DATES

European Data Space for Tourism

Deliverable D3.3

DESIGN FOR THE DATA SPACE FOR TOURISM

Funded by the European Union

31/10/2023

PROJECT INFORMATION

Project Number	101084007
Project Acronym	DATES
Project Full title	European Data Space for Tourism
Project Start Date	1 November 2022
Project End Date	31 October 2023
Project Duration	12 months
Funding Instrument	Digital Europe
Funding Scheme	CSA – Coordination and Support Action
Topic	DIGITAL-2021-PREPACTS-DS-01-TOURISM
Contract	101084007
Coordinator	AnySolution S.L.

DELIVERABLE INFORMATION

Deliverable No	D3.1
Deliverable Title	Design for the dataspace for tourism
Work-Package No	WP3
Work-Package Title	Design and Technical Framework
WP-Leader (Name and Short Org. Name)	TECNALIA - TEC
Tasks No	3.3 and 3.4
Tasks Title	Design of a Data Space for Tourism in Europe – Data space connector
Task Leader (Name and Short Org. Name)	TECNALIA -TECN
Main Author (Name and Short Org. Name)	Valentín Sánchez (TECN)
Other Authors (Name and Short Org. Name)	Andreas Eisenrauch (AMAD), Belén Martinez (TECN), Jesús Herrero (TECN), Martin Soutschek (OUTD), Dolores Ordóñez (ANYSOL)
Reviewers (Name and Short Org. Name)	Eric Pol (ANG), Umutcan Serles (STI), Tayrne Butler (ANYSOL)
Status	Draft ☒ Final ☐
Deliverable Type	Report ☒ Data ☐ Demonstration ☐ Data ☐
Dissemination Level	Public (PU) ☒ Sensitive (SEN) ☐ EU Classified (EUC) ☐
Date Approved by Coordinator	31/10/2023

DOCUMENT REVIEW HISTORY

Date	Version	Reviewer Name (Short Organisation Name)	Summary of main changes
01/10/2913	V0.1	Valentin Sánchez (TECN),	
15/10/2023	V0.2	Andreas Eisenrauch (AMAD), Jesús Herrero (TECN), Eric Pol (ANG), Belén Martinez (TECN), Umutcan Serles (STI), Martin Soutschek (OUTD)	First version
31/10/2023	V1.0	Valentin Sánchez (TECN)	Final version

FINAL ACCEPTANCE BY COORDINATOR: 31/10/2023

DATES CONSORTIUM			
Part. No	**Participant organisation name**	**Short Name**	**Country**
1	ANYSOLUTION SL	ANYSOL	ES
2	FUNDACION TECNALIA RESEARCH & INNOVATION	TECN	ES
3	IDC ITALIA SRL	IDC	IT
4	MINISTERO DEL TURISMO	MITUR	IT
5	INTELLERA CONSULTING SRL	INTELLERA	IT
6	AMADEUS SAS	AMAD	FR
7	INTERNATIONAL DATA SPACES EV	IDSA	DE
8	ARCTUR RACUNALNISKI INZENIRING DOO	ARCTUR	SI
9	NETWORK OF EUROPEAN REGIONS FOR COMPETITIVE AND SUSTAINABLE TOURISM ASBL	NECS	BE
10	OUTDOORACTIVE AG	OUTD	DE
11	DIH TOURISM 4.0, ZU	DIHT	CZ
12	AVORIS RETAIL DIVISION SL	AVOR	ES
15	ANEWGOVERNANCE	ANG	BE

INDEX

Index of Figures

List of Abbreviations and Acronyms	
AISBL	Gaia-X Association for Data and Cloud
API	Application Programming Interface
AI	Artificial Intelligence
BAE	Business Application Ecosystem
C2D	Compute-to-data
C2E	Compute-to-edge
CKAN	The open source data management system
CSV	Comma Separated Values
DCAT	Data Catalog Vocabulary
DSBA	Data Spaces Business Alliance
DGA	Data Governance Act
DID	Decentralized identifier
DMO	Destination Management Organization
DQV	Data Quality Vocabulary
DSA	Data Space Authority
DSSC	Data Spaces Support Center
DATES	Data Space for Tourism
EC	European Commission
EDIB	2.3.4 European data innovation board
eID	Electronic identification
eiDAS	electronic IDentification, Authentication and trust Services
EU	European Union
FAIR principles	Findable, Accessible, Interoperable and Re-usable
FIWARE	Future Internet Core Platform
GAIA-X	A Federated and Secure Data Infrastructure
GDPR	General Data Protection Regulation
GE	(Fiware) Generic Enabler
GXDCH	Gaia-X Digital Clearing House
GXFS	Gaia-X Federation Services
ICT	Information and Communication Technologies
IDM	Identity Management
IDS	International Data Space
IDSA	International Data Spaces Association
IDS-RAM	IDS Reference Architecture Model
ISO	International Organization for Standardization
MyData	Human-centric approach to personal data management
NGSI	Next Generation Service Interfaces
NGSI-LD	Next Generation Service Interfaces – Linked Data
ODRL	Open Digital Rights Language

OPENDEI initiative	Aligning Reference Architectures, Open Platforms and Large-Scale Pilots in Digitising European Industry
OTA	Open Travel Alliance / Open Travel Agency
PDP	Policy Decision Point
PEP	Policy Execution Point
PoC	Proof of Concept
POI	Point of Interest
REST	Representational state transfer
Simpl	EU project to design and implement the Smart middleware platform
SME	Small and Medium Enterprise
SOLID Pod	Decentralized data stores
SP	Service Provider
TA	Trust Anchor
TDS	Trusted Data Sources
UNWTO	United Nations World Tourism Organization
VC	Verifiable credential
VP	Verifiable presentation
WP	Work Package
W3C	World Wide Web Consortium
XSLT	Extensible Stylesheet Language Transformations

1 PURPOSE OF THE DOCUMENT

This report presents the results of Task 3.3 "Design of a Data Space for Tourism in Europe" and Task 3.4 "Data Space connector". Task 3.3 goal is to define the overall technical infrastructure of a Data Space for Tourism in Europe, taking into account previous results from the DATES project, the existing data spaces initiatives in Europe, the existing EU legislation and data policies, as well as common principles agreed at sector or local levels.

After the analysis of gaps and the inventory of data sets performed in WP2, WP3 analysed the Technical Framework required for the implementation of the Tourism Data Space.

This work has been done in collaboration with the main European bodies and instruments for the efficient deployment of data spaces in Europe, specifically, the **Data Spaces Support Centre** and the main European initiatives in the area of data spaces, mainly GAIA-X, IDSA and FIWARE, ensuring the alignment of the tourism data space with the European Data Spaces Technical Framework and the rest of the ecosystem of data spaces.

According to the proposal and grant agreement the activity of WP3 has focused mainly on:

- Identifying the **technical requirements** for the data infrastructure of tourism data spaces (e.g., technical design, functionality, operation and governance).
- Defining the guiding design principles for the creation of tourism data spaces.
- Identifying architecture and technical data governance frameworks, including the tourism domain's specific elements.
- Identifying the **common building blocks** essential for the creation of tourism data spaces and defining technical specifications.
- Identifying common standards, including semantic standards and interoperability protocols – both tourism-specific and crosscutting.
- Identifying the potential for synergies between data spaces and coordinate related cross-cutting exchanges between data spaces.
- Identifying common toolboxes that could be used across data spaces.
- Identify existing APIs that are relevant for a Tourism Data Space

The previous results have been the identification of the common building blocks, the specific requirements imposed by the tourism sector and some specific aspects of the Tourism Data Space and how they influence the data space technology to be deployed, gathered in D3.1. Furthermore, T3.2 has focused on interoperability, specifically in semantic interoperability based on standard models, ontologies along with methodological guidelines from current European initiatives and open standardization consortia.

Finally, T3.3 has worked on the definition of the overall technical infrastructure of a Data Space for Tourism in Europe, elaborating a technical blueprint for the data space for tourism based on existing EU legislation and data policies, as well as on common principles agreed at sector or local levels.

The report contains the final results of WP3. This document is based on the current state of each of the initiatives that are analysed along the document, the more detailed work in task T3.2, WP2 and WP4 achievements regarding data models and standards, data governance and tourism specific use cases as well as the collaboration with the sister CSA in tourism, which is working in parallel.

Many initiatives are running in parallel defining business cases, the scope, legal framework and technology needed to implement and deploy data spaces. Most of them are far from being stable and are producing open-source components which are not mature enough to be used in a commercial environment. Even worse, those components are evolving at a very high pace and are not interoperable with each other. Each initiative is deploying its own so-called Minimum Viable Product, built with components developed internally and highly dependent on specific versions of the components. Convergence or interoperability among the different initiatives is still an open issue. In this changing and uncertain environment, it is very difficult to know in advance what will be the "winning" initiative will be, if any.

Therefore, the roadmap for the definition and implementation of the Tourism data space and the recommendations included in the report could become obsolete in the future, so it is recommended to establish a technological surveillance process that allows adaptation to the future development of data space technology.

1.1 Strategy for the definition of the technical design (blueprint)

Since the Data Space blueprint (according to the DSSC definition) should contain "A consistent, coherent and comprehensive set of guidelines to support the implementation, deployment and maintenance of data spaces" three different but complementary points of view should be considered in the blueprint.

The data space governance authority point of view (*Governance Authority*).

Define rules and policies (dataspace governance) to become a data space participant, and also for services, resources and intermediaries.

Define/implement the onboarding process. Includes the compliance nodes to operationalize the onboarding process, to certify the rules and policies compliance and provide electronic proof.

Dataspace intermediary stakeholders (*Intermediary service provider*) including among others**,** Identity management, Trust anchors, Federated Catalogue, Logging and auditing, Contract negotiation service and Personal data intermediaries.

The data space participant point of view: Participant journey (*Participant*) definition composed by the following steps:

- **Onboarding**
  - Participant
  - Data product/service

- **Publish/search**
  - Data
  - Services
- **Data products/services**
  - Provider
  - Consumer
- **Monitoring**

Taking into account the current context of the data spaces technology, the different approaches and status of the "Data Space blueprints" being defined by the DSSC and other CSAs, we propose the following structure for the Tourism data space technical blueprint:

1. Section about the Tourism Dataspace Technical Design context
   - **Data space initiatives update**. Findings from the DSSC Blueprint including the conceptual model and the Building blocks taxonomy. Chapter about the updated status of the main data spaces initiatives, with special emphasis on the convergence process between FIWARE, Gaia-x and IDSA
   - **Legal and institutional context**.
2. Section to present the Tourism data space technical design
   - **Data space development roadmap**: Theoretical and pragmatic approaches
   - **Data space stakeholders perspective**: Data space governance authority, intermediaries and participants. Chapter for the description of the **participant journey**, describing how the technical building blocks are applied in each of the stages.
   - **Data models**: Generic and tourism sector specific.
3. Data space connector analysis.
4. Tourism data space specific requirements and open issues.
   - Section on **specific issues** of the "**tourism**" data spaces, including the four already identified and others.

Finally, a section about actionable recommendations has been included.

2 TOURISM DATASPACE TECHNICAL DESIGN CONTEXT

2.1 DATES initial findings

Our objective has been firstly to identify the **main technical issues and challenges to be solved** when implementing the data space concept and to analyse the main initiatives and how those issues are being tackled. These main technical issues are Identity, Self-description, Trust framework, On-boarding, Data and services discoverability, Data sharing, Data space monitoring and observability.

Several ongoing initiatives are now running in parallel to design and implement the data space concept. GAIA-X and Simpl provide by now the more comprehensive data space architectures and specifications including infrastructure, data and services in the same framework while IDSA and FIWARE provide solutions dealing with more specific aspects of the data space technology landscape. The DSBA convergence effort and Simpl project add even more complexity to the data space scenario.

However, even with this complex scenario some solutions and approaches to specific aspects of the data spaces architecture are common or quite similar in all the initiatives, emerging as the most promising ones:

- **Identity and self-description**: The need for a decentralised solution for identity management. Self-sovereign identity solutions along with verifiable presentation and verifiable credentials for self-descriptions (*GAIA-X*)
- **Trust framework and onboarding**: The concept of Data Space Authority defining and implementing the rules to be part of a data space (*all the initiatives*)
- **Data transfer**: Decoupling of control and data planes in data transfer technologies, making it possible to use any transfer protocol or technology available (*IDSA Dataspace protocol*)
- **Generic rule**: Use common and well-established standards if available. (*All the initiatives*). Here are some examples:
    - DCAT for data product
    - ODRL for data usage policies
    - W3C DQV for data quality (based on ISO)

DATES has also identified some specific issues that should be considered when setting up a tourism data space and some initial recommendations about the strategy to address them have been presented:

- **Personal data management**: People are at the centre of the tourism data space, so data protection legislation and initiatives play a very important role.
- **SMEs:** Most companies in the tourism sector are SMEs and the process to onboard in a data space and the technology needed to participate are too complex and very far from their usual business. They do not have the technical nor operational skills to deal with data spaces. Some approaches both in the IDS and GAIA-X context could be very useful for SMEs: **Connector as a service** and the more ambitious **Data Space as a service**.
- **Public and private actors**: Tourism sector stakeholders include both public administrations and private companies with different strategies and objectives regarding data sharing. Open data public administrations' strategy should be aligned with data protection, valorisation and monetisation strategies of private companies.
- **(STRONG) Relation with other data spaces:** Tourism sector is directly related to other sectors such as transport, mobility, energy, construction, health… In this way, data from these sectors should also feed the ecosystem of a tourism data space. Inter and intra data space interoperability is a critical issue for tourism data spaces.

2.2 Data spaces initiatives update

This initial section provides information about the current data spaces initiatives and their expected evolution in the near future. The analysis focuses on the following initiatives, deemed most significant for consideration: Data Space support centre (DSSC), Gaia-X, IDSA, the Data Space Business Alliance (DSBA) and SIMPL.

Diving into the specifics, the DSSC has been actively involved in shaping the Data Space Blueprint[1], placing particular emphasis on the DSSC Glossary[2], the data space Conceptual Model[3], a first description of the technical Building Blocks[4] as well as an initial Landscape of Technical Standards[5].

On the other side, IDSA has taken a distinctive path by publishing the IDSA Rulebook 2.0.[6] This publication offers a comprehensive understanding of the data space concept along with guidelines for its implementation across diverse technical strategies. Additionally, IDSA takes the lead in the development of the IDS dataspace protocol, which is being implemented by some of the most advanced data space connectors.

[1] https://dssc.eu/space/BPE/179175433/Data+Spaces+Blueprint+%7C+Version+0.5+%7C+September+2023
[2] DSSC Glossary | Version 2.0 | September 2023 - Glossary - Data Spaces Support Centre
[3] Conceptual Model of Data Spaces | Version 0.5 | September 2023 - Conceptual Model - Data Spaces Support Centre (dssc.eu)
[4] Building Blocks | Version 0.5 | September 2023 - Building Blocks - Data Spaces Support Centre (dssc.eu)
[5] Collection of Standards and Technologies landscape | Version 1.0 | October 2023; see https://dssc.eu/space/SE1/185794561/Collection+of+Standards+and+Technologies+landscape+%7C+Version+1.0+%7C+October+2023
[6] Cover - IDS Knowledge Base (internationaldataspaces.org)

Turning the focus to Gaia-X, this initiative distinguishes itself with the most mature trust framework, a rigorous compliance process along with the required software infrastructure to operationalise the onboarding process.

In contrast, DSBA is on a mission dedicated to fostering technical convergence among the main data space initiatives. It strives to define a common data space framework while defining the roles that each initiative can fulfil. The result of this analysis is included in the Technical Convergence Discussion Document (v2.0)[7].

Finally, Simpl emerges as a highly promising initiative in the field of data spaces. However, given that the project is yet to commence, and its initial outcomes are not anticipated until the final quarter of 2024, its inclusion in the current analysis remains unfeasible. It should be mentioned, that Simpl depends on the DSSC, as it is supposed to implement the DSSC specification in concrete software or take over already existing DSSC compliant SW.

2.2.1 Data Space Support Centre (DSSC)

The Data Space Support Centre (DSSC) is a three-year project funded by the European Commission as part of the Digital Europe Program. The DSSC will explore the needs of data space initiatives, define common requirements and establish best practices to accelerate the formation of sovereign data spaces as a crucial element of digital transformation in all areas..

Furthermore, the DSSC has started a collaborative process with the different beneficiaries of European-funded initiatives responsible of the preparatory actions for the data spaces of the different sectors. The objective is to co-create the following version of the "DSSC Data Spaces Blueprint" in collaboration with all the relevant initiatives and stakeholders in sectorial data spaces. Specifically, the two tourism CSAs DATES and DSFT projects, are actively contributing to the DSSC work by offering their expertise to ensure that the Data Spaces Blueprint incorporates the specific requirements of the tourism industry.

