🟨	🟩	🟩	🟦	⬜
Exploratory stage	Preparatory stage	Implementation stage	Operational stage	Scaling stage

City of Amsterdam / Braxwell.com

This work has been conceived as part of the preparatory actions for the Data Space for Smart and Sustainable Cities and Communities (DS4SSCC). The canvas was developed by Ron van der Lans and Jasper Soetendal of Braxwell.com in the role of external experts strategic data partnerships of the Directorate Digitalization & Innovation of the City of Amsterdam and has been further developed with and validated by other participants of DS4SSCC working groups.

Figure 38. Valencia Data Cooperation Canvas

3.6.4.3 Technical requirements

Valencia smart city platform users right now are:

- the citizens with open access to the public part,
- the civil workers registered in a LDAP[14] (Lightweight Directory Access Protocol) and accessing by user/passwd
- some universities accessing to the technical infrastructure (endpoint to the Context Broker, the component in the platform for data exchange)

The most relevant data exchanged by the city is:

- Consumed by the city:
  - noise sensors, owned by the city and public in the open data portal
  - traffic intensity sensors, owned by the city and public in the open data portal
  - real-time parking space monitoring, owned by the city and public in the municipality portal
- Provided by the city:
  - air quality observations, owned by the city and public in the municipality web page
  - number of passengers in urban buses, owned by the city public transport company, private data available through the platform dashboard

[14] https://en.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol

- number of administrative records for citizens at risk of water exclusion, owned by the city and public in the open data portal

They are using Smart Data Models (NoiseLevelObserved, TrafficFlowObserved, ParkingSpot, AirQualityObserved, KeyPerformanceIndicator) and NGSI-v2 as standards for interoperability; API key for identification and authorization in some cases and user/passwd and LDAP in other cases.

3.6.4.4 Scenario architecture

The figure below shows the current architecture of Valencia smart city platform. It represents a typical layered smart city architecture, followed by many cities across the world according to the standard ISO/IEC 240939:2022[15]. As can be seen in the picture, it relies on some FIWARE components which are already compliant with some of the recommended building blocks and standards included in the DS4SSCC Catalogue of Specifications (NGSI, CKAN, Smart Data Models…). Several vertical services (Impulso, Conecta, Valencia al minut,...) are integrated under the platform, providing a perfect *brownfield* scenario to validate the DS4SSCC architecture.

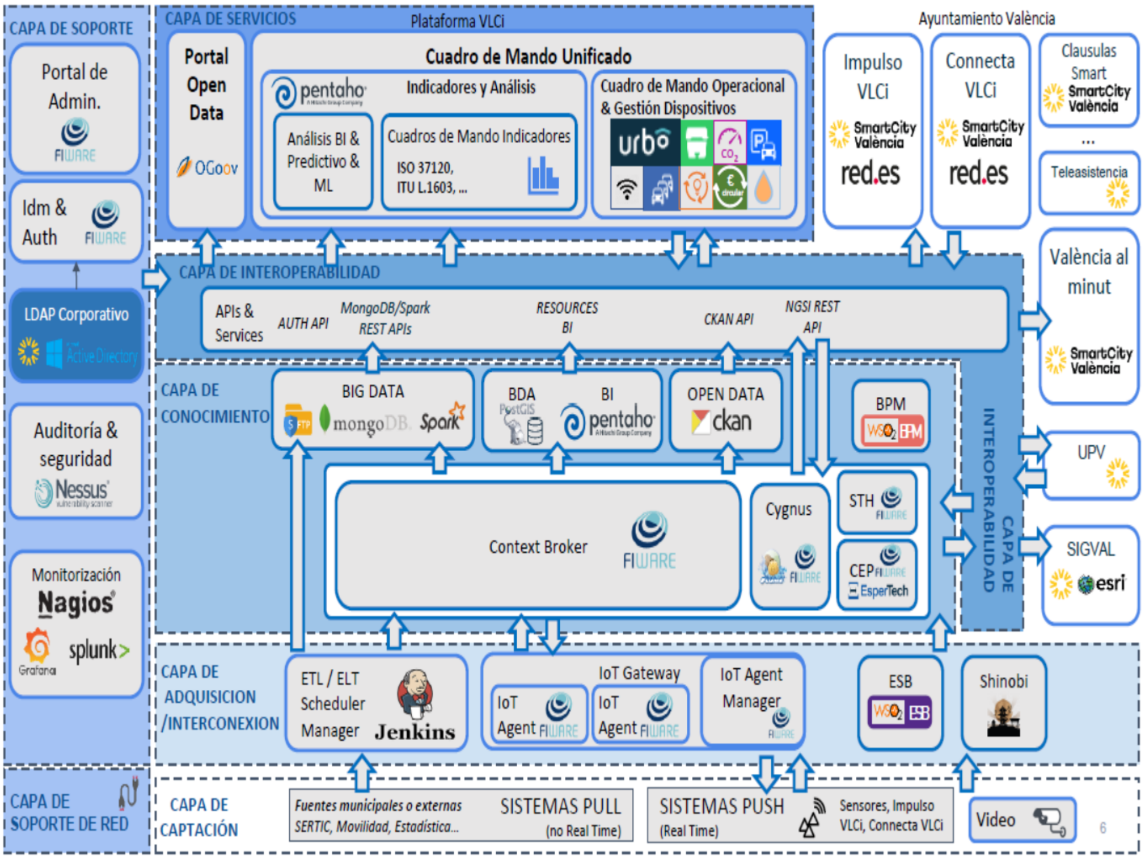


Figure 39. Existing high level architecture of Valencia city solution

[15] https://www.iso.org/standard/77621.html

The above presented architecture has been extended with the proposed components in the high level architecture in section 3.4, and the figure below depicts the resulting high level architecture. On one side the Identity Management component of the current architecture has to be connected to an Authorization Policies Store provided by the data space to which the city wants to engage, in order to be able to get access to the Universal Trust Data Registry. For that, it is also necessary that the current IAM register the Authorization policies previously in the Universal Registry. This mechanism will ensure that every data space participant is uniquely identified to get access to the right data and services.