According to the DSSC, the Data Space Blueprint is ***"...a consistent, coherent and comprehensive set of guidelines*** *to support the implementation, deployment and maintenance of data spaces. The blueprint contains the* ***conceptual model of data space, data space building blocks, and recommended selection of standards, specifications and reference implementations*** *identified in the data space technology landscape."*[8]

Figure 1 offers a graphical representation of the current scope of the DSSC blueprint. The first "official" version (v0.5) of this general data space blueprint was released at the end of September

[7] https://data-spaces-business-alliance.eu/wp-content/uploads/dlm_uploads/Data-Spaces-Business-Alliance-Technical-Convergence-V2.pdf
[8] Data Spaces Blueprint | Version 0.5 | September 2023 - Blueprint - Data Spaces Support Centre (dssc.eu)

2023. As can be seen in the Figure 1, this version encompasses a multifaceted array of components.

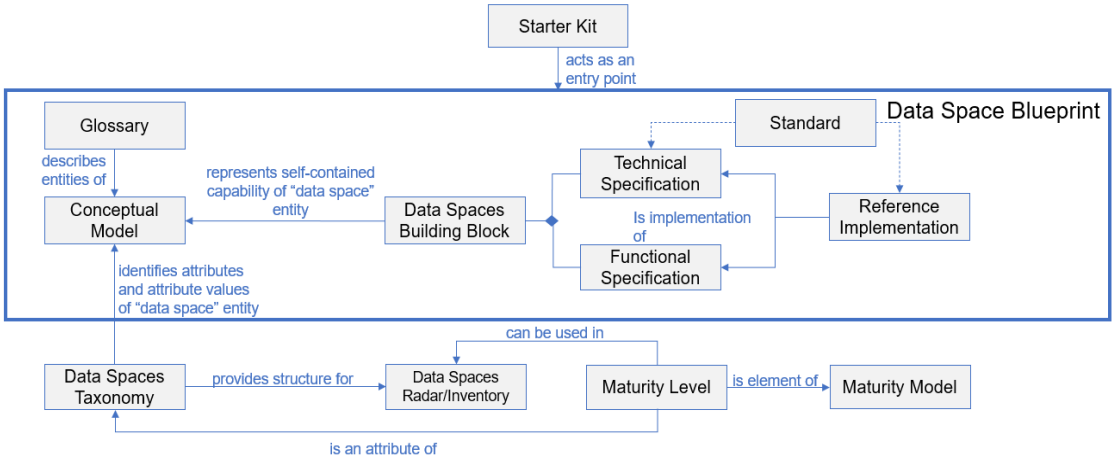

Figure 1 Scope of the DSSC blueprint (copyright: Data Spaces Support Centre)

The DSSC blueprint is still being developed further and the Figure 2 provides a visual representation showing the difference between the content already included in version 0.5 (highlighted in green colour) and the additional elements planned for version 1.0 (highlighted in orange colour).

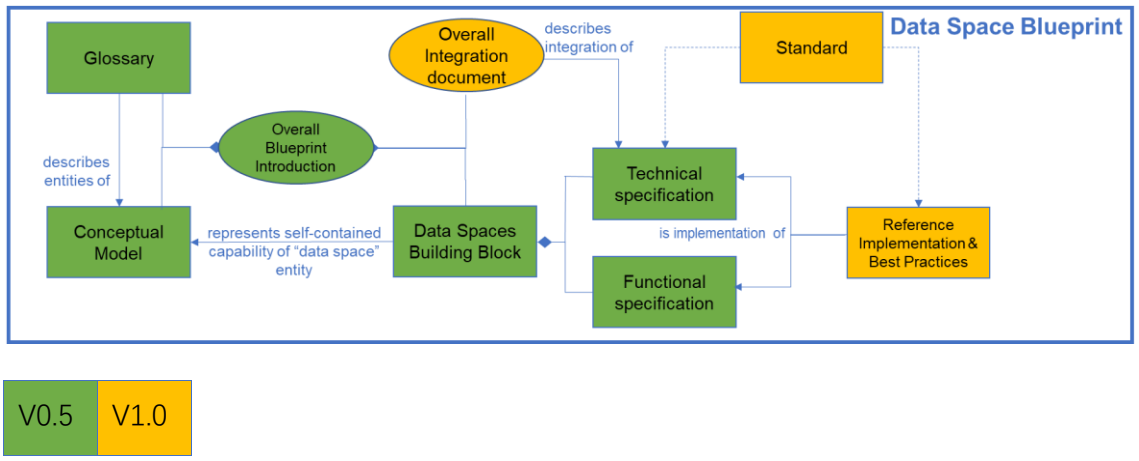

Figure 2 Scope of the Blueprint version 0.5 and 1.0

The DSSC Data Space Blueprint contains very detailed information about the **data space roles** and **concepts** along with the **technical**, governance, business and legal **building blocks** and their implementations in real data space deployments.

However, this blueprint does not address the specific characteristics and modules required for specific sectors like tourism (or any other industry). The present ETDS Blueprint endeavours to complement the general DSSC data space blueprint by offering tailored specific guidelines,

conceptual models, building blocks and specifications that are focused to the particular needs and requirements of the tourism sector.

Besides, the ETDS blueprint incorporates the participant journey perspective, which outlines the steps and experiences organisations go through when engaging with a data space. This would also involve offering guidance on the preparatory work that data space participants, both data providers and consumers, need to undertake before becoming part of a data space ecosystem.

Two main results of the DSSC will be used as a reference for the definition of the main data space related concepts:

- **Glossary 2.0**[9]: This glossary establishes a consistent and coherent terminology for DSSC communication and publications. Beyond the DSSC, this glossary also supports information sharing and co-development between the different data space initiatives and people involved and working with the DSSC. We hope that terminology from the glossary naturally spreads to the community of practice around data spaces, and we hope to get feedback and change requests from the community when needed.

- **Conceptual model Level 1**[10]: The conceptual model of data spaces provides a set of well-defined concepts and relationships between them, as well as a set of terms to refer to them. By relying on the conceptual model, the authors of the data spaces blueprint and the broader community of practice can clearly express data space related topics. Figure 3 shows the DSSC conceptual model level 1.

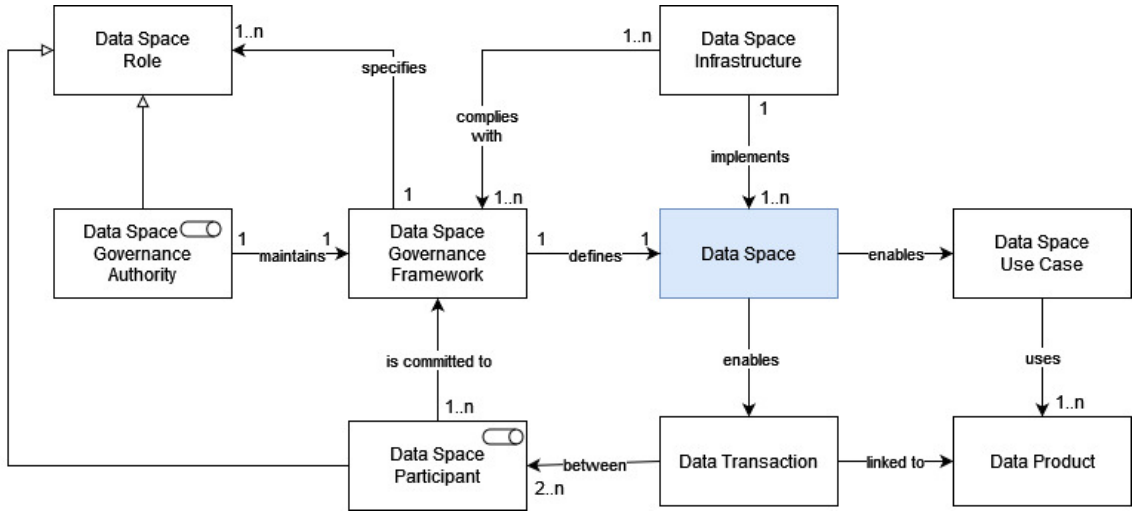

Figure 3 DSSC Conceptual model level 1

[9] https://dssc.eu/space/Glossary/176553985/DSSC+Glossary+%7C+Version+2.0+%7C+September+2023
[10] https://dssc.eu/space/CME/176554182/Conceptual+Model+of+Data+Spaces+%7C+Version+0.5+%7C+September+2023#Conceptual-Model-Level-1

Next, a description of the main concepts taken from the DSSC conceptual model is included. The information is taken directly from the DSSC conceptual model web page, and it is included here for reference and to be used along the document.[11] The concepts are "data space", "data space governance authority", the "data space governance framework" and "data space infrastructures"

A **data space** is a distributed system that enables the secure and trusted exchange and sharing of data. The idea is that individuals, businesses and other types of organisations (collectively called 'parties') can share data in a manner that respects privacy, security and intellectual property rights. Thus, data spaces enable data sovereignty. Data spaces help data users know where to get the data they need and reassure them that the data is valid for their purposes. This seemingly simple concept is, in practice, rather difficult to organise. Data space participants may not know and trust each other, and parties must be enabled to find the data they need. Such data come with the required assurances and policies. Also, there are various interoperability concerns at the technical, procedural, informational and legal levels.

To solve these issues, individual data spaces have their **data space governance framework**, i.e., the set of principles, standards, policies (rules/regulations), agreements and practices that apply to the governance, management, and operations of a data space as well as to the enforcement thereof and the resolution of any conflicts. That governance framework should address anything that participants in a particular data space are concerned about, and that is within the scope of the data space. Thus, the governance framework defines and distinguishes a data space from others.

The governance framework must also specify the requirements that the organisational, procedural and technical components, services and data that implement the data space must comply with. These components (such as connectors), services (such as catalogue services or identity services) and data (such as transaction logging schemas) enable the data space to function. Together, they form **data space infrastructures** that implement the data space.

One specific role regarding the data governance process is the so-called (using Gaia-X nomenclature) **clearing house**. Within the data governance framework, it serves as a pivotal component for operationalising data management. According to Gaia-X, the clearinghouse is the one-stop place to go and get verified against the data space governance rules to obtain compliance in an automated way.

Among many other things, a governance framework specifies several **data space roles**. The roles are distinct and logically consistent sets of responsibilities within a data space, encompassing associated rights and duties required to perform specific tasks and designed to be fulfilled by one or more parties. Every governance framework will at least define the roles of:

[11] Conceptual Model of Data Spaces | Version 0.5 | September 2023 - Conceptual Model - Data Spaces Support Centre (dssc.eu)

- **Data space participant**: the role performed by all parties that have committed themselves to comply with the content of the governance framework. As the governance framework is enforced, all participants can be sure that other participants behave according to the framework and follow the established principles, standards, policies, etc. It is to be noted that individual parties can participate in multiple data spaces and their roles within them can vary (as allowed by their respective governance frameworks).

- **Data space governance authority**: the role can be performed collectively by one or more participants. They are accountable for creating, developing, maintaining and enforcing that governance framework. While doing this, the data space governance authority shall consider the needs of potential future data space participants as well as what regulations are necessary based on the purposes the data space serves. A data space governance authority does not replace or diminish, in any way, the powers of other (e.g., public) governance authorities.

2.2.2 IDSA International Data Spaces Association

The vision of IDSA is to create the environment for trusted data exchange taking place through federated, international data spaces that are globally certified. IDS-certified products, services and systems open the door to a data economy in which businesses can share data up and down the value chain without security concerns. The aim is to create a global standard for Data Spaces, as well as fostering technologies and business models that will drive the data economy of the future in Europe and around the globe.

Two important publications of the IDSA are the IDSA Rulebook, focussing on governance of dataspaces and the IDS-RAM V4[12] which goes into more detail on the architecture of data spaces. The Dataspace Protocol[13] is a set of specifications designed to facilitate interoperable data 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, negotiate usage agreements, and access data as part of a federation of technical systems termed a dataspace.

Certification is a major aspect of the IDS-RAM V4 to achieve Data Sovereignty. Interoperability, and compliance criteria are tested and validated based on an Open-Source testing framework the IDS-Reference-Testbed[14]. The overarching view on a vibrant Data Space community as perceived by IDSA is part of the Data Space Landscape document[15].

Data Sovereignty is expressed in IDS based on human and machine-readable Usage Policies, which can be enforced in technology or organizational manners. Therefore, the definition, description, specification and validation of Data Space Connectors, which implement policy

[12] https://docs.internationaldataspaces.org/ids-ram-4/
[13] https://docs.internationaldataspaces.org/dataspace-protocol/overview/readme
[14] https://github.com/International-Data-Spaces-Association/IDS-testbed
[15] https://internationaldataspaces.org/download/39041/?tmstv=1681683482

negotiation and enforcement are the most important aspect of the IDSA work, including the interaction of the Data Space Connectors with other components and services in Data Spaces.

According to the IDSA Rulebook 2.0, the foundational concepts of a data space, shown in Figure 4, are the following: (i) establishing trust, (ii) data discoverability, (iii) data contract negotiation, (iv) data sharing & usage, (v) observability, (vi) vocabularies and semantic models. Additional elements that support these main functions of a data space are also represented as optional.

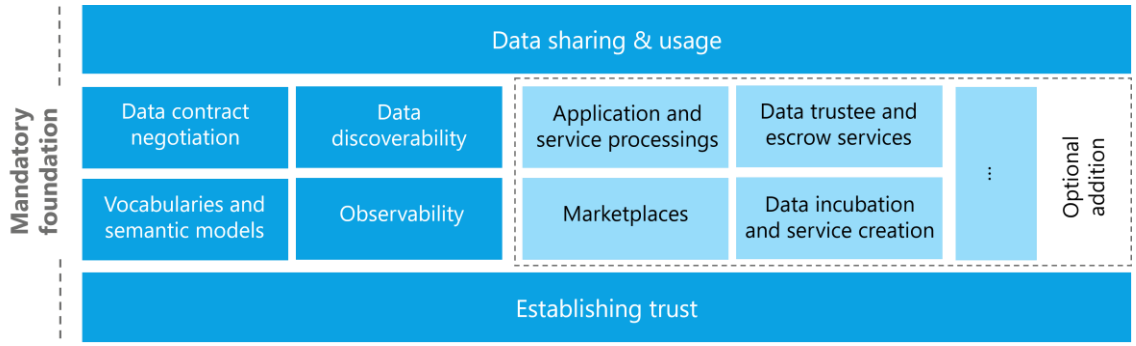

Figure 4: Foundational concepts in Data Spaces according to the IDSA Rulebook 2.0

One of the main achievements of the IDSA initiative is the IDS data space protocol, being used by The Eclipse Data Space components and the FIWARE data space connector (See section 0)

Furthermore, IDSA includes the vocabulary provider for sector data models, playing a pivotal role in fostering standardisation, interoperability, and common understanding within a specific industry or sector. This contributes to better data quality, integration, collaboration and compliance, ultimately facilitating data sharing and interoperability among organisations within that sector.

2.2.3 GAIA-X

Gaia-X aims to create a federated open data infrastructure based on European values regarding data and cloud sovereignty. The mission of Gaia-X is to design and implement a data sharing architecture that consists of common standards for data sharing, best practices, tools, and governance mechanisms.

From the technical point of view, Gaia-X aims to connect the Data and Infrastructure Ecosystems and relies on 3 conceptual pillars to achieve that:

- **Gaia-X Compliance**: Decentralized services to enable objective and measurable trust
- **Data Spaces / Federations**: Interoperable & portable (Cross-) Sector datasets and services
- **Data Exchange**: Anchored contract rules for access and data usage

Gaia-X is the most advanced initiative from the point of view of the technical governance framework and the onboarding process.

Gaia-X has recently announced the adoption of the Eclipse Data Space Connector Technology: *Eclipse Data Space Connector technology is a part of the Gaia-X ecosystem, designed to facilitate data exchange between different entities. It ensures that data can be shared and accessed securely and in a standardized way, aligning with the goals of Gaia-X for data sovereignty and control.*

Regarding the Data product model, he following models have been added to the Registry and can be validated against the Compliance Engine: Resource, PhysicalResource, VirtualResource, DataResource, SoftwareResource, InstantiatedVirtualResource, and ServiceAccessPoint. With these models, the service providers can offer more transparency on the services they provide, and those who offer data resources can now get their Gaia-X Compliant Verifiable Credentials for it.

The Gaia-X compliance wizard now supports eIDAS eIDs to sign proofs for Verifiable Credentials (VCs) and Verifiable Presentations (VPs), and users can now use their eID Smart Card to sign their VCs and VPs directly in the wizard. The compliance engine already had support for eIDAS certificates, so now the end-to-end functionality is available.

The last feature planned is adding the models to define labels as described in the Gaia-X compliance framework, which provides the optional criteria to participate in a Gaia-X data space.

Recently, CATENA-X has deployed a Gaia-X onboarding process for its participants.

Finally, Gaia-X next candidate topics to be developed are data Usage Agreement, Policies and Vocabulary, and Observability.

2.2.4 DSBA: technical convergence document V2.0

The DSBA convergence document is a document published by the Data Spaces Business Alliance (DSBA), a coalition of four European associations that aim to promote data spaces

across Europe and beyond.

The document defines a common reference technology framework for creating data spaces based on the technical convergence of existing architectures and models. The document discusses the key technology pillars for data spaces, such as data interoperability, data sovereignty and trust, and data value creation.

The aim of the convergence on dataspaces as a whole is to avoid dependencies while allowing the overlapping. If a given functionality or component is defined in more than one reference

document, the documents will be updated to allow compatibility and leave it up to every initiative to choose which one to take with no impact on the overall dataspace's architecture.

Next figure shows an example of how the different initiatives could contribute to the reference technology framework.

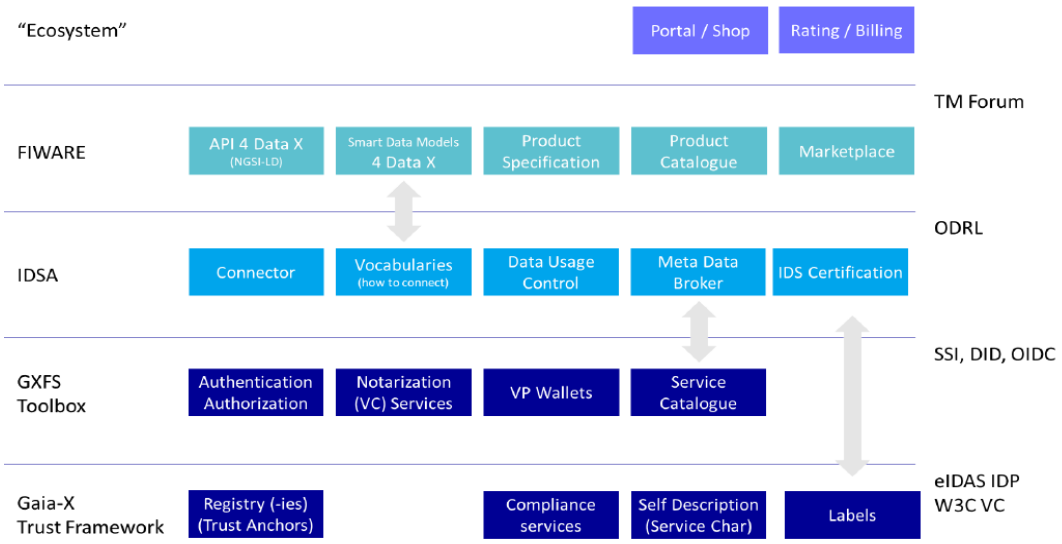

Figure 5 Technical convergence architecture example

FIWARE has implemented a DBSA-compliant data space for demonstration purposes. The data dpace is built with FIWARE-Components, using parts of the i4Trust-Framework and the Gaia-X Compliance Services[16].

2.2.5 FIWARE Building Blocks

The Fiware Foundation has developed a "curated framework of Open-Source Platform components to accelerate the development of Smart Solutions"[17]. With the Fiware Context Broker as its central and mandatory component, Fiware provides many building blocks for a complete data sharing ecosystem.

[16] fiware-gitops/aws/dsba at master · FIWARE-Ops/fiware-gitops (github.com)
[17] https://www.fiware.org/about-us/

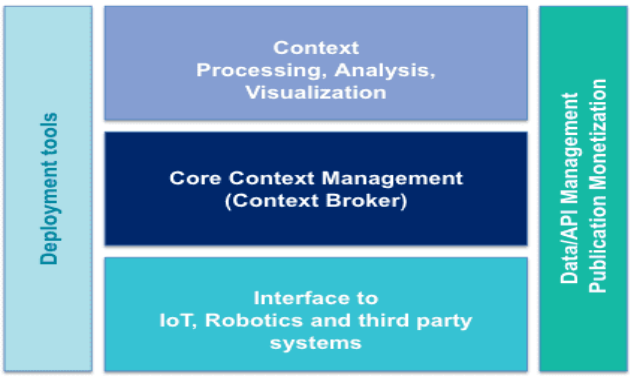

Figure 6 Fiware component schema[18]

Since in this context it is not possible to go too much into the details of their proposal, the following subsections should point out the most remarkable elements with respect to reuse and integration into future data spaces.

2.2.5.1 FIWARE Context Broker

The FIWARE Context broker ("Orion") is a C++ implementation of the NGSIv2 REST API binding developed as the central part of the FIWARE platform[19]. Orion Context Broker allows a Data Space to manage the entire lifecycle of its context information including updates, queries, registrations and subscriptions. With the FIWARE Context Broker in centre, a rich suite of complementary open source FIWARE Generic Enablers has been made available, dealing with the following features:

- **Interfacing with the Internet of Things (IoT), Robots and third-party systems**, for capturing updates on context information and translating required actuations.
- **Context Data/API management, publication, and monetization**, bringing support to usage control and the opportunity to publish and monetize part of managed context data.
- **Processing, analysis, and visualization of context information** implementing the expected smart behaviour of applications and/or assisting end users in making smart decisions.

The FIWARE components are modules that can be combined with and integrated to various third party elements, to build a hybrid data-exchange platform at the user's convenience.

2.2.5.2 Smart Open Data Models

Data Models play a crucial role because they define the harmonized representation formats and semantics that will be used by applications both to consume and to publish data.

The Smart Data Models board (FIWARE Foundation, TM Forum, OASC and IUDX) are leading a joint collaboration program to support the adoption of a reference architecture and compatible common data models that underpin a digital market of interoperable and replicable smart

[18] https://www.fiware.org/catalogue/
[19] https://github.com/telefonicaid/fiware-orion/

solutions in multiple sectors. Indeed, a smart data model for tourism is available from the FIWARE GitHub repository[20]. A smart data model includes three elements: The schema, or technical representation of the model defining the technical data types and structure, the specification of a written document for human readers, and the examples of the payloads for NGSIv2 and NGSI-LD versions.

2.2.5.3 Business Application Ecosystem (BAE)

FIWARE proposes a Business Application Ecosystem (BAE). This is a marketplace component that is made up of the combination of the FIWARE Business Framework and a set of APIs provided by the TMForum. It supports 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.

See the following list of backend components and APIs associated to the FIWARE BAE Marketplace[21]:

- Backend implementing standard TM Forum APIs supporting configuration of the marketplace:
  - Catalog Management API
  - Product Ordering Management API
  - Product Inventory Management API
  - Party Management API
  - Customer Management API
  - Billing Management API
  - Usage Management API
- Rating, Charging, and Billing backend;
- Revenue Settlement and Sharing System;
- Authentication, API Orchestrator, and Web portal.

2.2.5.4 FIWARE CKAN monetization architecture

FIWARE has defined and implemented an open-source architecture for data products the is able to include open and private data and static data sets along with "real time" data via APIs.

CKAN is used as the catalogue where data can be published, discovered, managed, and consumed. Within the FIWARE platform, not only static datasets (CSV, XLSX, etc) are supported, but also real (right) time context information can be published in CKAN as Context Broker queries.

[20] https://github.com/smart-data-models/SmartDestination
[21] https://www.fiware.org/wp-content/uploads/FF_PositionPaper_FIWARE4DataSpaces.pdf

In addition, using the Biz Ecosystem GE, FIWARE provides the components that allow to manage the access and monetisation of published datasets.

The following picture shows the architecture followed in FIWARE for the publication and monetization of data.

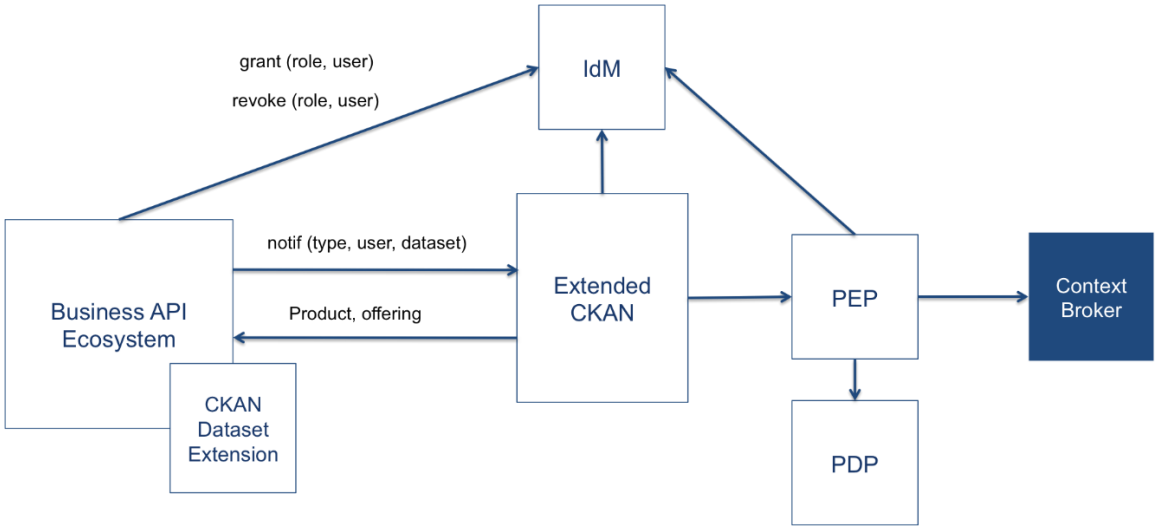

Figure 7 FIWARE Monetization architecture

The following components are integrated in order to manage and monetize data:

- **Extended CKAN**: Used for the publication of the static datasets and the Context Broker queries.
- **Biz Ecosystem**: Used for the creation and acquisition of data offerings, including customer charging, and access grant
- **IdM + PEP + PDP** (Security Framework): Used for the provision of identities, authentication, and authorization of users.

2.2.5.5 FIWARE technology update regarding data spaces

Several FIWARE technologies are well positioned to take part of the future data space implementations:

- FIWARE Data space connector (see section 0)
- FIWARE CKAN monetization architecture, including the Context Broker and the BAE Marketplace

FIWARE has developed the first DBSA-compliant dataspace for demonstration purposes using the FIWARE data space connector[22]

[22] fiware-gitops/aws/dsba at master · FIWARE-Ops/fiware-gitops (github.com)

Furthermore, the BAE Marketplace is being used by DOME, the project developing the market place for Cloud Services in Europe changing the security framework by a new one using DIDs and Verifiable Credentials.

Finally, the BAE APIs are being used to negotiate contracts within the FIWARE connector.

2.2.6 Simpl

As an important element in implementing its Data Strategy, the European Commission announced its intention to develop a Smart middleware platform (so called "Simpl"). The information in this section has been taken from the tender documentation.

Simpl is a project financed by the Digital Europe program of the European Commission, endowed with 150 million euros and with an execution period of three years.

Its objective is to make software (middleware) available to society for the construction of data ecosystems and infrastructure services in the cloud that support the European values of data sovereignty, privacy and a fair market. Such a middleware would enable cloud-to-edge federations and support major data initiatives funded by the European Commission, such as common European data spaces. In particular, its main goal is to provide data space operators with a common denominator technical foundation that will ensure adequate levels of trust, security, ease of access, adaptability, and interoperability.

Under the name of **Simpl**, the tender specifications refer to three distinct products:

- **Simpl-Open**: the core product of Simpl. It is an open-source software stack over which Simpl-Labs and Simpl-Live are built.
- **Simpl-Labs**: the provision and management of a pre-installed demonstration/playground environment where third parties (typically sectoral data spaces in their early phases of deployment) can experiment with/on the most up-to-date version of Simpl-Open before deploying it for their own needs.
- **Simpl-Live**: the provision of several instantiations of Simpl-Open in the form of customised production environments for sectoral data spaces where the European Commission itself plays an active role in their management.

Next figure shows the Simpl-Open Conceptual architecture showing the scope of the project.

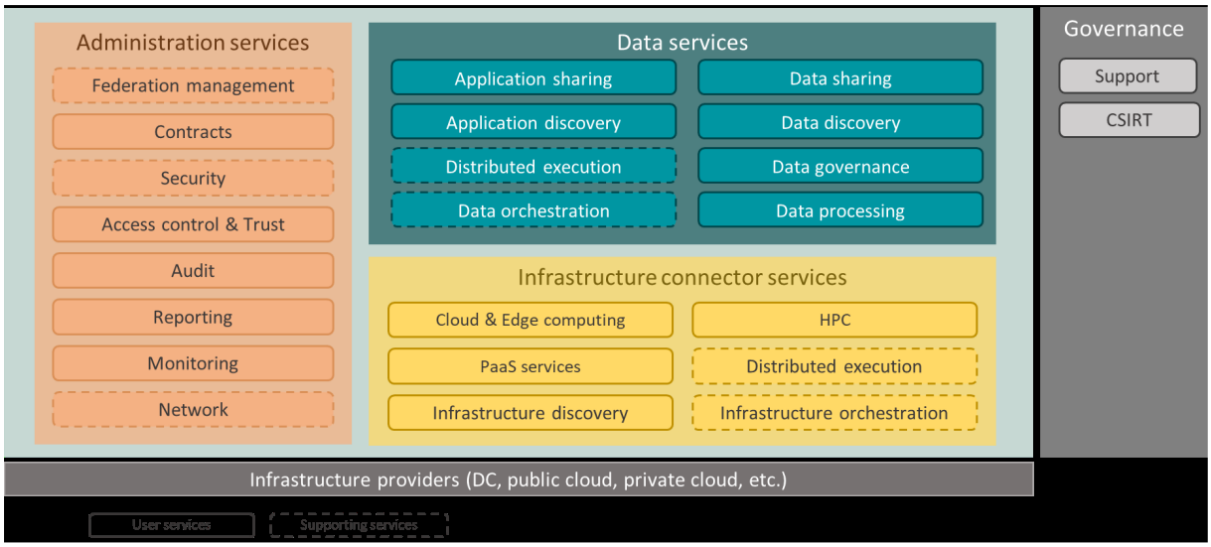

Figure 8 Simpl-Open Conceptual architecture

The architecture is based on four layers:

- **The administration services layer** is a support layer providing user and supporting administrative services that are used by the data and infrastructure service layers. Management and supervision services are included here to enable better control and interoperability among Simpl-Open's services. Issues such as security, personal data protection, access control, federation management, contract management, monitoring or reporting are integrated.
- **The data services layer** enables the exchange of data assets and applications by offering capabilities such as application sharing, data sharing, discovery (of data and applications) and data-relative functionalities such as data processing, data governance and data orchestration. More specifically, the data service layer also includes the implementation of a set of mechanisms to allow different modalities of data access (e.g., online streaming, batch download, ...), transfer (bulk transfer and data streaming) and integration capabilities to merge data coming from various sources or data spaces.
- **The infrastructure connector services layer** manages the infrastructure–related assets and services, allowing to also connect a variety of heterogeneous third-party infrastructural services. The services provided herein allow the finding and provisioning of computing and storage resources so that application workloads can be efficiently and securely managed end-to-end. This layer presents capabilities such as the discovery of infrastructural elements from the catalogue, infrastructure orchestration, deployment and execution of applications, infrastructural services (such as containers and Virtual Machines), and PaaS-related services (e.g., Databases, AI capabilities among others).
- **Governance layer** is a transversal layer supporting the layers above, which focuses on the provisioning of security contingency and personal data protection measures and support capabilities.

The tender also includes an analysis of the relation of Simpl with other initiatives like the EU Digital Building blocks and GAIA-X.

Regarding Gaia-X, the analysis concludes that Gaia-X developments, by means of its Federation Services, are today mainly addressing Simpl-Open Administration Services Access control and trust features.

Beyond these capabilities, other Gaia-X developments such as Self-descriptions and Federated Catalogue could integrate the implementation of the Simpl-Open Infrastructure discovery building blocks along with certain aspects of contracts management, monitoring and infrastructure orchestration.

Overall, the analysis shows that Simpl-Open has a broader perimeter than the current Gaia-X developments. However, existing results have the potential to contribute to Simpl-Open implementation. Gaia-X developments and specifically its Federation Services can act as a source for open-source components which could be integrated, re-utilised and extended as part of Simpl-Open implementation.

The introduction of SIMPL (Smart Middleware) marks a significant advancement in technological infrastructure. In its initial phase, set to be procured by year-end, SIMPL is positioned as a fundamental requirement for all Call of Proposals for the data spaces deployments, ensuring its widespread adoption across diverse domains. Additionally, the mandate to integrate SIMPL technology within all Data Spaces underscores its critical role in facilitating efficient data processing and management.

2.3 Legal and institutional context

This section provides a brief overview of the legal an institutional context in Europe applicable to the data spaces field.

2.3.1 DSSC work on legal building blocks

The DSSC Blueprint v0.5 includes a legal building blocks category which provides guidance and resources for the data space initiatives to ensure compliance with legislation and establish a robust contractual framework. One of the building block is **Regulatory compliance** which present the data space initiatives with an awareness of the legal landscape and aids them in assessing applicable regulatory requirements to ensure legal compliance and alignment with EU values.

Next figure shows a first analysis of the legal framework related to data spaces.

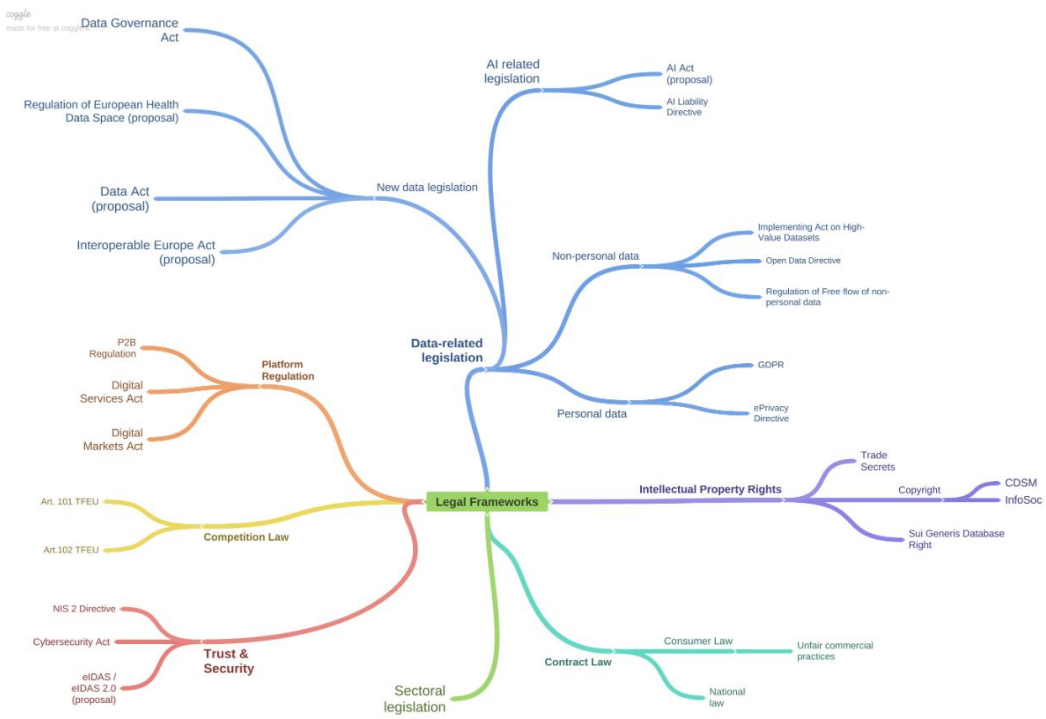

Figure 9 Legal initial mapping of the primary European cross-cutting legislation for data spaces.

2.3.2 Data governance act

The European regulation that aims to create a framework to facilitate European data spaces and increase trust between actors in the data market. The DGA entered into force in June 2022 and applies from Sept 2023. The DGA defines the European Data Innovation Board (EDIB).