On the other hand, the city platform needs to use the Federation Layer component if it wants to federate their Publication and Discovery services into the data space to get access to data from other participants in the data space or from other data spaces. This data space can be at the local, regional, national or European level, depending on the desired ecosystem and strategy of the city.

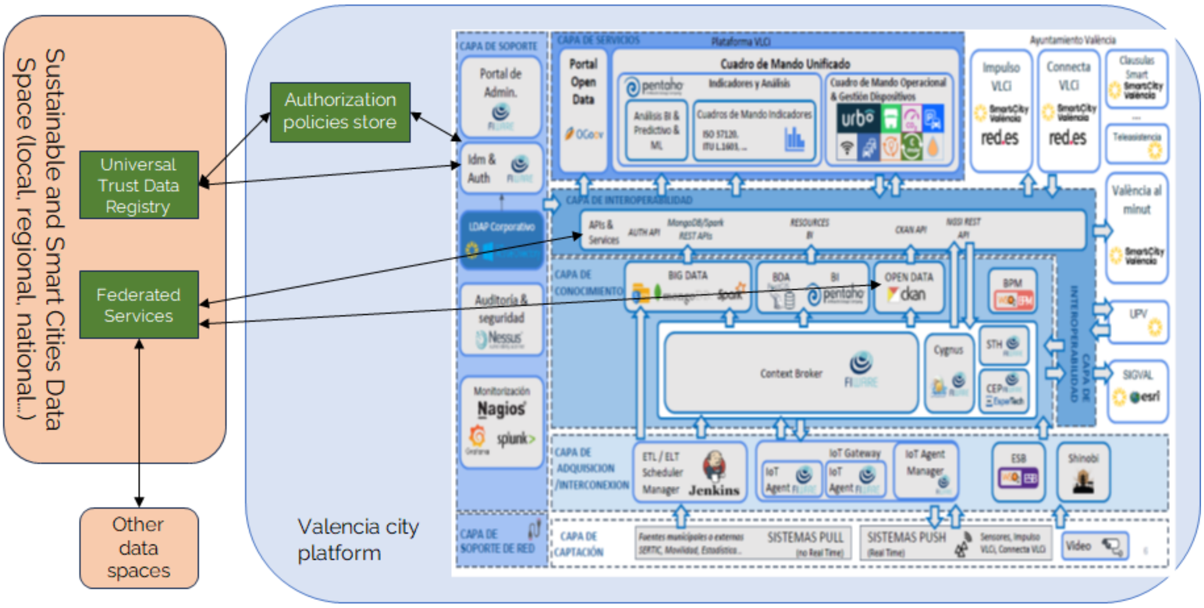


Figure 40. Customized high level architecture of Valencia data space

3.6.4.5 Implementation steps

To carry out this evolution from the existing architecture to the engagement in a data space with the identified ecosystem, following concrete implementation steps are recommended:

- Evolving the current architecture to the most updated versions of technological components to ensure the alignment with most adopted standards.
- Mapping the existing components in the architecture to the building blocks recommended in the Catalogue of Specifications (in alignment with the DSSC

building blocks). Identify which one is missing and evaluate if it is required. It is also recommended to check in the Catalogue the use of recommended standards and reference implementations by DS4SSCC.

- Deploying the required building blocks, either if they already exist and need to be updated; or if they have been identified now.
- Installing the DS4SSCC high level architecture components required to evolve the current architecture towards the data space concept (in green in the picture). If the components are not yet available, it is recommended to follow the latest developments of open source communities under the scope of data spaces and find collaboration with other cities and technological providers to agree on their development.
- Testing the proposed architecture with a simple use case and incrementally enlarge the experiment with more complicated scenarios. The Citcom.ai project, where Valencia is participating, may for sure support the city in this process.

3.7 Required alignments with DSSC

During the architecture definition process, the DSSC has released an updated taxonomy of the Building Blocks. This section refers to all the required changes in the Catalogue online to keep aligned the DS4SSCC blueprint with the DSSC one.