The **Data Governance Act** (**DGA**) **¡Error! No se encuentra el origen de la referencia.** is a legislative proposal of the European Commission that aims to create a framework which will facilitate data-sharing, fostering the availability of data for use by increasing trust in **data intermediaries** and by strengthening data-sharing mechanisms across the EU. The instrument addresses the following situations:

- Making public sector data available for re-use, in situations where such data is subject to rights of others.
- Sharing of data among businesses, against remuneration in any form.
- Allowing personal data to be used with the help of a **'personal data-sharing intermediary'**, designed to help individuals exercise their rights under the General Data Protection Regulation (GDPR).
- Allowing data use on altruistic grounds.

According to the DGA **a Personal data sharing intermediary** is a specific category of data intermediaries which provides data sharing services to data subjects in the sense of Regulation (EU) 2016/679 (GDPR: General Data Protection Regulation).

Such providers focus exclusively on personal data and seek to enhance individual agency and the individuals' control over the data pertaining to them. They would assist individuals in exercising their rights under Regulation (EU) 2016/679, in particular managing their consent to data processing, the right of access to their own data, the right to the rectification of inaccurate personal data, the right of erasure or right 'to be forgotten', the right to restrict processing and the data portability right, which allows data subjects to move their personal data from one controller to the other.

In this context, it is important that their business model ensures that there are no misaligned incentives that encourage individuals to make more data available for processing than what is in the individuals' own interest. This could include advising individuals on uses of their data they could allow and making due diligence checks on data users before allowing them to contact data subjects, in order to avoid fraudulent practices. In certain situations, it could be desirable to collate actual data within a personal data storage space, or 'personal data space' so that processing can happen within that space without personal data being transmitted to third parties in order to maximise the protection of personal data and privacy.

2.3.3 Data act

Proposed European Union regulation introducing EU-wide rules for standardised access to the product or related service data to the user of that connected product or service. The proposed Regulation also includes essential requirements for the interoperability of data spaces (Article 28) and essential requirements for smart contracts to implement data sharing agreements (Article 30).

The description relies on the final compromise text. Political agreement is reached, but the act is not yet adopted (in September 2023)

2.3.4 European data innovation board (EDIB)

The EIDB is the expert group established by the Data Governance Act (DGA) to assist the European Commission in the sharing of best practices, in particular on data intermediation, data altruism and the use of public data that cannot be made available as open data, as well as on the prioritisation of cross-sectoral interoperability standards, which includes proposing guidelines for common European data spaces (Article 30, DGA).

2.3.5 GDPR

The GDPR is a European Union law that was implemented May 25, 2018 and requires organizations to safeguard personal data and uphold the privacy rights of anyone in EU territory.

The regulation includes seven principles of data protection that must be implemented and eight privacy rights that must be facilitated:

1. **Lawfulness, fairness and transparency** — Processing must be lawful, fair, and transparent to the data subject.
2. **Purpose limitation** — You must process data for the legitimate purposes specified explicitly to the data subject when you collected it.
3. **Data minimization** — You should collect and process only as much data as absolutely necessary for the purposes specified.
4. **Accuracy** — You must keep personal data accurate and up to date.
5. **Storage limitation** — You may only store personally identifying data for as long as necessary for the specified purpose.
6. **Integrity and confidentiality** — Processing must be done in such a way as to ensure appropriate security, integrity, and confidentiality (e.g. by using encryption).
7. **Accountability** — The data controller is responsible for being able to demonstrate GDPR compliance with all of these principles.

The GDPR recognizes a set of privacy rights for data subjects, which aim to give individuals more control over the data they loan to organizations:

- The right to be informed
- The right of access
- The right to rectification
- The right to erasure (‘right to be forgotten’)
- The right to restrict processing
- The right to data portability
- The right to object
- Rights in relation to automated decision making and profiling.

3 TOURISM DATASPACE TECHNICAL DESIGN

Data is an essential building block of modern societies, and we are constantly producing more of it: citizens create data while browsing the internet, booking a trip, or buying clothes online, and public bodies generate data from weather tracking and traffic monitoring. The European Commission has developed a European data strategy[23] to help unlock its benefits. One key component of this strategy is the development of common European data spaces in strategic economic sectors and domains.

According to the DSSC glossary, a **data space** is a framework that supports data sharing within a data ecosystem. It provides a clear structure for participants to share, trade, and collaborate on data assets in a way that is compliant with relevant laws and regulations and ensures fair treatment for all involved.

In business, data spaces enable the trusted and secure sharing of commercial data assets with automated controls on legal compliance and remuneration. This can create a market among participants or facilitate collaboration among diverse, interconnected parties who rely on each other for mutual benefit. Personal data spaces, on the other hand, allow data subjects and holders to control their data and its subsequent use, ensuring that legislation surrounding the handling of personal data is followed.

According to the IDSA Rulebook, the **foundational concepts** of a data space are the following:

- Establishing of trust
- Data discoverability
- Data contract negotiation
- Data sharing & usage
- Observability
- Vocabularies and semantic models

Additional elements that support these main functions of a data space can include these optional functional areas:

- Application and processing services
- Marketplaces
- Data trustee and escrow services
- Data incubation and service creation

[23] https://digital-strategy.ec.europa.eu/en/policies/strategy-data

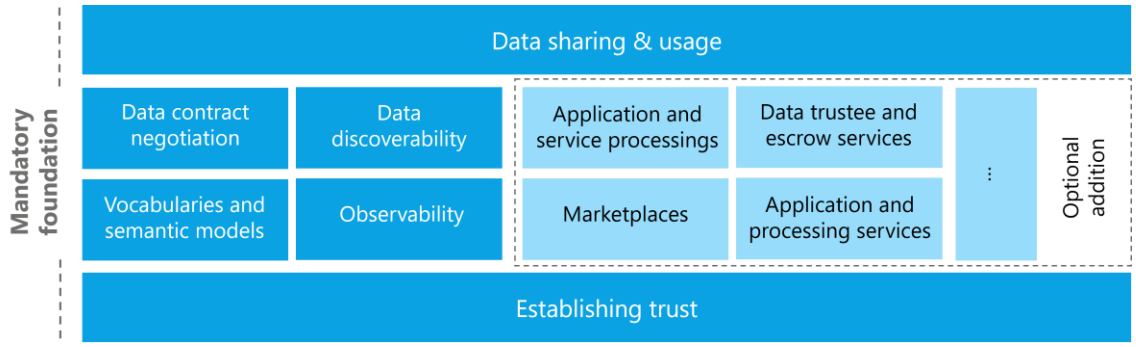

Figure 10 Data spaces foundational concept. Source: https://github.com/International-Data-Spaces-Association/IDSA-Rulebook/blob/main/documentation/media/Foundational_Concepts_in_data_spaces.png

According to the OPENDEI position paper: *"**The technical building blocks** enable the implementation of the technical architecture of a data space. They include network protocols, middleware components, (standardized) APIs, and more, facilitating the sharing of data between different parties in a secure and trustworthy fashion"*. They have defined what they call "Technical pillars" which address most of the technical concerns associated with the creation of data spaces:

- **Data interoperability**, covering aspects such as data exchange APIs, data representation formats as well as data provenance and traceability
- **Data sovereignty**, covering aspects such as identity management, trustworthiness of participants, as well as data access and usage control
- **Data value creation**, covering aspects such as publication of data offerings, discovery of such offerings based on metadata and data access/usage accounting, which are essential to handle data as an economic asset

Based on these high-level categories and the discussions taking place now in the current data space initiatives, we have identified the list of concepts below, that we have used to compare the different approaches. The selected topics do not cover all the concepts and blocks, but they are the ones that are currently under discussion and development among the main data space initiatives.

- Identity
- Self-description
- Trust framework
- On-boarding
- Data and services discoverability: Catalogue
- Data sharing
- Data space monitoring and observability

Two more topics have been included in the analysis, since we think they are very useful in the context of the tourism data space.

- **Service orchestration**: It facilitates the SMEs with limited technical skills to compose several services to obtain the expected result.

- **Personal data management**. Personal data is the core of the tourism data space so it must be managed according to the ethical and legal constraints applicable, especially in Europe.

3.1 Tourism Dataspace development roadmap: technical view

This section includes the description of the steps needed to deploy a data space from the **technical point of view**, providing the context for the recommendations for each of the steps included in next sections.

The recommendations will be based on the current state of the main data space initiatives in Europe, particularly the Gaia-X initiative, which is the most advanced from the point of view of the technical governance framework and the onboarding process.

Most of the technical building blocks needed to operationalize the governance model are sector agnostic. However, some specific requirements from the Tourism sector will be highlighted.

First, the "ideal" process of defining and implementing a data space governance framework from the technical point of view is described. Then, a more pragmatic approach is described.

3.1.1 Data space development formal roadmap

The following steps define the roadmap to define and deploy a "generic" data space governance model, along with some specific steps where requirements from the Tourism sector can be added. These steps are highlighted in red.

It is very important to note that this is a **technical roadmap**, which means that the organizational aspects of the data space are not taken into account. For example, the first step is to create the data space governance authority. If this organization is only one company, a public administration or a consortium is not directly related to the technical approach, so it is not discussed in this section.

1. Define/create the **data space governance authority**.
2. Define the **data space governance framework**.
    a. Compliance process (from the technical perspective)
        i. Mandatory and optional criteria
        ii. Sector specific criteria
    b. Define data models needed to implement the compliance processes
        i. Generic models to describe Data products, natural and legal persons, services, contracts and so on
        ii. Tourism specific data models

3. Operationalize the on-boarding process.
    a. Define and deploy the on-boarding services.
        i. Deploy the necessary Data Space infrastructure
4. Deploy the needed data space intermediaries:
    a. Catalogue
    b. Identity provider(s)
        i. For natural and legal persons
    c. Personal data intermediaries
    d. Logging service
        i. Observability and audit
5. Define and deploy the standard "connector" for the data space, with the following features:
    a. Secure identification and authorization
    b. Secure data transfer
    c. Data sovereignty enforcement

The first three steps are related to the **data space governance authority** role, and it is described more in depth in section 0

The fourth step is related to the data space intermediaries, described in section 0.

The final step related to data space connector is covered in section 0.

3.1.2 Current pragmatic data space deployment

The previous roadmap defines the theoretical roadmap to deploy a data space. However, some of the data spaces already in place have been defined and developed following a more pragmatic approach, based on an incremental technology adoption, not the theorical roadmap.

These are some of the characteristics of these data spaces:

- Some of them do not follow a well-defined onboarding process.
- The connector technology plays a central role from a technological point of view, and the governance framework or Rulebook elements are not addressed in the initial phase.
- The data space has been defined around a use case with a small and "closed" set of companies with a common business case or belonging to a supply-chain. Then, other companies have joined the data space.
- Some of the intermediaries have not been deployed, like for example the catalogue or the logging/audit service.

3.2 Stakeholder perspectives

Since the Data Space blueprint (according to the DSSC definition) should contain "A consistent, coherent and comprehensive set of guidelines to support the implementation, deployment and maintenance of data spaces" three different but complementary points of view should be considered in the blueprint.

The data space governance authority point of view (*Governance Authority*).

Define rules and policies (dataspace governance) to become a data space participant, and also for services, resources and intermediaries.

Define/implement the onboarding process. Includes the compliance nodes to operationalize the onboarding process, to certify the rules and policies compliance and provide electronic proof.

Dataspace intermediary stakeholders (*Intermediary service provider*) including among others**,** Identity management, Trust anchors, Federated Catalogue, Logging and auditing, Contract negotiation service and Personal data intermediaries.

The data space participant point of view: Participant journey (*Participant*) definition composed by the following steps:

- **Onboarding**
  - Participant
  - Data product/service
- **Publish/search**
  - Data
  - Services
- **Data products/services**
  - Provider
  - Consumer
- **Monitoring**

3.2.1 Data space governance authority

According the IDSA Rulebook 2.0 the **data space authority** (DSA) is responsible for establishing the policies and rules of the data space. This role can be carried out by one entity, but also by multiple or even all participants. In a centralized data space, this could be the operating company. In a federated data space, this function would be performed by the federator(s) agreeing on the rules, while in a fully decentralized data space, various mechanisms are available to the participants. The mechanisms in a decentralized data space enable participants to agree on the

set of policies and their enforcement, thus sharing responsibility for the data space authority function.

The **Gaia-X European Association for Data and Cloud AISBL** represents the core of the Gaia-X organisational structure. It is an international non-profit association under Belgian law (French: association internationale sans but lucratif, shortened to AISBL). It was founded to develop the technical framework and operate the Gaia-X Federation services.

In the context of Gaia-X, **Gaia-X AISBL plays the role of the Data space governance authority** defining the data space governance framework for the Gaia-X based data spaces. This **data space governance framework corresponds with the Gaia-X Compliance pillar of the Gaia-X framework.**

Gaia-X Compliance rules is split in two main subsystems:

1. The **'Trust framework'** – which is mandatory and verifies the existence and veracity of any service characteristics. The mandatory criteria are the following:
    a) serialization format and syntax.
    b) cryptographic signature validation and validation of the keypair associated identity.
    c) attribute value consistency.
    d) attribute veracity verification.
2. The **'Policy Rules & Label document**– which is optional and allows to verify adherence to rule-sets that fulfil specific market needs.

Those rules apply to all Gaia-X Self-Descriptions and there is a Self-Description for all the entities defined as part of the Gaia-X Conceptual model described in the Gaia-X Architecture document: This list mainly comprises:

- Participant including Consumer, Federator, Provider
- Service Offering
- Resource

Based on the three basic labels further Gaia-X Labels can be created to fit new needs, in particular using extension profiles for country and **domain specific requirements**. Extension profiles can also leverage the labelling criteria by adding and defining on-top requirements for particular purposes.

To ensure impact and consistency of Gaia-X Labels, new labels and extensions have to be authorized by the Gaia-X Association. Another subsystem can be added to include some sector specific rules and policies, as is shown in the next figure:

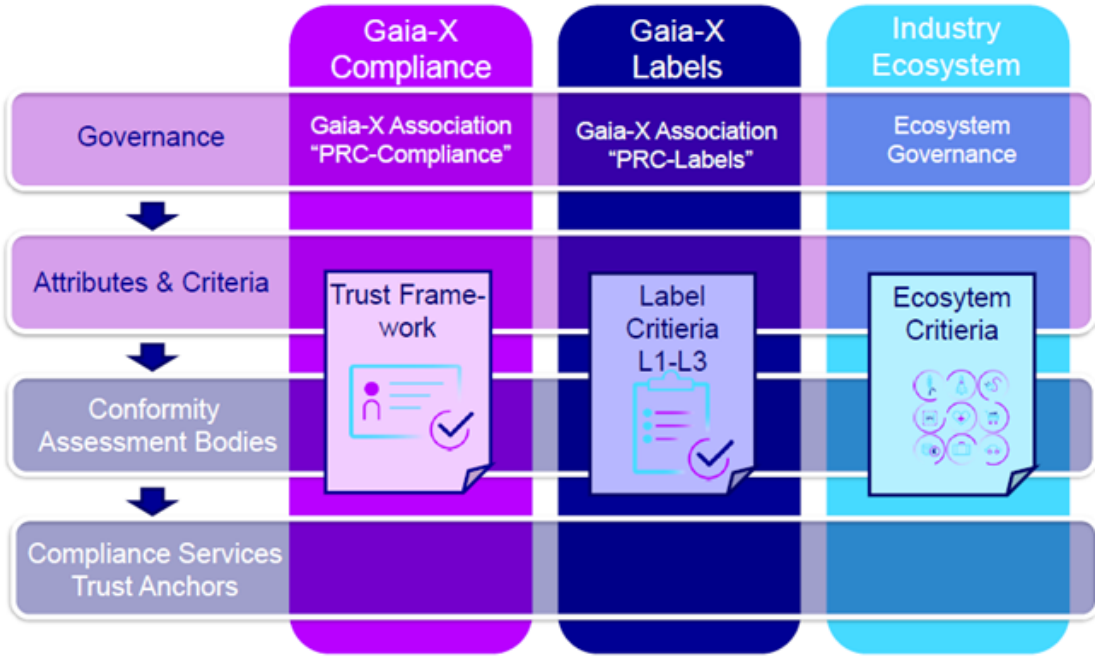

Figure 11 Gaia-X compliance model

Therefore, this model allows the tourism data space to add any specific criteria valid for Tourism stakeholders.

The current version of the **Policy Rules & Label document** contains criteria only for cloud providers. Gaia-X anticipates that additional rules will be defined for the participants in data spaces and data sharing ecosystems.

3.2.1.1 Gaia-X Digital Clearing House Clearing House (GXDHC)

In order to operationalize the onboarding process a specific data space intermediary or infrastructure has been defined in Gaia-X, the so-called Gaia-X Digital Clearing House (GXDHC)

According to Gaia-X[24], the GXDCH is the necessary element to operationalize Gaia-X in the market. The Gaia-X Framework describes functional specifications, technical requirements, and SW assets necessary to be Gaia-X compliant.

The GXDCH are a network of execution nodes for the compliance components that we have developed. This safeguards the distributed, decentralised ways of running the Gaia-X compliance, not operated centrally by the Association, and where anybody can benefit from the open, transparent, and secure federated digital ecosystem – thus making the Gaia-X mission a reality.

It involves a Participant and Trust Anchor Registry, a compliance service and a notarization service.

[24] GXDCH - Gaia-X: A Federated Secure Data Infrastructure

1. The GXDCH is a node of verification of the Gaia-X rules;
2. It is the go-to place to obtain Gaia-X compliance and become part of the Gaia-X ecosystem;
3. The GXDCH are non-exclusive, interchangeable multiple nodes operated by market operators, acting as a Gaia-X Federator;
4. They operate and run services of the Gaia-X Framework (compulsory and optional), necessary to achieve compliance and support the onboarding of any Gaia-X adopter;
5. They integrate to external TA (Trust Anchors), including CAB (Conformance Assessment Bodies) for external asseverations, Identity Verification (like eIDAS), and other TDS (Trusted Data Sources) as defined by the AISBL.