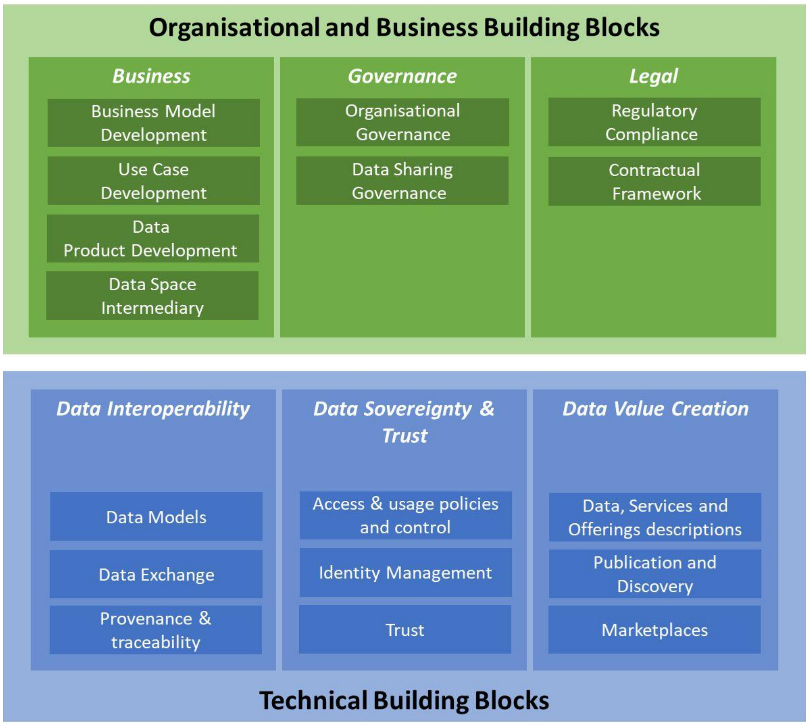


Figure 41. DSSC Building Blocks Taxonomy

Additionally, the DSSC has recently endorsed the DSBA Technical Convergence document to which DS4SSCC is already aligned, so no need to further update due to this fact.

The changes made in the online Catalogue have been the following:

- The Trusted Exchange BB has been renamed to Trust
- The Data Models & Formats has been renamed to Data Models
- The Data Exchange API BB has been renamed to Data Exchange
- The Access and Usage Control BB has been renamed to Access & Usage policies and control
- Adjustments of the Data Value Creation category which imply:
  - Join Publication and Discovery in the same BB, thus Metadata and Discovery Services has been renamed to Publication and Discovery;
  - Publication and Marketplaces Service has been renamed to Marketplaces
  - Move the standards and reference implementations from Data usage accounting to the renamed Marketplaces BB.
- The Organisational and Business part has been split into Governance, Business and Legal.

We have also updated the descriptions of all the BBs to the proposed descriptions by the DSSC in the taxonomy paper shared with the data spaces.

4 CookBook, recommendations and tips for use

Despite the relevance of the technical work delivered both in D3.1 with the Catalogue of Specifications and here in the present document about the Reference Architecture, cities and communities require some guidelines to overcome the challenge to deploy or engage in a data space.

Therefore, it is crucial to provide what is called in DS4SSCC the CookBook which accompanies the Catalogue and the Architecture Model. We have split the CookBook into a short introductory guide, the recipes for each of the 3 scenarios identified in the document and a Frequently Asked Questions section that provides a set of tips for using the blueprint.

Additionally, we have included the collected feedback through several iterations from use cases and diverse stakeholders about the architecture.

4.1 Short guidance

This guide provides a brief summary of the high level steps which are required to set up and deploy a data space. The fine grained process for each step can be found in this and other deliverables of the project.

Would your city or community be interested in getting engaged in a data space? We recommend the following steps and pointers to specific material.

1) *Why does my city/community want to engage in a data space?*

We recommend exploring your motivation, ecosystem, existing data and technologies by using the Data Cooperation Canvas tool proposed by DS4SSCC. Through this tool you will consolidate in one unique diagram the most relevant points and information to make the right decision about your data space. Refer to section 3.1 in D2.2 Multi Stakeholder Governance Scheme.

2) *Which is the technical maturity of my city/community to be engaged in a data space?*

In order to fill the technical dimension of your Data Cooperation Canvas, you would also have to collect the different data sources in your city/community, identify the level of digitalization of your data process (if you have a data platform, or some digitalized services, or no digitalization at all) and the existing technical infrastructures. In many cases the engagement to a data space will be an evolution from your current scenario (brownfield, greenfield, digital twin). The technical analysis should also include the collection of standards and reference implementations you are using currently in your technical infrastructure. The DS4SSCC Catalogue of Specifications collects all the standards, specifications and reference implementations that are currently being used by more than 80 cities interviewed during the project. You may have a look at the catalogue, identify which ones you are using and provide inputs about existing or new standards/implementations. Refer to online Catalogue of Specifications.

This analysis will give you the starting point for your technical evolution towards a data space.

3) *How can my city/community organise the surrounding ecosystem with the best governance and fully compliant with EU regulations?*

We recommend looking at the analysed use cases and identity which one is closer to your current scenario. You can then follow the provided recommendations at local, regional, national and European level to set up the most suitable governance for you and follow the suggested rules for establishing the Code of Conduct. Refer to various sections in D2.2 Multi Stakeholder Governance Scheme.

4) *Which are the architectural and technical evolutions that my city/community should do in order to engage in a data space?*

DS4SSCC proposes a high level reference architecture assuming that the city/community has already a certain level of digitalization, this means, at least a data platform where the data sources are integrated or some digitalized vertical services (parking, gardening, cadastre, water,...). If the city/community is not yet at this stage, a first step would be to set up the digital capture and process of the data in order to be able to share this data with others.

Once this pre-requisite has been fulfilled, the mandatory action is to connect to an Universal Trust Data Registry where universal credentials of data space users are stored. The identity and access management system of the city/community data platform must connect to this registry through an Authorization Policies Store, which includes the access rights allowed to the data and services in the platform for each user. Optionally, the city/community platform can be also connected to the Federation Layer to federate the publication and discovery services of the platform with the ones in the data space, as well as the marketplace, if it exists. Explanation about the evolution of the architecture, description of the mentioned components and four examples of application can be found in this document. Refer to D3.2 Reference Architecture Model, especially section 3.