Next figure shows the current set of GXDCH nodes:

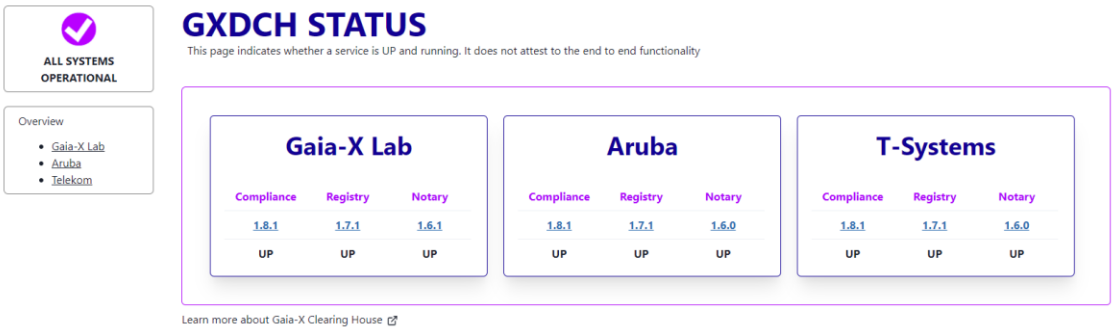

Figure 12 Current GXDCH nodes

3.2.2 Intermediaries/Federators

According to Gaia-X[25] the intermediaries or federators provide enabling services to facilitate the operation of data ecosystems under a specific rule of an ecosystem governance. The intermediaries can be implemented using multiple technologies, products and implementations.

Data product usage in Gaia-X is enabled by a set of Data Exchange Services that are realized by each Participant and can be supported as Federation Services. This is the Gaia-X list of intermediaries in the context of data sharing:

Authentication (mandatory) is essential to connect two Participants. Authentication is provided by the Identities and Trust Framework: Identities provide general information on the Participant, and the Trust Framework appends additional claims, like verified location, or verified application of other standards or regulations.

Policy negotiation and contracting (mandatory) includes the ability to negotiate access and usage policies between two parties. This could be deployed as a sequence between the parties implemented by the control plane of a data space connector. However, an external contracting service can support here, when one or multiple parties do not have the technical abilities for this. ODRL is used to support contracting as (a) it provides interoperability (all parties must be

[25] Enabling and Federation Services - Gaia-X Architecture Document - master Release

able to understand the policies to enforce them later on) and (b) it enables computerised policies enforcement during the transaction.

A **catalogue** (mandatory) provides mechanisms to publish Data Product Descriptions (metadata) and support search or query of the Descriptions. A catalogue may be realized as a centralized or decentralized service, but the capability can also be realized as a distributed functionality.

Vocabularies (optional) provides additional metadata to the Data Product Descriptions. The Descriptions should contain a limited amount of **information** as a common denominator but must be extensible with vocabularies from different (business or technical) domains.

Observability abilities (optional) - logging and audit data - are required to provide an auditable framework for transactions.

Data Exchange protocols are required to exchange data between Participants and enable Data Usage. Data exchanges are realized on a peer-to-peer basis. Gaia-X does not promote any technical protocol - the actual protocol must be agreed between the parties during the contracting phase.

Next, a more detailed information about the main intermediaries is included.

3.2.2.1 Identity provider

Identity plays a critical role in the context of data spaces. Data spaces are designed to enable efficient and effective sharing and management of data across different systems and applications, which requires a secure and reliable system for identifying and authenticating users.

Identification mechanisms are the basis for finding attributes of a participant in a data space and they allow the participant to exert control, to choose which data to share with whom, when and under what conditions. This ensures the participant has agency over its assets, i.e., data sovereignty.

In a data space, identity is used to establish trust between different users and systems, and to ensure that access to data is controlled and secure. For example, users may need to authenticate themselves in order to access certain data or perform certain actions within the data space. This helps to prevent unauthorized access and ensure that data is only accessed by authorized users.

In addition, identity can be used to manage data ownership and attribution within data spaces. For example, a user's identity may be associated with the data they create or contribute to, which can help to ensure that they receive proper credit and recognition for their contributions.

The design of the identity provider is the first decision for the design of the data space. If a central identity provider is chosen to manage the identities for all participants, every other service depends on this central verification, and decentralized designs are no longer fully feasible.

Which mechanism to use to identify participants is the most fundamental design decision. It impacts policies on autonomy and sovereignty as well as technical solution architectures for other components of a data space.

Regarding identity, GAIA-X proposes the use of a decentralized identity model, based on the principles of self-sovereign identity (SSI). This approach enables individuals and organizations to control their own identity information, without relying on centralized authorities or third-party providers.

In the GAIA-X model, individuals and organizations would be able to create and manage their own digital identities, which could be verified by trusted third-party providers. These identities would be based on open standards and would be interoperable across different systems and applications.

The use of decentralized identities would enable individuals and organizations to have greater control over their personal data and how it is shared. It would also help to increase security and reduce the risk of identity theft and other types of fraud.

However, identity is not only used for natural and legal persons, every dataspace component, including data sets, resources, applications, services, etc., can have an identity. Even the data space itself should have an identity since several independent data spaces can be deployed and collaborate with each other.

A Verifiable Credential is *Gaia-X conformant* if the issuer of the Verifiable Credential itself has an identity coming from one of the Trust Anchors.

Here is the list of the current accepted identity trust anchors[26]:

State	The Trust Service Providers (TSP) must be state validated identity issuers or EV SSL issuers. - For participant, if the legalAddress.countryCode is in EEA, valid state identity issuers are **eiDAS** ones. - **Gaia-X** Association may also be a valid TSP for Gaia-X Association members.
eiDAS	Issuers of Qualified Certificate for Electronic Signature as defined in eIDAS Regulation (EU) No 910/2014 (homepage: https://esignature.ec.europa.eu/efda/tl-browser/#/screen/home) (machine: https://ec.europa.eu/tools/lotl/eu-lotl.xml)
EV SSL	Extended Validation (EV) Secure Sockets Layer (SSL) certificate issuers are considered to be temporarily valid Trust Service Providers. (homepage: https://wiki.mozilla.org/CA/Included_Certificates) (machine: https://ccadb-public.secure.force.com/mozilla/IncludedCACertificateReportPEMCSV)
registrationNumberIssuer	During the pilot phase, the Gaia-X Association nominated itself as a valid Trust Anchor under https://notary.gaia-x.eu

[26] Trust Anchors - Gaia-X Trust Framework - main version (fb420580)

3.2.2.2 Catalogue

Discoverability refers to the ability to locate and access relevant data and services within a data space. Discoverability is critical for enabling efficient and effective data sharing and collaboration among users.

To achieve discoverability, data and services must be appropriately labelled, categorized, and indexed to facilitate search and retrieval. This involves the use of metadata, which provides information about the data, such as its format, content, quality, and source. Metadata also enables data and services discovery through the use of search engines and other discovery tools that allow users to browse, query, and filter resources based on different criteria.

In the case of data products, discoverability can also be enhanced through the use of data standards, ontologies, and other data management best practices that promote consistency, interoperability, and reuse of data. These practices help ensure that data can be easily discovered and understood by a wide range of users, regardless of their background or expertise.

The Catalogue is a key component of a data space and plays a critical role in facilitating data discovery, access, and sharing. It is essentially a repository or database that stores information about the data assets within the data space.

Two participants can share data by communicating directly online or offline without the need for a catalogue. But for more participants a catalogue function greatly increases the discoverability of data assets and services. If there is more than one catalogue due to a federated or decentralized design, the catalogue must allow federated searches of data assets in catalogues at multiple sites.

When choosing a target architecture for a data space, the design of the catalogue function can fall somewhere along the spectrum between a central catalogue, multiple federated catalogues, and many decentralized catalogues.

In the case of **IDS**, two different catalogues have been included in the reference architecture: the Meta data broker for data and the App broker for applications to be used within connectors.

GAIA-X is still defining the Federated Catalogue and the Eclipse Dataspaces Components git also includes an EDC Catalogue.

3.2.2.3 Logging service: Data space monitoring and observability

In data spaces with highly regulated data, it is necessary to make the data sharing process observable. This can be done for legal reasons to prove that data has been processed only by authorized entities, or for business reasons to provide a marketplace and billing function through a trusted third party.

Depending on the architecture of the data space, multiple solutions are possible. For a centralized architecture a central observer (clearing house, auditor or monitoring agent) can be implemented. Another option is to define a federated model of observers to distribute the

information, load, and potential for error. To go a step further, a decentralized architecture can minimize the risks associated with a centralized or federated observer model.

In the case of GAIA-X, the contract negotiation can lead to both parties agreeing on a Data Transmission Logging Service which is then used by both sides to log data transmission details. The logs might also include information needed for billing and service level details.

IDSA has included in the reference architecture the "clearing house", which receive and store the logs produced during the data transfer process.

3.2.2.4 Vocabulary hub

The concept of Vocabulary provided was introduce by IDSA to facilitate semantic interoperability to data spaces. Gaia-X still do not have a similar intermediary.

3.2.2.5 Personal data intermediary

The protection of personal data is a critical concern in Europe, and data spaces need to be designed to comply with the EU's General Data Protection Regulation (GDPR) and other data protection laws. Ensuring that data is processed in a way that is secure and respects individual rights is a key issue for data spaces.

3.2.2.6 Service orchestration

GAIA-X scope is not only to provide services and resources but also the possibility to compose and orchestrate different services. For example, in GAIA-X one company could use the cloud infrastructure from one provider to run an algorithm offered by other provider with the data available from a third participant and get the final result. In this scenario a workflow editor and engine should be provided to define and execute the different processes needed. This is a very complex scenario but very useful for a company without the technical skills needed to participate in a data space.

3.2.3 Participants: Participant journey

One of the main challenges we face is to explain to companies, especially those without technical skills, the data space concept and how the "data spaces" paradigm is different from previous data sharing approaches. Furthermore, it must be clearly demonstrated what are the real benefits of using the data space approach and how it can improve data valorisation both inside the company and externally.

The conceptual model and building blocks approach provide a good and detailed overview of the main concepts like trust, data sovereignty, interoperability, business models and data governance but they do not provide a non-technical overview of the data space processes, i.e., the steps a company must follow to participate in a data space either as a data provider or as a consumer. Main concepts are easily understood but the operational level, the way in which they apply to the data sharing processes, is not so clear.

Explaining the data space concept as a "participant journey" is very helpful to understand how the building blocks cooperate to provide the whole experience of sharing data in an interoperable and standardized way. Since this deliverable is related to technical issues, the difference among Data producer/Data owner and data provider is not taken into account.

The main steps of this journey, that can be considered as the main phases of the data space engagement life cycle, are the following ones:

- **Onboarding**
  - Participant
  - Data product/service
- **Publish/search/purchase**
  - Data
  - Services
- **Data products/services**
  - Provider (offers the data/service through the catalogue)
  - Consumer (accesses the data/service from the catalogue)
  - Owner (the holder of the rights to access and use the data/service)
- **Monitoring**
  - Tracking the data exchanges and contract fulfilment

3.2.3.1 On-boarding

This section describes the steps a participant should follow to onboard in a Gaia-X compliant dataspace. Although the technical details are Gaia-X specific, the general steps can be applicable to any other data space following the Self Sovereignty Identity approach.

The **first step** is selecting a clearinghouse to formalize the onboarding. Next steps can be done with the wizard developed by Gaia-X: wizard.lab.gaia-x.eu/userGuide

The **second step** is to define the participant entity and follow the compliance process for the participant.

If the participant is a data provider, the data service offerings must be defined and included in the compliance process. However, these steps are only a small part of the overall journey. Companies need to understand that the journey begins long before they begin the process of engaging in a data space. The journey starts when the company begins to manage data like any other company asset.

Most companies do not manage their data in the same way, there is no comprehensive management of data including the processes needed to maintain data clean, up to date and

ready to be shared. Even the definition of "quality" data is not clear. Unfortunately, most data space initiatives do not take into consideration this first step.

From a technical point of view, if a company wants to participate as a data provider, it is necessary to **create a data product and install and configure the available connectors to provide the data service**.

A generic process compliance process includes the following steps:

1. Generate a DID (Decentralized Identifier) for each of the entities participating in the data space.
2. For each of the entities create a self-description according to the format defined by the data space governance authority.
3. Call the compliance service of the clearing house to get the signed verifiable credentials.

This is the current onboarding process in Gaia-X:

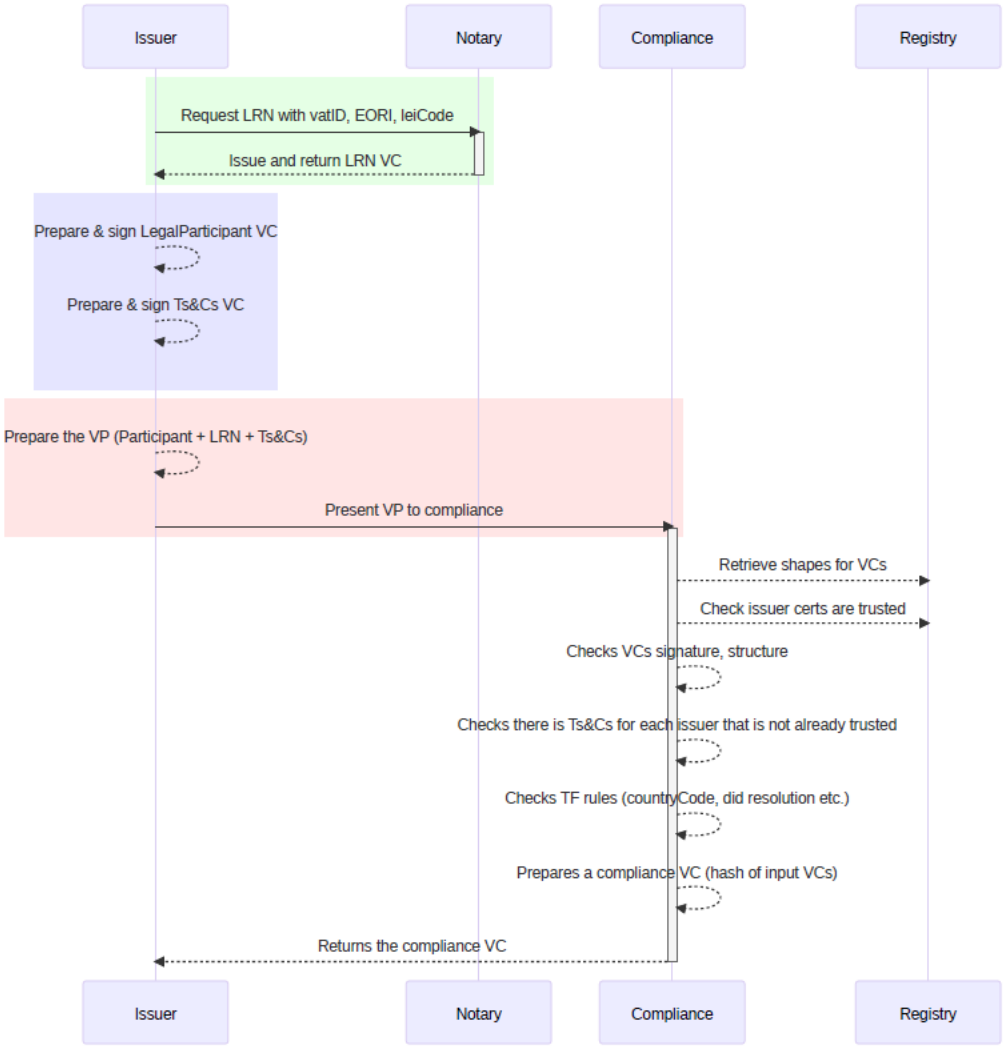

Figure 13 Gaia-X onboarding workflow

3.2.3.2 Publish/search/purchase:

Once the on-boarding process is finished the participant can start publishing or finding services using a data space catalogue and participating in data transfer processes.

3.2.3.3 Data product/services

Once a participant has joined a data space and discovered available data contract offers, the mechanism of data sharing is initiated. Data sharing is the core activity to enable further data processing and value generation by using the data.

Before the actual data transfer starts, two previous steps are mandatory:

1. **Authentication:** The identities and Trust Framework are essential. Without this, you cannot connect two Participants. Identities provide general information on the Participant, and the Trust Framework appends additional claims, like verified location, or verified application of other standards or regulations.
2. **Policy negotiation and contracting** include the ability to negotiate access and usage policies between 2 parties.

Data sharing must accommodate a wide range of scenarios. From a simple file transfer between two storage providers, to API access for streaming or eventing, to quite complex implementations with secure execution environments through confidential compute enclaves, environment attestations, signed code, custom encryption algorithms, and more. The most appropriate solution depends on the data protection needs and the trust level between the participants.

3.3 Data space related (semantic) models

We will first introduce a conceptual model for data spaces, based on the existing data space architectures such as IDSA and GAIA-X. Then, we will briefly define “data product” and explain different kinds of data that can be relevant in the context of tourism. Finally, we will give an overview of tourism related (semantic) models, including some prominent ontologies.

3.3.1 Data space conceptual model

The Data Space Support centre (DSSC) provides a comprehensive blueprint for data spaces including a conceptual model that contains different components (technical and organisational) of data spaces and the relationship between them. This blueprint nicely consolidates the definitions made by different data space initiatives such as GAIA-X and IDSA. In the following, we will briefly summarise these definitions and provide examples of each component in the context of tourism, when applicable. An overview of the conceptual model according to DSSC blueprint can be seen in Figure 14.