5) *Which are the stages of a data space and what should my city/community do in each of the stages?*

Following the recommendations from DSSC, 5 stages have been defined (exploratory, preparatory, implementation, operational, scaling). For each of the stages, the city/community needs to address different steps for governance, architecture and datasets that have been defined in a roadmap document. Refer to the action plan included in D4.2 Roadmap for implementing a European data space for smart and sustainable cities and communities.

4.2 Recipes for the different scenarios

This section recommends concrete steps for each of the three possible scenarios identified in this document. They have been consolidated from the experience in the customization of the four use cases conducted in section 3.6, and generalised for common use.

Data Space from scratch (greenfield)

This scenario represents all the cities and communities which have not yet digitised their data collection and processing or are at a very early stage of that, just having few data sources digitally treated. In this case, the cities need to develop their data space from scratch and require some certain steps to be accomplished in advance:

- Define a strategic plan for the digitalization of the city. This plan must include the analysis of the current infrastructures, the desired digital services to be developed, the required data sources to be used by the digital services, the required infrastructure for implementing the services, including sensors, devices and data platform, and a clear agenda and funds to carry out all the actions.
- Launch the corresponding tenders and elaborate the due contracts with providers. It is recommended to include in the public tenders documents, the requirements in line with the data spaces approach. In this way, the implemented infrastructure is already prepared and compatible with the standards and technologies that will be required later to engage in a data space.
- Another possibility could be hiring (internal) experts, consultants or reuse of external developments (by other cities for example). Or some localities might use support from a governmental digitalisation entity (as is the case in Flanders with Digital Flanders) or collaboration between cities.
- Once all of this is in place, steps in the following case can be followed.

Existing Data Platform (brownfield)

In this case, the city or community already has a data platform in place. This means that they have all data sources integrated in the data platform, in some cases interoperable amongst them, and city services using and exploiting such datasets. This data platform would have an architecture and an implementation following certain standards and technologies, so an evolution from the existing platform is required. Following steps are recommended:

- Evolving the current architecture to the most updated versions of technological components to ensure the alignment with most adopted standards.
- Mapping the existing components in the architecture to the building blocks recommended in the Catalogue of Specifications (in alignment with the DSSC

building blocks). Identify which one is missing and evaluate if it is required. It is also recommended to check in the Catalogue the use of recommended standards and reference implementations by DS4SSCC.

- Deploying the required building blocks, either if they already exist and need to be updated; or if they have been identified now.
- Installing the DS4SSCC high level architecture components required to evolve the current architecture towards the data space concept (in green in the high level architecture picture). If the components are not yet available, it is recommended to follow the latest developments of open source communities under the scope of data spaces and find collaboration with other cities and technological providers to agree on their development.
- Testing the proposed architecture with a simple use case and incrementally enlarge the experiment with more complicated scenarios.

Existing Digital Twin (advanced)

This case is an advanced scenario from the previous one, where the existing data platform is a virtual representation of the city or community by using the context of the data. Thus, the steps than previously mentioned are valid here. Other additional steps to advance the data space further as to enable new applications of the digital twins maybe:

- Determine a long-term digital vision.
- Develop a business case with the stakeholders.
- Roadmap including the data space enrichment.

4.3 FAQ - Tips for use

This section includes all possible questions that a concrete use case may have at the time to deploy a data space instance

- **What is a data space?**

According to the DSSC Glossary, a data space is "*a distributed system defined by a governance framework that enables trustworthy data transactions between participants while supporting trust and data sovereignty*". A data space then consists of several actors who want to share and exchange data as an asset, and they need all the instruments to do that in a trusted and smooth manner. A data space needs to facilitate the agreements amongst participants to carry out the data transaction, which not necessarily has to be monetized. All the participants need to gain some value out of the data exchange.

- **Which is the value of a data space for a city or a community?**

The data spaces are the natural step for cities and communities with a certain degree of digitalization which have already integrated the data sources from the city in a data platform or digital services. It can also be an ambition for those cities still in digitalization progress but looking at the future.

With the evolution from data platforms to data spaces, the cities and communities will facilitate their access to data from third parties and vice-versa. This will happen without a significant investment in the integration of these third parties (city providers, e.g.) and without any implementation and any additional cost. Additionally, the cities could also monetizate their data which is not openly available in their open data portals, allowing to reinvest these means in further enriching their digital offering.

Ultimately, thanks to these benefits, they could provide more efficient and sustainable services to the citizens and could engage them as data space participants, either as data providers and/or data consumers.

- **Why should I engage in a data space?**

Before embarking on this endeavour, be sure about your readiness and motivation for that. We recommend following the steps indicated in section 4.1.

- **What is a data space blueprint?**

According to the DSSC Glossary, a data space blueprint is "*a consistent, coherent and comprehensive set of guidelines to support the implementation, deployment and maintenance of data spaces*". Every data space may define different elements to include in the blueprint, but usually it should include a glossary, a conceptual model, a reference architecture, an inventory of components, standards and technologies and all the guidelines to use this material.