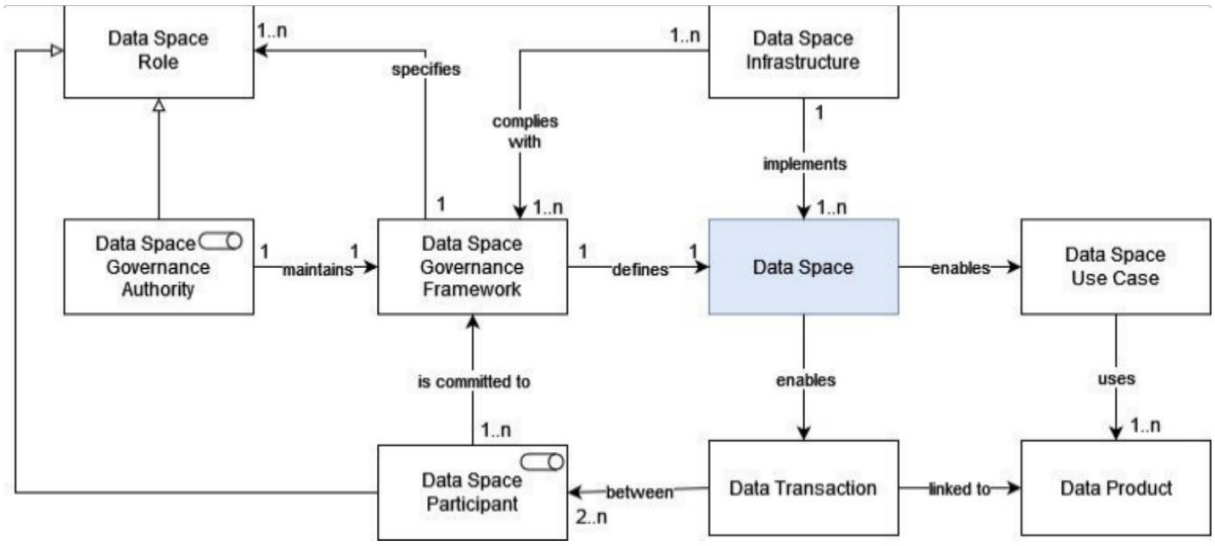

Figure 14 An overview of the conceptual model (level 1) of data spaces from the DSSC

Data space governance framework is defined by DSSC as "the set of principles, standards, policies (rules/regulations), agreements and practices that apply to the governance, management, and operations (including business and technology aspects) of a data space as well as to the enforcement thereof, and the resolution of any conflicts". This definition is quite generic and can apply to the tourism domain in a broad sense. For example, the high-level communication protocols between data space participants to exchange data products, minimum metadata the self-descriptions should contain (e.g., licensing information) can be part of this framework. Since the tourism domain may involve sensitive personal data, how to deal with such data according to regulations is also part of this framework.

Data space use case: The guidance for creating the governance framework typically comes from use cases. Use cases are scenarios where a group of stakeholders create some value (e.g., business, societal) from sharing one or more data products. Some examples for use case data types that may be relevant for sharing in tourism data space, but will require a common European semantic interoperability standard, are listed below:

- **E-bike charging stations**: Recommendation for using the same semantic data model for any public information that is made available for e-bike charging stations, no matter where they are located and which provider they are operated by (e.g. charging stations at alpine huts and in urban areas)

- **Tourism bus routes**: Routes that operate in addition to public transport but are not yet available in public transport timetable data (e.g. could be shared via the Mobility Data Space)

- **Protected areas**: Rules and regulations to support sustainable tourism behaviour (e.g. Digitize the Planet e.V.)

- **Mountain resorts**: Ski lifts and runs (operational status open/closed); It should be possible to do intermodal routing from home address to the mountain station of a

mountain railway, using train, bus, walking, ropeway, so the data should be interoperable with public transit systems (e.g. AlpineBits Alliance https://www.alpinebits.org/destinationdata/)

- **Current conditions and closures**: Information about track closures and current condition on trails (e.g. snow and ice)
- **Sustainability data:** Information about transparent and measurable key performance indicators that cover economic, social and environmental aspects of sustainability for accommodations, tour operators and tourism destinations (e.g. Tourism Impact Alliance)
- **Events:** Information about events that are interesting for tourists. (e.g. AlpineBits Alliance https://www.alpinebits.org/destinationdata/)

Data space role: One thing a data space governance framework defines is a set of data space roles. There can be many roles defined in a data space, but it should at least define two roles, according to DSSC:

Data space governance authority: This role can be taken by one or more stakeholders to create, maintain and enforce the data space governance framework. In tourism data spaces, this could be regional or national tourism boards, special department in ministries and so on.

Data space participant: This role is taken by any stakeholder that interacts with the data space and follows the framework defined by the data space governance authority. In the context of tourism, the participants can be from different verticals, for example event organisers and aggregators, accommodation providers, regional marketing organisations and many other long-tail product and service providers.

Data space infrastructure: The governance framework not only defines the regulations of a data space, but also specifies the requirements with which organisational, technical and data-related components should comply. These components such as connectors, catalogue and identity services and metadata about various data assets comprise the data space infrastructure. The data space infrastructure implements the data space.

Data transactions: According to the DSSC blueprint, a data transaction is a data sharing activity among involved participants in a data space. It involves technical, legal, financial and organisational arrangements made for data sharing between the participants.

Data product: According to the generic definition of DSSC, a data product is “a standardised data unit packaging relevant data resources and services into a consumable form complying with data product specifications”. This definition is also valid for tourism data spaces, however, there are some nuances regarding nature of the data, which we elaborate in section 0 3.3.2 Data **product**. The data product specifications are defined as part of the data governance framework and may include aspects like vocabularies used for self-descriptions. Since tourism is a highly fragmented domain, it is important to use machine understandable self-descriptions that use (de facto) standardised vocabularies (e.g. DCAT, schema.org, PROV-O) to address any interoperability and scalability issues.

Data space: A data space conceptually consists of the components mentioned above and enables data sharing between participants to create value.

3.3.2 Data product

In the context of tourism data spaces, a data product can be seen as a representation of any product and service that can be exchanged between the participants of data spaces. In different data space specifications, a data product can have names like asset, artefact or resource. As a highly fragmented and heterogenous sector, tourism domain may deal with three kinds of data:

Static data: The data that do not change too often, for example address information of touristic establishments.

Dynamic data: Data with high update frequency. For example, prices of hotel rooms, plane tickets, traffic and weather data. This data is typically provided via web services.

Active data: Data that changes the state of a system. Tourism domain makes intensive use of active data, for example, web services that enable the booking of services like hotel rooms or event tickets would be an example to this. In this case the data product can be the function, behavioural and non-functional descriptions of web services.

One of the most important features of dataspaces is that the data products can be in different formats, structure and level of formality. A data product may be a piece of text, a bundle of CSV or JSON files, or semantic data products like knowledge graphs, linked datasets, or the (semantic) descriptions of web services. For semantic data products, there are many tourism-specific ontologies which we elaborate in Section .

3.3.3 Tourism specific models

In order to define a Tourism specific data model to be used in the data space several strategies can be used

- **Incremental:** A domain model, for example an ontology is a living artifact. The development should be guided by the domain expertise and conducted incrementally. This is aligned with the modern engineering approaches in many disciplines, such as software and knowledge engineering, where the products developed in iterations and updated as the use cases require.
- **Pluralistic and flexible:** A single data model for the tourism domain is not realistic. In the best case scenario, there are discrepancies between the points of view of different countries and regions about how touristic products and services should be modelled. Therefore a pluralistic approach must be supported, where different models can be aligned and harmonized within or across dataspaces. To enable this pluralistic approach, the models must be flexible. Flexibility implies how easy it is to extend a model with more specific terms and how easy it is to align it with other models. As a rule of thumb, models should describe the domain from their perspective as wide as possible while being not more restrictive than necessary. For example, a model can

define a type called Event, and leave more specific types to different stakeholders as they may differ across regions.

- **Holistic:** At present, there is no single common European ontology in tourism that has been universally adopted and standardized across all European countries. Nevertheless, serious efforts have been made to develop ontologies and standards that can be used as common reference points for tourism-related data and interoperability within Europe.

3.3.3.1 Examples of current Tourism Ontologies

Notable initiatives have been carried out by organizations like the World Tourism Organization (UNWTO) to promote interoperability and data sharing in the tourism sector. For example, they promoted the multilingual "Thesaurus on Tourism & Leisure Activities". This thesaurus can be used as a guide to tourism terminology, as well as for the standardization and normalization of a common indexation and research language, at an international level.

Furthermore, many research projects, mainly at universities, have worked on the creation of ontologies for destination management. They are very interesting works, but, in practice, they have the deficit of their limited use. Based on this circumstance, it is crucial that some entity brings together public interests (eg. for public DMOs) and private interests (eg. for tourism SMEs) and assumes the responsibility to promote and to maintain this common ontology in tourism.

Fortunately, initiatives of this type are emerging in different countries in Europe. The initial approach and conceptualization of these ontologies is different, which makes it difficult to provide interoperability among them, but at least, it is a necessary effort to try to build an initial semantic model.

These are some examples of these ontologies:

ODTA Tourism domain specifications (odta.sti2.org)

Open Data Tourism Alliance (ODTA) is an association of tourism marketing organisations from the DACH region (Germany, Austria and Switzerland) that aims to create standardised schemas for semantically describing tourism related data and services. The schemas are based on schema.org, more specifically they are an extended subset of schema.org. From a syntactical point of view, the schemas are represented as SHACL shapes, which enables existing tools in the semantic web stack to work with them seamlessly. Currently there are schemas already published for Events, POIs, Accommodations and Trails, with new ones coming in 2024. The ODTA schemas can be used as a basis for creating semantic data products to address interoperability issues.

Digital Tourism Hub in Italy (www.italiadomani.gov.it/en/Interventi/investimenti/hub-del-turismo-digitale.html)

As part of the National Recovery and Resilience Plan (PNRR), the Ministry of Tourism has launched the Tourism Digital Hub (TDH) Program, with the aim of creating a digital tourism

hub, accessible through a dedicated multi-channel web platform, which, like a virtual bridge, allows the meeting between the needs of tourists and the vast Italian tourism ecosystem.

This tool, through the exploitation of data, enables the communication of digital supply and demand by offering national and international tourist services and interests tailored to different types of users in order to improve their travel experience.

The collection and aggregation of data is enabled by the interoperability platform realized within the framework of the Program, called TDH022, which enables the exchange of contents and the integration of services provided by public and private Partners through API calls. The TDH022 technology platform allows public and private partners, subject to agreements with the Ministry, to interoperate with the Tourism Digital Hub to exchange data and services.

The TDH022 Ecosystem Guidelines, identify digital standards and technologies at a national level, and guide the exchange of data and content between participants, acting as an integration interface between sector operators who wish to be part of the digital ecosystem and who are operational in Italy.

Ontology DATAtourisme in France (www.datatourisme.fr)

DATAtourisme is a research and development project led by ADN Tourisme, the national federation of institutional tourism organisations. It aims to aggregate, standardize and facilitate the dissemination of institutional tourism data in open data. This project is, by nature, collaborative and evolutionary. Consequently, this version of the ontology and its documentation will evolve regularly, enriched by the suggestions of data producers and re-users.

The central concept of the ontology is the term ":PointOfInterest". It is defined as any tourist element that deserves to be described and valued. A POI (Point of Interest) is a tourist element which is managed by an Agent and which can be consumed via Products and Services. It is the minimal class to instantiate for a product to be managed in the DATAtourisme information system.

A POI can be divided into 4 different subtypes:

- ":Product": a tourist object that can be consumed (e.g. a hotel room, an activity, a guided tour, etc.).
- ":Tour": a tourist route is a POI that proposes an itinerary made up of stops forming a route.
- ":EntertainmentAndEvent": events, festivals, exhibitions, or any other event with a beginning and an end.
- ":PlaceOfInterest": a place of tourist interest (e.g. a natural site, a cultural site, a village, a restaurant, etc.).

Segittur Tourism Conceptual Reference Model in Spain (www.segittur.es/transformacion-digital/proyectos-transformacion-digital/modelo-conceptual-de-referencia-para-el-desarrollo-de-una-red-de-ontologias-del-sector-turistico/)

The SEGITTUR Tourism Semantic Conceptual Reference Model provides definitions and a formal structure to:

- describe the implicit and explicit concepts and relationships used in the tourism domain and the sub-domains defined in the scope,
- promote a shared understanding of the information on this domain by providing a common and extensible semantic framework to which any information on tourism can conform.

It is intended to be a common language for subject matter experts and implementers to formulate requirements for information systems and to serve as a guide for good practice in conceptual modelling. In this way, it can provide the "semantic glue" needed to mediate between different sources of tourism information, such as that published by different tourism stakeholders.

This model is based on the UNE 178503:2019 Standard on 'Smart Tourist Destinations. Semantics applied to tourism' and previous research on the state of the art of other semantic reference models related to tourism.

The goal of the project is to have a representation of the "tourism" knowledge domain that includes the most important entities that make it up, as well as the relationships between them, and that serves to model the major domains or subdomains of the system:

- **Subdomain 1: Supply**. This subdomain comprises the providers of products and services that build tourism activity. These elements can be found both in the destination and outside it (permanent and non-permanent elements). The representation of this domain acts as a digital reflection of the analogue world, contemplating entities such as: accommodation, travel agencies, tour guides, tour operators, car rental companies, catering establishments, events, etc.
- **Subdomain 2: Tourist**. This subdomain refers only to the user, placing them at the heart of the activity. They have different travel and consumption motivations and, therefore, the offer must respond to these expectations and needs. This subdomain encompasses entities that are transversal to the tourist's travel cycle and interacts closely with the supply subdomain.
- **Subdomain 3: Destination**. This subdomain refers to the sum of public spaces and services involved in tourism activity and are generally provided by the local administration. The representation of entities in this sub-domain considers that the objective of destinations is to attract visitors, offer them a satisfactory experience and ensure their loyalty. This domain encompasses the sub-domains of supply and tourist. This subdomain acts as the context in which tourism interaction takes place.

The Conceptual Tourism Reference Model comprises 305 classes in the defined subdomains (supply, tourist, destination), 433 properties and 103 lists of terms.

SEGITTUR is strongly involved in the creation of a Smart Destinations Platform, a project through which Spain aims to place itself at the forefront of tourism countries by using digitalization to improve the tourism experience and offer integrated services of value at the destination.

The aim of the platform is to integrate, collate and combine public and private data to build more competitive intelligence on the ecosystems of Spanish destinations. The platform will promote

continuous innovation by interconnecting and meeting the needs of all agents (tourists, destinations and companies) along the tourism value chain.

This platform will integrate this Ontology into its systems, so that the provision of common services from the central node of the platform to the tourist destinations will allow the construction of interoperable digital layers of tourism information between the different agents of the ecosystem.

4 DATA SPACE CONNECTORS

Data space connectors facilitate and orchestrate the sharing of data assets, while enforcing requirements set by the data provider. A connector includes policies, configuration and other metadata artifacts that can run on any cloud infrastructure, on premises or on an edge device.

The existing connectors can be classified under two main categories:

- Connectors based on the previous IDS specifications.
- Connectors following the new IDS specifications: Dataspace Protocol v0.8 - IDS Knowledge Base (internationaldataspaces.org)

These are the main characteristics of the new protocol:

- The new specification relies in well-known standards to define the data product and data usage policies: **DCAT** (Version 3.0) and usage control expressed as **ODRL** Policies.
- Decoupling of control and data planes in data transfer technologies, making it possible to use any transfer protocol or technology available.

There are two main connectors following the principles of the new IDS protocol: the Eclipse Data Space Connector and the new FIWARE Data Space Connector.

4.1 IDS Data Space Connector (previous specification)

The available connectors based on the previous IDS specifications are in general more mature and stable than the connectors implementing the new protocol. The main drawbacks of this connectors:

- Only https-based data transfer protocols with maximum size of data
- Federated authorization schema (No SSI)
- Deprecated specifications
- Coupling of control and data plane

Furthermore, the new IDSs data space protocol is being adopted by the main data spaces initiatives, Gaia-X, FIWARE and DSBA.

Therefore, our recommendation is to use one of the new IDS Dataspace Protocol connectors. Next section provides some information about the current status of the two available connectors.

4.2 New IDS Dataspace Protocol connectors

4.2.1 Eclipse Data Space Components (including the connector)

The Eclipse Dataspace Components (EDC) is a comprehensive framework (concept, architecture, code, samples) providing a basic set of features (functional and non-functional) that dataspace implementations can re-use and customize by leveraging the framework's defined APIs and ensure interoperability by design. It is powered by the specifications of the **Gaia-X AISBL Trust Framework** and the **IDSA Dataspace protocol**.

Several initiatives implementing dataspaces are using the Eclipse Data Space connector, including EONA-X, CATENA-X. Furthermore, this connector has been selected by several Gaia-X proof of concepts.

Recently Gaia-X announced the plans for integrating Gaia-X related features with EDC (Eclipse Data Components) to simplify the utilization of Gaia-X Verifiable Credentials for Participant Compliance in contract negotiations and access control within this ecosystem. This integration aims to enhance the accessibility of Gaia-X Verifiable Credentials for participants, enabling service providers to exclusively grant access to their services to Gaia-X compliant participants. In simpler terms, this means that service providers can limit access to their services solely to Gaia-X compliant participants.

4.2.2 FIWARE Data Space Connector

Recently FIWARE has announced the development of a new Data space connector that follows the principles, requirements and specifications included in the second version of the DSBA convergence document. The information about this new connector is based solely in the public documentation provided by FIWARE since it has been no time to test it.

The FIWARE Data Space Connector is an integrated suite of components every organization participating in a data space should deploy to "connect" to a data space. Following the DSBA recommendations, it allows to:

- Interface with Trust Services aligned with EBSI specifications
- Implement authentication based on W3C DID with VC/VP standards and SIOPv2/OIDC4VP protocols
- Implement authorization based on attribute-based access control (ABAC) following an XACML P*P architecture
- Provide compatibility with ETSI NGSI-LD as data exchange API
- Supports the TMForum APIs for contract negotiation

5 SPECIFIC ISSUES OF THE "TOURISM" DATA SPACES

This section includes some specific issues and recommendations that should been taking into account in the process of creating the ETDS.

- **Personal data management**: People are at the centre of the tourism data space, so privacy protection and compliance with data protection legislation and initiatives play a very important role.
- **SMEs:** Most companies in the tourism sector are SMEs. Both the process to onboard in a data space and the technology required to participate in it are too complex and very far from their usual business. They have neither the technical nor operational skills to handle the complexities of a data space. Some approaches, both in the context of IDS and Gaia-X, could be very useful for SMEs, like **Connector-as-a-Service** and more **Tourism Data Spaces**.
- **Public and private actors**: Tourism sector stakeholders include both public administrations and private companies with different strategies and objectives regarding data sharing. Open data public administrations' strategy should be aligned with data protection, value generation and monetisation strategies of private companies.
- **(STRONG) Relation with other data spaces:** The tourism sector is directly related to other sectors such as transport, mobility, environment, energy, cultural heritage and construction. In this way, data from these sectors should also feed the ecosystem of the ETDS. Inter and intra data space interoperability is a critical issue for tourism data spaces.

- **Geographic data**: Almost every tourism related data has a geographic scope, that can be a region, a specific geographic address, or a set of addresses forming a line or a path. Also vector and raster data are used. The geographic data can be organised in layer of different types that facilitates geographic based reasoning and analysis. Geographic data uses some specific models and standards.

- **Local vs regional vs national vs international scope**. The geographic scope mentioned before might also induce a tendency to set up local, regional and/or national tourism data spaces. The implementation of such should be thoroughly controlled, in order to avoid a too fine-grained substructure of the European tourism data ecosystem on one hand, which will inevitably challenge interoperability, and might reduce the cross-border data visibility. On the other hand, a certain degree of decentralisation facilitates resilience by avoiding single points of failure. It may be a task for EDIB and the future EDIC for Tourism to define the right equilibrium of (de)centralisation taking into account both strategic and technical arguments.

5.1 ETDS-specific requirements technical challenges: Personal data

The importance of personal data to data sharing

The EU Data Strategy's success relies heavily on sharing personal data, which is expected to drive adoption across various sectors such as health, administration, education, and mobility, and tourism.

GDPR compliant personal data sharing will also help ensure that the EU's values, principles, and regulations are fully implemented. As the EU digital single market is grounded in human-centricity, it aims to give back value to citizens by creating decentralized alternatives to the Big Tech platforms and their lock-in effect. By prioritizing the needs and interests of citizens, the EU can foster a more transparent and equitable digital landscape that benefits everyone.

For the tourism sector

For the tourism sector tourists will need to be able to share data across multiple organizations from different sub sectors like mobility, hospitality, and tourism activities, through a seamless user experience (tourist identity tools), while protecting their privacy (GDPR consent).

The challenges of personal data sharing

Human-centricity represents a paradigm shift in how we think about managing data and its potential. It stands in stark contrast to the prevalent "organization-centric" approach by placing the focus on the individuals involved in generating the data, rather than the organization responsible for capturing it, such as a company or government agency. Human-centricity encompasses concepts such as (self-)sovereignty, self-determination, self-governance, autonomy, and agency, which derive from the notion of human rights. At its core, a human-centric approach acknowledges that individuals have the right to determine, without coercion or compulsion, what happens to their personal data.

But human-centricity also comes with several challenges: technical issues (identity management, standardization, user experience, etc.), business issues (costs, IP strategy, etc.), legal issues (compliance with GDPR, DGA, etc.), as well as psychological factors like trust and digital resignation. Given this complexity, few existing data spaces today are actually processing personal data, despite personal data-sharing being a priority for most of them.

Addressing these challenges will require collaboration between stakeholders from a variety of levels and domains, including technical experts, legal professionals, business leaders, and psychologists. Only through such cross-disciplinary efforts can we hope to realize the full potential of a human-centric approach to data management. Given the coming widespread use of AI models in data spaces for purposes such as recommendation or personalization, it is crucial that individuals have effective means to maintain control over their personal data, which may be stored and processed by multiple organizations.

In most data spaces (e.g., mobility, agriculture), the data are generated by industrial applications. Instead, in the ETDS the tourist takes centre stage as the primary generator of data. Hence, the

ETDS is distinctively human-centric. This makes the ETDS akin to health or education data spaces, where individuals (patients or students) play a leading role in generating data through their activities and interactions.

Given this human-centric nature, it becomes essential to adapt how data are accessed and processed. Tourists generate information by consuming services offered by private companies and public authorities, making data sovereignty and privacy paramount considerations. The design of the ETDS takes inspiration from the broader global community of data spaces, but it is uniquely tailored to address the specific needs and challenges that arise from managing human-generated data in the context of tourism.

By emphasizing privacy, data sovereignty, and the centrality of the tourist in data generation, the ETDS aims to create a secure and trusted environment for data sharing and collaboration. This approach ensures that data are used responsibly and ethically, empowering tourists while also fostering innovation and providing valuable insights for stakeholders in the tourism ecosystem and other cross-sectorial data spaces as well.

Personal data sharing and regulation

The Data Governance Act (DGA) introduces the concept of data intermediaries tools that will allow the sharing of data within a data space, and that need to be notified to competent authorities. A subset of the data intermediaries will enable the management of personal data: the *personal data intermediaries (PDI)*. Beyond facilitating personal data-sharing, the personal data intermediaries will also provide data subjects with standard mechanisms to protect data subjects' privacy and rights (GDPR): right to be informed, right of access, right to rectification, right to erasure, right to restrict processing, right to data portability, right to object.

Tools for sharing personal data

The issue of personal data sharing is particularly complex due to specific concerns around privacy and GDPR compliance, and also given the fact that the individual is, in this case, the central point of data integration, not the organizations controlling data. The individual can potentially interact with data spaces of all locations and all sectors. Since he/she does not bear technical capabilities in itself the individual has 2 options:

- the individual authenticates to organizations controlling his/her data separately, and gives consent separately each time, which can be cumbersome.
- the individual uses dedicated tools, personal data intermediaries (PDI) for managing his/her identity and data, that aggregate data and simplify identity and consent management.

There exist different approaches/paradigms to design PDI tools in Europe, here are some of the major ones:

1. **MyData operators** that are tools for GDPR consent management and personal data stores respecting the MyData declaration and label. MyData: MyData is a prominent movement advocating for human-centricity. Originally emerging from open data

activism in Finland, it has since expanded into an international movement that is now run by "MyData Global," a non-profit organization. MyData provides guiding principles aimed at giving individuals greater control over the data trails they leave behind in their everyday activities. The goal is to enable individuals to see what happens with their personal data, specify who can use it, and modify those decisions over time. MyData is not solely focused on individuals' perspectives but also aims to serve commercial interests by promoting business opportunities for personal data. The MyData Principles strive to make privacy, data security, and data minimization standard practices in application design. The movement also seeks to empower individuals to understand privacy policies and to give, deny, or revoke their consent to share data based on a clear understanding of why, how, and for how long their data will be used. The Declaration of MyData Principles outlines ethical principles for personal data management and has been endorsed by over a thousand organizations and individuals worldwide.

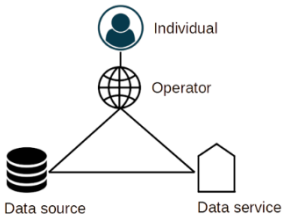

Figure 15 MyData operator concept

2. **SOLID Pods** that are tools following the SOLID specification/protocol and that let people store their data securely in decentralized data stores. The Social Linked Data project (SOLID) is a web decentralization initiative led by Sir Tim Berners-Lee, the inventor of the World Wide Web, and developed collaboratively in an open-source project consisting of multiple commercial and independent contributors. It aims at realizing Tim Berners-Lee's original vision for the Web as a medium for the secure and decentralized exchange of data. SOLID is at its core a specification/protocol that lets people store their data securely in decentralized data stores called pods (personal online data stores) (Buyle et al. 2020). Apart from a focus on separation between the application, identity provider and data storage as three interconnected entities, Solid has a strong focus on machine-readable linked data to ensure interoperability between different applications that reuse the same data source. The Linked Data architecture also helps to create (quite literally) links between one data set and another, including links between personal data and public data from various semantically related data spaces.

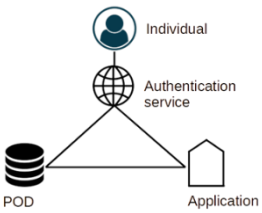

Figure 16 Solid POD concept

3. **POTENTIAL EU Wallets** allowing all European citizens to store and manage personal data in an ecosystem of standard wallets. POTENTIAL unites 148 participants from 19 EU member states from Northern, Western, Eastern, Central, and Southern Europe – representing more than 70% of the European population – and Ukraine. POTENTIAL's pilots drive European digitalization and ease numerous administrative as well as tedious identification processes in everyday situations. The aim is to vividly illustrate the possibilities, functionalities, and added value of a European Digital Identity Wallet. By involving relevant market players, POTENTIAL quickly scales solutions which build on existing market-relevant national solutions.

Within a paradigm PDI must be interoperable (the user can switch from one to another), discussions about cross-paradigms interoperability are still at a very early stage. The two main topics to address regarding interoperability for personal data in the data spaces will be about federating the individual's identity and consent. Some standardization discussions are happening now in organizations like Decentralized Identity Foundation, Kantara (consent receipt standards that inspired ISO), and MyData.

At this stage, Gaia-X, IDSA, FIWARE and other related data space support organization, all explore IAA (Identity, authentication & authorization) capabilities, but they do not consider personal data apart and to not address the question in depth.

Focus on consent

Explicit consent is a cornerstone of personal data sharing, as it is the lawful basis (GDPR) of personal transfers between independent data controllers in a wide range of use-cases. When it comes to data sharing, GDPR requires that individuals are informed about the data being shared, the purpose/finality for which it will be used, and the recipients of the data. Additionally, the individual must provide their explicit consent for their data to be shared for that specific purpose. Explicit consent means that the individual must take a clear and affirmative action, such as checking a box or signing a form, to indicate their consent. Consent must be freely given, meaning that individuals cannot be forced or coerced into giving their consent, and they must be informed of their right to withdraw their consent at any time. It's important to note that GDPR applies to any organization that collects or processes personal data of individuals residing in the EU, regardless of whether the organization is based in the EU or not. Failing to obtain explicit consent for data sharing can result in significant fines and other legal consequences. The major hurdle for consent management within the data spaces landscape now is the variety of consent formats and the lack of interoperability. In the tourism context, we can suppose that a same individual will share data across borders, which will be complex if consents are incompatible.

5.2 ETDS-specific requirements technical challenges: SMEs

For most SMEs the process to on board in a data space and the technology needed to participate are too complex and very far from their usual business. They do not have the technical nor operational skills to deal with data spaces.

Some approaches both in the IDS and Gaia-X context could be very useful for SMEs: **Connector as a service** and the more ambitious **Data Space as a service**. The objective of both approaches can be summarized by the slogan used recently by Sovity[27]: "Setting up data space technology in minutes instead of months".

The strategy is to provide the SMEs with consultancy services and the software needed to share (both provide and consume) data. The software includes a user interface that hides the technical and operational complexity of data spaces showing the final user a high-level overview of the available data, the data contracts and the data transfer processes taking place, facilitating the connection with the internal applications used to produce or consume data.

Next Figure 17 shows an example of the Eclipse Data Space components data dashboard, an example dev frontend application for EDC Data Management API.

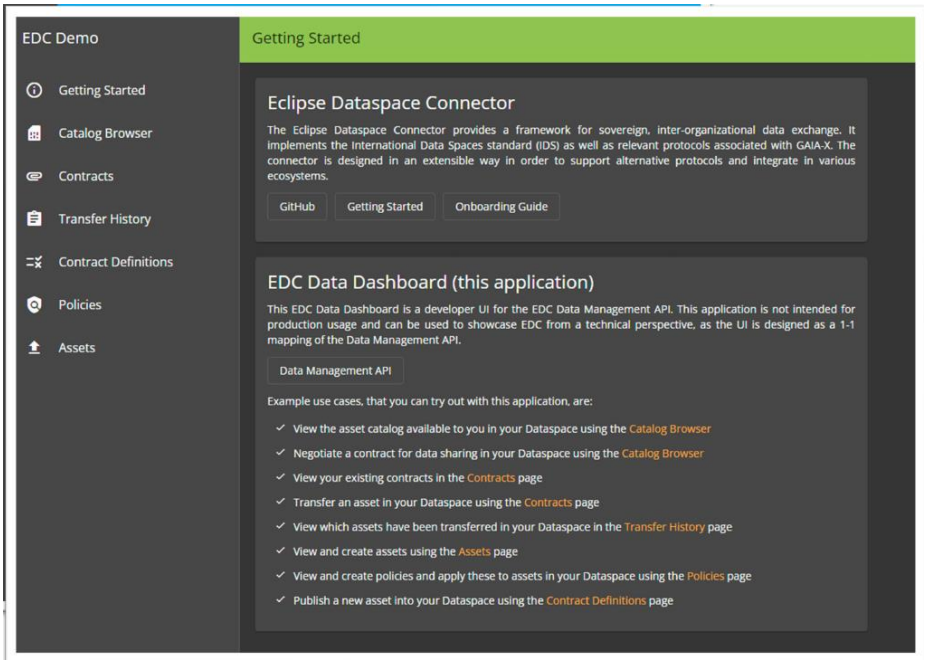

Figure 17 Eclipse Data Space components data dashboard

To better demonstrate the possibilities of dataspaces and their management, the Eclipse Dataspace components project has created under the minimum viable dataspace the so-called **Vision Demonstrator** to showcase a possible user interface that would enable end-to-end interaction - all the way from joining a dataspace to being able to publish a new data asset for others to consume.

This vision was based on 7 scenarios essential to managing a dataspace:

1. Manage my Dataspaces
2. Discover Data Shared by Others

[27] https://sovity.de/about/

3. Negotiate a Data Contract
4. Create a new Policy
5. Create a new Data Asset
6. Create a Data Contract
7. Review existing Data Contract and Manage Notifications

Figure 18 shows the main window of the vision´s mock-up.

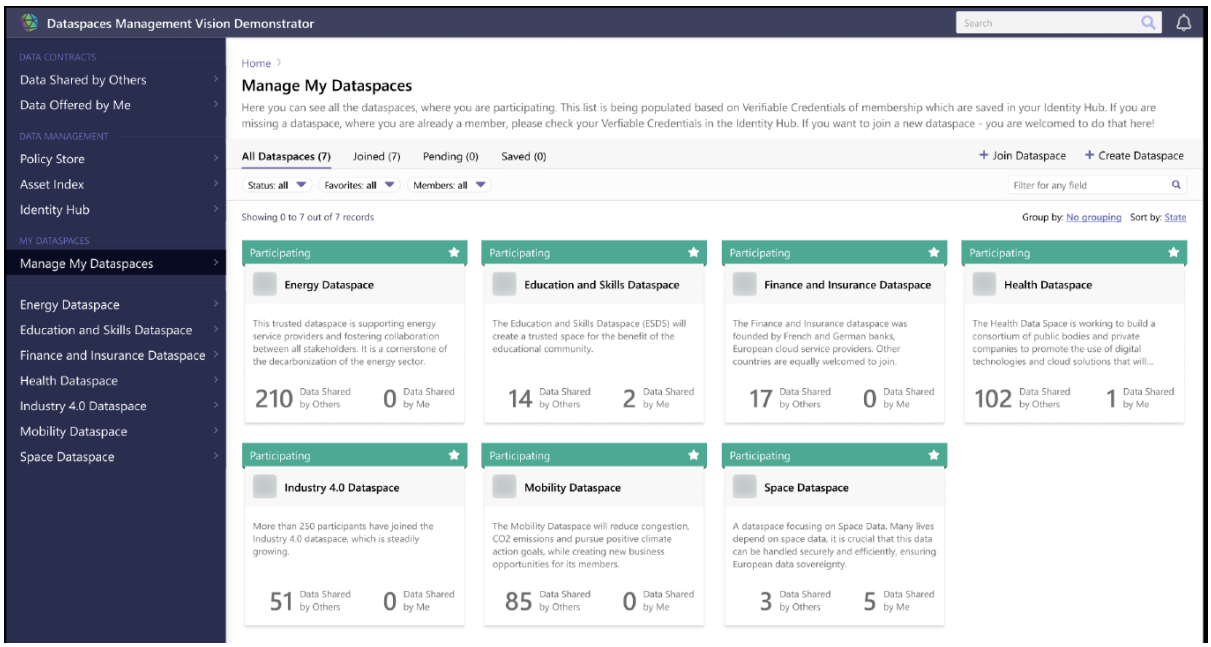

Figure 18 Data Space Vision mock-up

Gaia-X and SIMPL go a step forward offering cloud infrastructure needed to run the software, data products that encapsulate data sets or access to data APIs, applications and services providing data-based applications or algorithms and the orchestration functionality needed to define specific services composition and workflows.

The Data space as a service concept includes all the steps needed to participate in a Data Space.

- Participant on boarding
- Compliance and certification (data and services)
- Catalogue registration
- Data sharing functionalities and dashboard
    - Search for data products/infrastructure/applications or services
    - Use or provide data products/infrastructure/applications or services
    - Monitor data spaces

The adoption of this kind of technology by SMEs depends highly on the perceived added value of data management, both internal data and external sources participating in a data space.

5.3 Tourism specific requirements technical challenges: Public and private actors

Tourism sector stakeholders include both public administrations and private companies with different strategies and objectives regarding data sharing.

Public administrations' concern is about improving tourism destinations, improving users experience both tourists and people living in the destination and assuring tourism sustainability in the long term both from the economic and environmental points of view.

Regarding data, the main objective of public administrations is to make data as open and available as possible, so open data is the main approach. An example of this strategy is the recently approved EU Commission Implementing Act on High-Value Datasets.

The Regulation is set up under the Open Data Directive, which defines six categories of such high-value datasets: geospatial, earth observation and environment, meteorological, statistics, companies, and mobility. The datasets will be available in machine-readable format, via an Application Programming Interface and, where relevant, as bulk download. Some of this high-value datasets are very relevant for the tourism sector.