- **What does the DS4SSCC Technical Blueprint include?**

The DS4SSCC technical blueprint includes a Reference Architecture Model for smart cities and communities, a Catalogue of Specifications with the standards and reference implementations for each building block and a CookBook with all the recipes for the emerging data spaces to be deployed. This blueprint is complemented with the Governance, Use Cases and Roadmap delivered in other documents of the project.

- **How should I use the Catalogue of Specifications?**

The Catalogue is available online for consultation and evolution. Every city or community may look at the Catalogue to find which are most adopted standards and reference implementations used in other cities and communities. In this way, it can decide to follow them for the sake of interoperability or provide extensions to this catalogue by filling the online form. A documented version of the Catalogue is also available for reading all the details about the sources and process to build the Catalogue.

- **How can I identify which scenario is my city/community?**

Refer to section 3.1 where the scenarios are described and find the closest. The recipes for each specific scenario can be found in section 4.2 of this document.

- **How should I apply the proposed high level architecture to my concrete scenario?**

In the case of the brownfield or digital twin scenarios, the basic concept is to use the three recommended components in the high level architecture (coloured in green). Examples about how to use these components in several architectures can be found in section 3.6 where four cases are shown. In the case of the greenfield scenario, we recommend first to follow the steps indicated in section 4.2 for this case of scenarios.

- **How can I find support to deploy the data space instance in my city/community?**

You may find support at the Data Spaces Support Center (DSSC) by mailing to support@dssc.eu or via this link. If you apply to some of the open calls to be launched by the deployment project for the DS4SSCC, you will also have support from this project (information to be provided).

- **How can I keep posted about future evolutions?**

In order to follow all that is happening, you may stay tuned by following the Linkedin or Twitter accounts for DS4SSCC, or via web site.

4.4 Feedback

Under this section, we have collected all the feedback received both directly from the use cases involved in the validation of the architecture and the general comments collected through the various workshops organised with the Stakeholders forum.

Use cases validation

The 4 use cases selected for the customization of the architecture have provided their feedback on the process and result.

Helsinki Feedback

During the reference architecture discussion, initially introduced use case centred around the Heat and Climate Atlasthe evolved into the Helsinki Real-Time Data Space, a broader concept that allows for a phased approach. The Helsinki team found the Data Cooperation Canvas particularly valuable, as it facilitated a multi-faceted exploration of use cases and proved to be user-friendly.

Helsinki city provided their existing real-time data architecture that would be upgraded to be compatible with the Architecture Model. We have upgraded existing architecture to include components of the Architecture Model. Upgrade was very straightforward since their existing architecture was already modular.

In September, we reviewed the customised architecture, compatible with Architecture Model with their team and received the following feedback:

- Proposed Architecture Model compliant architecture is what they were expecting.
- Universal Trust Data Registry is an important cornerstone of authorization and authentication. Development of Verified Credentials related standards on the EU level should be monitored and integrated in the Architecture Model. They see the Gaia-X Registry component as a possible implementation for this component.
- Data Translation Component is important for data interoperability. Data platform vendor should be responsible for its implementation.
- Data Translation Component should connect to the digital platform APIs and not directly to the data store of the digital platform. Architecture Model is designed in this way, but maybe this detail is not seen clearly on the high level architecture.
- Role of the Data Space Connector was not clear to them. We discussed that Data Space Connector is only a concept in this iteration of the Architecture Model and will be added in later iterations.
- Personal Data Intermediary (PDI) should be also defined in the architecture.

Flanders feedback

Flanders as a region represents the digital twin scenario. Within the region however they have cities and municipalities ranging from greenfield, to brownfield to digital twin.

For Flanders, the Agency Digital Flanders was interviewed several times, each time with another focus.

From the smart and sustainable cities and communities perspective they stated the diversity in the field of digital maturity and how they are succeeding as a region to increase collaboration between cities and communities with the Smart Region Office and subsidies. How they ensure cities and communities are left behind through sessions on OSLO and VLOCA (Flemish Open City Architecture). How they encourage the more mature cities to take the lead through spotlighting their achievements, collaborating with them in European projects and participating in the EDIC “Local Digital Twin towards the CitiVerse”. All proving that a smart region approach can make the difference and help truly leverage the efforts of Europe and the cities and communities.

Highlights of the feedback from Flanders are:

- Take into account the diversity in digital maturity also within the greenfields, brownfields and digital twins, it would be useful to state what the minimum requirements and support to measure based on each gradation.
- Make sure to take the context sufficiently in consideration. What is being developed that they can use? What are services that are available that can help? How are solutions being structurally embedded?
- Try to include use cases with a higher level of complexity covering thematic dimensions, not just trying to solve use cases that seem feasible. Start with problems, not the available data.

Valencia feedback

Valencia city represents a brownfield scenario as they already have a well proven data platform which integrates data from many city services. An interview on technical aspects was held with them at the end of July, where they provided inputs for the questions in Annex I. Based on their answers and explanations during the interview, we filled a Data Cooperation Canvas for them.

Additionally, they provided us with their current architecture and some envisaged scenarios of application in their case. We proceed then to customise the proposed high level architecture by expanding their architecture of data platform with the proposed components in the architecture.

We shared both the Canvas and the customised architecture with them and in a new meeting at the beginning of September, we validated with them the proposal.

Main highlights from their feedback were:

- Main benefit of moving to a data space approach would be to save time and money in integration of providers and their data into their platform. Every time that they have to integrate a new service, they have to put in place an implementation project which could be avoided due to the Universal Registry of participants credentials.
- From the two proposed scenarios to them (get access to third-party applications and provide access to third-party applications), the second one fits better into their interest. In their view, accessing data from third parties is less relevant as they prefer to have the data in their data platform, so providers must provide a copy of the required data. However, providing access to third parties in a trusted manner allows the unification of the access to city data and they could even monetize the data.
- An additional scenario is of their interest: get access to federated data. In this way, they could get access to data from other cities and in the other way around. This would provide services to the citizens for example to benchmark their city with other cities (i.e. number of trees per city).
- The proposed technical insights are arriving just in time as they are now defining the smart clauses in the contracts with providers.

- They appreciate any proposal that avoids any new implementation in the city platform. In our approach, it is mostly a matter of configuring some existing components (if they rely on existing implementations of the Authorization Policies Store).
- In relation to the Data Space Connector, they would like to follow how it is finally developed and open to test it when ready.
- Final reflection on the commonality of standards. In their view, the only way to reach full interoperability is to agree on common standards for exchanging data.

Amsterdam feedback

Throughout the DS4SSCC blueprint development, the IDEA project from Amsterdam has been a strong partner to provide input and share experiences throughout stakeholder forums, workshops, and online meetings throughout the spring and summer period of 2023. They presented the current state of the project, outlined technical details and challenges they experienced, that have been collected and presented in deliverable D4.1 in detail. Due to this course, the cooperation grew stronger and continued together with IDEA, supporting the development of the Data Cooperation Canvas and adding significant contributions to the establishment of the action plan (D4.2).

IDEA represents a brownfield scenario, as the use case is up and running and has been proven for the City of Amsterdam. Besides the completed Data Cooperation Canvas, they provided us with their current architecture and from interviews and meetings, we could define a high-level scenario for aims and goals of the use case. Based on this, we were then able to suggest a high-level customised architecture by the extension of the IDEA architecture, and shared the outcome in the beginning of September 2023 with involved IDEA stakeholders. Here main take-aways of the development and their feedback:

- Proposed architecture was as expected and they were excited to see how IDEA can be implemented into the DS4SSCC.
- As long as stakeholders were less familiar with the components, there was some additional information needed to each of the added components and what their purpose was.
- During development of the customised architecture, the role of the Data Space Connector was not clear and had to be adapted, as long as the Data Space Connector has to be seen as a concept in this iteration of the model which will be detailed in later stages.

Stakeholders validation

The Stakeholders Forum have provided their feedback through two of the workshops organised for them.

In one of them (on 7th June), we presented the High Level architecture approach and through a MIRO Board, we collected their feedback through a set of three questions, which can be summarised as follows:

Questions	Inputs
Are the proposed existing scenarios (brownfield, greenfield, digital twin) fitting with your case? If not, which is your case?	Some participants confirmed the use of data platforms in their cities. Some are just at the definition of their data space architecture, so this architecture is welcome. Concern about how to fit the data platforms in cities into the common national infrastructure of each country.
Does the presented approach sound to you? What do you miss? What would you change?	More details on how to guarantee the Data Exchange and Trust mechanisms. Sharing data with neighbour cities. How to deal with the annual wheel of operations. Good as the first step, but needed to look at the potential software to do it. Build on top of existing platforms.
What would you appreciate to get as a CookBook to facilitate the deployment of this architecture in your city/community?	How to move from paper to reality. Guidelines, support and validation of implementation. Avoid excluding rare cases. Monthly report about results. Provide a “hello world” use case to test.

According to the provided feedback in this workshop:

- The presented architecture is welcome, although it represents only the first step and further detail is expected in the future.
- There is a common concern about leveraging existing data platforms in the cities and being inclusive with those cities that are not already at that level of digitalization.
- The stakeholders are expecting from DS4SSCC to provide guidelines and support to land this approach into their specific case.

In a second workshop (on 6th September), one example of how to customise the high level architecture in one of the selected use cases was presented to the stakeholders and they provided also their feedback below:

Questions	Inputs
Have you perceived the architecture comprehensive? If not, why?	Too generic components that can be misinterpreted. Useful to have use case examples and sequence diagrams.
Have you seen the presented example illustrative for your own case? Which are the similarities and differences?	Too focused on digital exchange, individuals are missing. Helsinki use case presented with few details, not clear how FIWARE components are used. Other different examples would have been welcome.
Which are the main challenges and difficulties you envisage in your case in applying this approach?	Existing building blocks are not so interoperable (from a practitioner in i4Trust project). Missing building block in personal data control (example of fairsfair.org). No clear what gain who share data Migrating identities to a new system (IDM).

Following the received feedback, the main conclusions were:

- The presented components are too generic, more detail is desired
- Citizens and personal data are missing in the architecture
- The stakeholders would like to see other examples beyond the presented about Helsinki
- Existing solutions are not very much interoperable, so this will be a big challenge

5 CAMSS Self-Assessment

CAMSS is the European guide for assessing and selecting standards and specifications for an eGovernment project, a reference when building an architecture and an enabler for justifying the choice of standards and specifications in terms of interoperability needs and requirements.

The main objective of CAMSS is achieving interoperability and avoiding vendor lock-in by establishing a neutral and unbiased method for the assessment of technical specifications and standards in the field of ICT. This method will be compliant with the Regulation 1025/2012 on European Standardisation.