Many public administrations provide open-data portals and open-data based applications based on public data, also including private data from companies when possible. The list of current data sharing initiatives analysed in WP2 includes many examples of this kind of "open data lake" style approaches.

However, public administrations are not allowed to include commercial data or include data products for profit. Besides, public administration data sharing approaches does not facilitate B2B data sharing, which is one the main data space goals.

The public administration strategy and the technology used to implement are not aligned with two of the main characteristics of a data space: Data monetization and valorisation and data sovereignty.

- Regarding data transfer, since business is at the core of the current data space initiatives, these initiatives include contract negotiation as a mandatory step, that is not needed in an open data context.
- Open data access normally has a very low security requirements, just download files or a rest interface with no security is enough.

The only example of merging both open and proprietary data (for profit) in the same platform is the FIWARE monetization architecture.

One possible approach could be to adapt the current open data initiatives to the data space initiatives requirements, constraints, and compliance rules, becoming a participant in the data space.

The public administration managing the open data portal should follow the data space on-boarding process and all the data sets/products included should be defined and certified according to the trust framework defined in the data space.

Furthermore, another channel to get the data should be added to the current ones, a data connector compliant with the data space requirements.

5.4 Tourism specific requirements technical challenges: Relation with other dataspaces

The strong interconnection between the tourism data space and other data spaces holds immense significance due to several compelling reasons. Firstly, data sharing among these spaces is imperative because datasets are inherently linked but value comes from specific data semantics on each particular data space. For example, in the relation between the mobility and tourism domains, it's crucial to not only track people moving between countries using any transportation medium but also discern who among them are tourists. Identifying the origin and destination of tourists as well as purpose is vital for understanding their economic impact and environmental footprint within the tourism sector.

Tourism is intricately intertwined with various other data spaces such as mobility, healthcare, sustainability, and smart cities. The interactions between these domains are profound as tourists' movements and activities have wide-ranging implications. Their healthcare needs, influence on urban infrastructure, and contributions to sustainability practices all rely on data from these interconnected spaces.

To facilitate effective interoperability and data exchange among these spaces, the establishment of standardised protocols and formats is paramount. These standards serve as a compatibility layer, enabling data from diverse sources to harmonise while preserving the unique tourism-specific aspects of each dataset. This approach fosters data sharing and collaboration, facilitating insights, innovation, and a comprehensive understanding of the tourism ecosystem within the broader context of interconnected data spaces.

In this context, data intermediaries play a pivotal role. They bridge the gap between data producers and consumers, enriching and standardising tourism data while facilitating seamless data exchange. These intermediaries bring valuable analytics, predictive models, and actionable insights, empowering tourism stakeholders to make informed decisions, optimise operations, and enhance traveller experiences. Their adaptability and transformation capabilities are critical for ensuring compatibility and integration across data spaces, thus maximising the potential of these interconnected ecosystems.

These are some examples of the connection between tourism and these other sectors:

- **Transport:** The volume of air passengers and the reservation forecasts are very useful indicators for any tourism manager.
- **Mobility**: Data from mobile phone operators are used to measure the flow of tourists at a specific destination or point of interest.

- **Environment**: Tourism activity data is used to monitor the carbon footprint of a destination.
- **Energy:** Monitoring of energy consumption data (e.g. resort) should optimise energy efficiency systems.
- **Cultural and environmental heritage**: In protected natural resources is essential to identify the maximum carrying capacity to preserve the original conditions and usually maximum entry quotas are established.

In this way, **data from these sectors should also feed the ecosystem** of a tourism data space. The current approach in Europe is to build separate data spaces for each sector or domain or in some cases even more than one depending on the data space scope and characteristics.

Inter and intra data space interoperability is a critical issue for tourism data spaces.

The technical challenge to solve is: What happens if one company needs data from a company in another participating in other data space? The key topics to address regarding inter data spaces technical interoperability are:

- Technical: How find data in another data space. What connectors are being used?
- Semantic: Different data models, both generic (i.e., data product, participant, etc) and specific tourism data models
- Onboarding processes
- Monitoring

Two main strategies can be applied:

- **Participate in several data spaces**. The company needs to follow several on-boarding processes. This strategy probably implies to use different data models, identity providers, connectors, etc. This approach is not scalable but nowadays is the more pragmatic one.
- **Federation among data spaces**. With this strategy the participation in several data spaces is transparent for the companies. However, this approach is not feasible from the technical point of view unless the data spaces use the same data space governance framework.

6 ACTIONABLE RECOMMENDATIONS

This section presents some final conclusions and recommendations to develop the ETDS. The recommendations are based on the current status or the main data space initiatives.

Frist of all, the previous results included in deliverable D3.1 are still valid.

- **Identity and self-description**: The need for a decentralised solution for identity management. Self-sovereign identity solutions along with verifiable presentation and verifiable credentials for self-descriptions (*GAIA-X*)
- **Trust framework and onboarding**: The concept of Data Space Governance Authority defining and implementing the rules to be part of a data space (*all the initiatives*)
- **Data transfer**: Decoupling of control and data planes in data transfer technologies, making it possible to use any transfer protocol or technology available (*IDSA Dataspace protocol*)
- **Generic rule about data models**: Use common and well-established standards if available. (*All the initiatives*). Here are some examples:
    - DCAT for data product
    - ODRL for data usage policies
    - W3C DQV for data quality (based on ISO)

Furthermore, two more specific issues of the Tourism data space have been added: Geographic data and Local vs regional vs national vs international scope. The geographic data usage in the Tourism data space adds some specific data models and the tourism different scopes could lead to more than one data space covering different regions.

As a summary, the recommended approach is to follow the onboarding process defined by Gaia-X including the compliance framework and the clearinghouse.

The second step would be to add some sector specific criteria and trust anchor, for example the adoption of the CODE OF CONDUCT ON DATA SHARING IN TOURISM[28].

The intermediaries are still being developed and the only available and stable solutions are the solutions provided by IDS: the IDS metadata broker, the vocabulary provider and de IDS clearinghouse (different from the Gaia-X clearinghouse) for logging. However, these solutions are not being integrated with the current Gaia-X framework.

Regarding the data space connector, the recommended one is the Eclipse Data Space connector which implements the new IDS data space protocol.

[28] Code of Conduct on Data Sharing in Tourism. European Travel Commission. Retrieved September 2023 from: https://etc-corporate.org/reports/code-of-conduct-on-data-sharing-in-tourism/#:~:text=The%20goal%20of%20the%20Code,partnerships%20in%20the%20tourism%20industry

Finally, it should be taken into account that several initiatives are running in parallel defining business cases, the scope, legal framework and technology needed to implement and deploy data spaces. The components being developed are evolving at a very high pace and are not interoperable with each other and convergence or interoperability among the different initiatives is still an open issue.

In this changing and uncertain environment, it is very difficult to know in advance what will be the "winning" initiative will be, if any.

Therefore, the roadmap for the definition and implementation of the Tourism data space and the recommendations included in the report could become obsolete in the future, so it is recommended to establish a **technological surveillance process** that allows adaptation to the future development of data space technology.

Next, some more detailed recommendations are presented. It is very important to note that this is a **technical design and roadmap**, which means that the organizational aspects of the data space are not taken into account. For example, the first recommendation is related to create the data space governance authority. If this organization is only one company, a public administration or a consortium is not directly related to the technical approach, so it is not discussed in this section.

6.1 Recommendation on Data Space Governance Framework and Data space governance authority

The recommended approach is to follow the onboarding process defined by Gaia-X including the compliance framework and the clearinghouse.

The **Gaia-X Governance framework** is being adopted widely as the base for mandatory compliance criteria. It is a good starting point because of its maturity and flexibility, adopting Gaia-X will allow the future ETDS to gain interoperability with a wide range of data spaces initiatives while maintaining a non-opinionated strategy for adopting connectors and other components into the implementation. Gaia-X Compliance process is composed by the Trust Framework, and the Policy Rules & Label document.

In addition, Gaia-X explicitly encourages adopting domain specific compliance criteria, that could improve the coverage of the specific challenges presented in this blueprint such as the policies to ensure GDPR compliant when dealing with personal data within the use cases or the integration of traceability practices when accessing personal data.

Also, Gaia-X federation and decentralisation architecture will allow an inclusive strategy to onboard different regional tourism data space initiatives into a shared implementation of the ETDS.

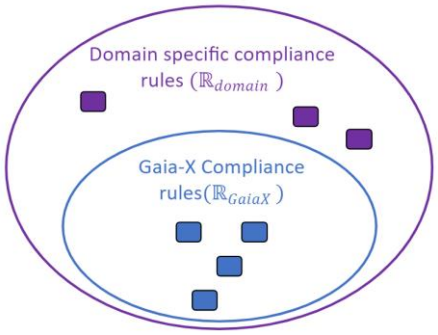

Figure 19 Gaia-X Governance framework as a base for mandatory and optional criteria with additional Tourism specific criteria

According to Gaia-X The GXDCH is the necessary element to operationalize Gaia-X in the market. The Gaia-X Framework describes functional specifications, technical requirements, and SW assets necessary to be Gaia-X compliant.

The GXDCH are a network of execution nodes for the compliance components that Gaia-X has developed. This safeguards the distributed, decentralised ways of running the Gaia-X compliance, not operated centrally by the Association, and where anybody can benefit from the open, transparent, and secure federated digital ecosystem.

Currently there are three clearing houses up and running, and several organizations are working to became new clearing houses all over Europe.

Figure 20 Current Gaia-X clearing houses.

One interesting feature of the clearing house is the development of the Gaia-X wizard that will facilitate the companies, especially the SMSs the onboarding process.

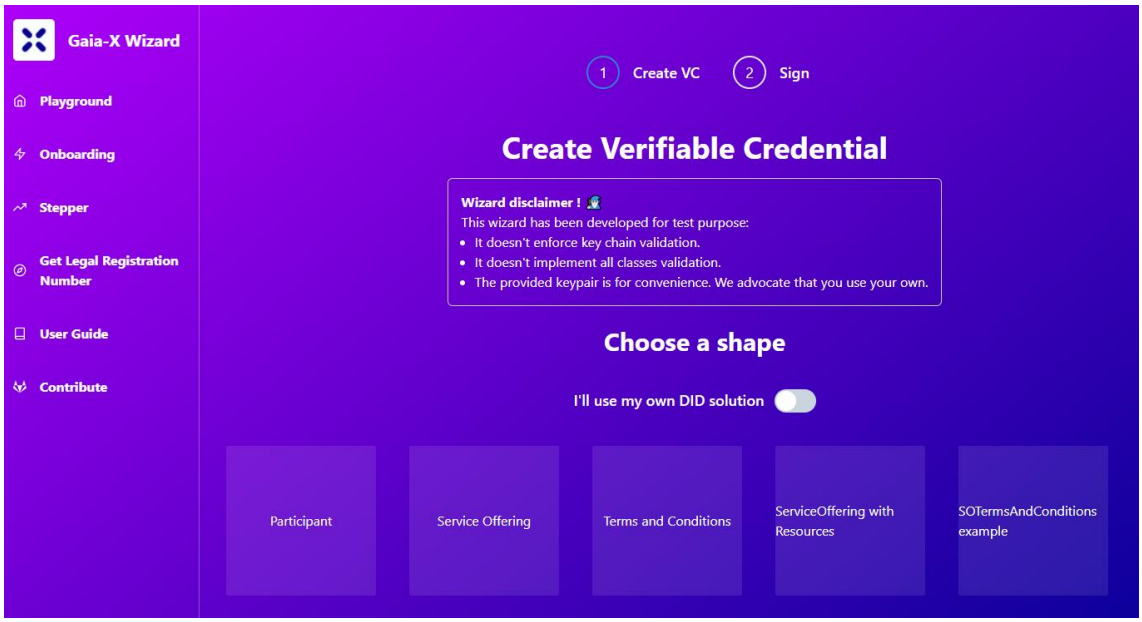

Figure 21 Gaia-X wizard

6.2 Recommendations on data space intermediaries

6.2.1 Services discoverability

In order to provide service discoverability all the data spaces initiatives include the Catalogue Federation Technology. This technology is used to create a unified, federated catalogue of available data and services across different entities within the data space. It allows Data Space Entities to discover, access, and utilize resources and data assets across the network easily.

The IDS architecture includes the Meta Data broker and Gaia-X is defining and developing a federated catalogue.

The Gaia-X federated catalogue technology is still being developed and the IDS metadata broker is not compatible with the Gaia-X data product model. Therefore, services discoverability is still an open issue.

The goal of the Gaia-X Federated Catalogue is to:

- enable Consumers to find best-matching offerings and to monitor them for relevant changes of the offerings
- enable Producers to promote their offering while keeping full control of the level of visibility and privacy of their offerings.
- avoid a gravity effect with a lock-out and lock-in effect around a handful of catalogue instances.

For the Gaia-X Catalogues to be notified about new, updated, revoked credentials, a central Publication / Subscription service (pubsub service) is being defined and must be deployed via

the GXDCH instances. However, the software stack to create and deploy such a service will be detailed in further technical specifications.

6.2.2 Data space monitoring and observability

Observability refers to the capability to monitor, log, and analyse system activities and performance.

In Gaia-X, a logging service is used to provide transparency and visibility into the data-sharing processes, ensuring that the ecosystem is operating securely and efficiently. This is crucial for compliance, auditing, and issue resolution.

However, the Gaia-X logging service[29] has not been developed by Gaia-x and it is not clear if it will be finally part of the official Gaia-X software.

6.3 Recommendation on connectors

The current trend regarding data space protocols is Decoupling of control and data planes in data transfer technologies, making it possible to use any transfer protocol or technology available.

According to the DSSC blueprint, it is important to distinguish between a control plane and a data plane. The control plane is responsible for deciding how data is managed, routed and processed. The data plane is responsible for the actual moving of data.

For example, the control plane handles the identification of users and the handling of access and usage policies. The data plane handles the actual exchange of data.

Therefore, our recommendation is to use one of the connectors using this approach. The connector being used by the main data space implementations is the **Eclipse Data Space connector**.

6.4 Recommendation on data models

One of the general trends for data models is to use common and well-established standards if available, both for generic (sector agnostic) models as well as tourism specific domain models. In the case of generic data models, the following ones are being adopted by Gaia-X and are necessary to follow the onboarding process.

- **Data product model**
    - DCAT 3.0 for data product
    - ODRL 2.0 for data usage policies
- **Participant model** → Participant self-description defined by Gaia-X
- **Service offering model.**

[29] https://gitlab.eclipse.org/eclipse/xfsc/del

However, the definition of the generic Data product is one of the main ongoing processes in Gaia-X and in other data space initiatives.

6.4.1 MTourism data models.

At present, there is no single common European ontology in tourism that has been universally adopted and standardized across all European countries. Nevertheless, serious efforts have been made to develop ontologies and standards that can be used as common reference points for tourism-related data and interoperability within Europe.

Notable initiatives in this regard have been carried out by organizations like the World Tourism Organization (UNWTO) to promote interoperability and data sharing in the tourism sector. For example, they promoted the multilingual "Thesaurus on Tourism & Leisure Activities". This thesaurus can be used as a guide to tourism terminology, as well as for the standardization and normalization of a common indexation and research language, at an international level.

Our recommendation is not to create a new "tourism data model" but use the current national initiatives. In case there is not a national initiative in place, select and adapt an existing "tourism ontology" from other country. Leave for a second stage the data models federation among countries.

These are some examples of these ontologies:

- **Digital Tourism Hub in Italy**: www.italiadomani.gov.it/en/Interventi/investimenti/hub-del-turismo-digitale.html
- **Ontology DATAtourisme in France:** www.datatourisme.fr
- **Segittur Tourism Conceptual Reference Model in Spain:**
    - www.segittur.es/transformacion-digital/proyectos-transformacion-digital/modelo-conceptual-de-referencia-para-el-desarrollo-de-una-red-de-ontologias-del-sector-turistico/

6.5 Tourism data space deployment roadmap recommendation

Our recommendation regarding the data space deployment is to follow the two main paths on parallel. Follow the "theoretical path" to the creation of the data space while using a more pragmatic approach for the deployment of use cases.

Theoretical roadmap:

- Use the Gaia-X onboard framework since it is the more advance one and it is following an open development process.
- Regarding the operationalization of the onboarding, a Gaia-X clearing house should be deployed or use one of the existing ones.
- The companies should create the participant self-descriptions and data services self-descriptions (in the case of data providers) and use the clearing house services to follow the compliance process.

Pragmatic roadmap:

- Select a specific stable version of a connector implementing the new IDS data space protocol with a Federated Authentication schema (Keycloack or DAPS). Even though connectors using the same data space protocol should be interoperable (and some of them even claim they are interoperable) the real situation now is that even different version of the same connector can´t interoperate.
- Deploy and configure the connector in the participants (both providers and consumers) and start the data transfer processes defined by the us cases.
- Eventually change the authentication and authorization methods by an SSI approach when available.

7 REFERENCES

[1] "*Study on mastering data for tourism by EU destinations*", EU project Smart Tourism Destinations, July 2022: https://data.europa.eu/doi/10.2873/23880

[2] "*Holistic view of platforms and Intelligent Destination Platform solutions",* Segittur (Spanish Mercantile Society for the Management of Innovation and Tourist Technologies), February 2023: https://www.segittur.es/wp-content/uploads/2023/02/20230602-Vision-Holitica-Plataformas-y-soluciones-PID.pdf

[3] Results of the workshop on "*Use Cases of the Spanish Tourism Data Space*" held in October 2022 by Globaldit in collaboration with the Spanish Data Office / SEDIA: https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2022/Infografias_WS_Turismo.pdf

[4] "*Radiography of the Spanish Tourism Data Space in Spain*", Gaia-X Hub Spain, December 2022: https://portal.mineco.gob.es/RecursosNoticia/mineco/prensa/noticias/2022/Radiografia_Espacio_Datos_Turismo_Espana.pdf