As identified in section 2.4, the EIF is of high relevance for DS4SSCC, since it provides a reference framework for interoperability in public administrations. Therefore, CAMSS provides a self-assessment to validate the alignment with EIF scenario, so relevant as well for DS4SSCC architecture which should be compliant with EIF.

The CAMSS Assessment of EIF is available for use on self-assessments via Joinup. The CAMSS Team uses the CAMSS Assessment EIF Scenario on a regular basis to produce assessments using this scenario. The use by CAMSS Team allows for the detection and improvement of the scenario as well as the current document. A repository with all the assessed architectures is also available here.

The Annex II includes the self-assessment checklist for assessing the architecture against CAMSS EIF Scenario which is the most relevant for the DS4SSCC project. The assessment of the DS4SSCC architecture will be carried out during the deployment phase when the architecture will be consolidated and proven by the selected pilots.

6 Conclusions, future challenges and recommendations

The elaboration of this Reference Architecture Model has been a great challenge. This is due to the different typologies of cities and communities that have been analysed and the difficulty to come up with a unique yet abstract enough architecture. Another challenge is the existence of many reference architectures and technologies to define and implement data spaces. Finally, we were also met with the challenge to evolve the existing data platforms which are using different standards and technologies not always compatible. Our ambition is to propose a solution for the commonality of the cases, but also to provide recipes to those cases that are not yet ready for this scale of evolution.

As a result, the document collects the most adopted reference architectures in the domain, analyses the possible technical scenarios and proposes an evolution of the existing data platforms in communities towards the data spaces. To do that, three main components are defined: Universal Trust Data Registry, Authorization Policies Store and Federated Layer. The high level architecture is complemented with an emulation flow to show how the architecture would work in a sequence. Additionally, some future evolutions are mentioned like the use of Data Space Connectors for an easy deployment of these and other componentes; the evolution of Verifiable Credentials and towards decentralised marketplaces.

With the purpose to illustrate how the proposed architecture would land in concrete cases, four use cases have been selected to customise the architecture (Helsinki, Amsterdam, Flanders and Valencia). The Data Cooperation Canvas was completed for all the cases and a set of recommendations were listed for the participant cities and regions. These use cases have also provided very useful feedback for the future evolution of the architecture.

Finally, a Cookbook was produced. It includes a set of short guidelines as a starter kit; then a set of recipes to follow for each of the identified scenarios and finally, a set of FAQs with major tips to manage data spaces in the smart communities field.

The presented Reference Architecture Model and CookBook are just the starting point of a long path which will be continued in the deployment of the data space under the DS4SSCC-DEP project. Under this new endeavour the architecture will be extended and piloted through several pilots in open calls. Besides, this material will also be of relevance for the Testing Experimental Facilities (TEF) project called Citcom.ai which is providing testing infrastructures for cities and communities to deploy their data spaces.

The main recommendations for these further steps projects, in line what is recommended by the European Interoperability Framework are:

- Establish and keep a common language to foster the interoperability in data sharing (see Catalogue of Specifications - Data Models BB).

- Define and use stable and standard interfaces (APIs) for data exchange, relying on which are already commonly adopted by the SSCC community (see Catalogue of Specifications - Data Exchange BB).
- Rely on existing BBs, standards and reference implementations (see overall Catalogue of Specifications).
- Follow the architecture template proposed in this document to ensure the alignment with Data Spaces Support Center, MIMs and most recent trends and market adopted technologies. Expand it with still evolving components and concepts up to the highest level of concreteness and usability by cities and communities.

Annex I: Technical questionnaire for selected use cases to customise the architecture

Aiming at providing a customised architecture for some selected use cases of data spaces in smart cities and communities which represent the most frequent scenarios in the domain, the WP3 team of DS4SSCC project kindly request these selected use cases to provide the information below. This information is essential to understand exactly what is already available at each location and provide the right architecture and recipes to become a data space. The collected information will be only used by this purpose, and the produced result will be duly validated with the stakeholders before publication.

- Please name end users of your solution (please group them as much as possible like individual/company):
  - Name of end user no 1:
    - What are you using for identification?
    - What are you using for authorisation?
  - Name of end user no 2:
  - …
- Please list all datasets most frequently consumed by you (think from perspective, from which solutions you are getting the data, where you are in a data consumer role):
  - Name of dataset no. 1:
    - What is the solution name?
    - Who owns the data (stakeholder name)?
    - Where is the data published? Is the data publicly available?
    - Is this data about individuals (GDPR)?
    - What are the data usage rights?
    - What do you use for data model interoperability (example: Smart Data Models)?
    - What do you use for API interoperability (example: NGSI-LD)?
    - What are you using for identification?
    - What are you using for authorisation?
  - Name of dataset no. 2:
  - …
- Please list all datasets most frequently provided by you (think from perspective, which solutions you are providing the data to, where you are in a data provider role):
  - Name of dataset no. 1:
    - What is the solution name?
    - Do you own the data? If not, who owns the data (stakeholder name)?
    - Where is the data published? Is the data publicly available?
    - Is this data about individuals (GDPR)?
    - What are the data usage rights?
    - What do you use for data model interoperability (example: Smart Data Models)?
    - What do you use for API interoperability (example: NGSI-LD)?
    - What are you using for identification?
    - What are you using for authorisation?
  - Name of dataset no. 2:
  - …

Annex II: CAMMS EIF Scenario self-assessment checklist

Below is the checklist for all the categories listed in the document CAMSS EIF Scenario v6.0.0.

Category	Compliant (Full/Partially/Not)	Compliant with EIF Recommendation	Comments
SUBSIDIARITY AND PROPORTIONALITY			
Criterion 1 (A1) – To what extent has the specification been included in a national catalogue from a Member State whose National Interoperability Framework has a high performance on interoperability according to National Interoperability Framework Observatory factsheets?			
OPENESS			
Criterion 2 (A2) – Does the specification facilitate the publication of data on the web?			
CriteCriterion 3 (A3) – To what extent do stakeholders have the opportunity to contribute to the development of the			

specification?			
Criterion 4 (A4) – To what extent is a public review part of the release lifecycle?			
Criterion 5 (A5) – To what extent do restrictions and royalties apply to the specification's use?			
Criterion 6 (A6) – To what extent is the specification sufficiently mature for its use in the development of digital solutions/services?			
Criterion 7 (A7) – To what extent has the specification sufficient market acceptance for its use in the development of digital solutions/services?			
Criterion 8 (A8) – To what extent has the specification support from at least one community?			
TRANSPARENCY			
Criterion 9 (A9) – To what extent does the specification enable the visibility of administrative procedures, rules data, and services?			
Criterion 10 (A10) – To what extent does the			

specification scope comprehensibly administrative procedures, rules data, and services?			
Criterion 11 (A11) – To what extent does the specification enable the exposure of interfaces to access the public administration's services?			
REUSABILITY			
Criterion 12 (A12) – To what extent is the specification usable beyond the business-specific domain, allowing its usage and implementation across business domains?			
TECHNOLOGICAL NEUTRALITY AND DATA PORTABILITY			
Criterion 13 (A13) – Is the specification technology agnostic?			
Criterion 14 (A14) – Is the specification platform agnostic?			
Criterion 15 (A15) – To what extent does the specification allow for partial implementations?			
Criterion 16 (A16) – Does the specification allow customisation?			

Criterion 17 (A17) – Does the specification allow extension?			
Criterion 18 (A18) – To what extent does the specification enable data portability between systems / applications supporting the implementation or evolution of European public services?			
USER-CENTRICITY			
Criterion 19 (A19) – To what extent does the specification allow relevant information to be reused when needed?			
INCLUSION AND ACCESSIBILITY			
Criterion 20 (A20) – To what extent does the specification enable the e-accessibility?			
PRIVACY			
Criterion 21 (A21) – To what extent does the specification ensure the protection of personal data managed by Public Administrations?			
Criterion 22 (A22) - Does the specification provide means for restriction to access to information/data			

Criterion 23 (A23) - Is the specification included in any initiative at European or National level covering privacy aspects?			
DATA EXCHANGE AND PROCESSING			
Criterion 24 (A24) – To what extent does the specification enable the secure exchange of data?			
Criterion 25 (A25) – To what extent does the specification enable the secure processing of data?			
DATA AUTHENTICITY			
Criterion 26 (A26) – To what extent the specification guarantees the authenticity and authentication of the agents involved in data transactions?			
DATA INTEGRITY			
Criterion 27 (A27) – To what extent is information protected against unauthorised changes?			
DATA ACCURACY			

Criterion 28 (A28) – To what extent does the specification ensures and enables data processing accuracy?			
ACCESS CONTROL			
Criterion 29 (A29) – To what extent does the specification provide an access control mechanism?			
MULTILINGUALISM			
Criterion 30 (A30) – To what extent could the specification be used in a multilingual context?			
ADMINISTRATIVE SIMPLIFICATION			
Criterion 31 (A31) – Does the specification simplify the delivery of European public services?			
Criterion 32 (A32) – Does the specification enable digital service delivery channels?			
PRESERVATION OF INFORMATION			
Criterion 33 (A33) – To what extent does the specification enable the long-term preservation of			

data/information/knowledge (electronic records included)?			
ASSESSMENT OF EFFECTIVENESS AND EFFICIENCY			
Criterion 34 (A34) – To what extent are there assessments of the specification's effectiveness?			
Criterion 35 (A35) – To what extent are there assessments of the specification's efficiency?			
INTEROPERABILITY GOVERNANCE			
Criterion 36 (A36) – Is the (or could it be) specification mapped to the European Interoperability Architecture (EIRA)?			
Criterion 37 (A37) – To what extent can the conformance of the specification's implementations be assessed?			
Criterion 38 (A38) – Is the specification recommended by an European Member State?			
Criterion 39 (A39) – Is the specification selected for its use in an European Cross-border			

project/initiative?			
Criterion 40 (A40) – Is the specification included in an open repository/catalogue of standards at national level?			
Criterion 41 (A41) – Is the specification included in an open repository/catalogue of standards at European level?			
LEGAL INTEROPERABILITY			
Criterion 42 (A39) – Is the specification a European Standard?			
ORGANISATIONAL INTEROPERABILITY			
Criterion 43 (A43) – Does the specification facilitate the modelling of business processes?			
Criterion 44 (A44) – To what extent does the specification facilitate organisational interoperability agreements?			
SEMANTIC INTEROPERABILITY			
Criterion 45 (A45) – Does the specification encourage the creation of communities along with the sharing of			

their data and results in national and/or European platforms?			

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	EURO CITIES	European Network of Living Labs

FIWARE	AIT AUSTRIAN INSTITUTE OF TECHNOLOGY	Gospodarska zbornica Slovenije Chamber of Commerce and Industry of Slovenia