Deliverable D3.2

Technical specifications for interoperability

31/07/2023

PROJECT INFORMATION

Project Number	101084007
Project Acronym	DATES
Project Full title	European Data Space for Tourism
Project Start Date	November 1 2022
Project End Date	October 31 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	D3.2
Deliverable Title	Technical specifications for interoperability
Work-Package No	WP3
Work-Package Title	Technical Framework and Design
WP-Leader (Name and Short Org. Name)	Valentin Sanchez (TECN)
Tasks No	T3.2
Tasks Title	Common standards and interoperability
Task Leader (Name and Short Org. Name)	Dr. Andreas Eisenrauch (AMAD)
Main Author (Name and Short Org. Name)	Dr. Andreas Eisenrauch (AMAD)
Other Authors (Name and Short Org. Name)	Jens Schwarz (OUTD) Tatiana Semenova (MITUR) Adele Pia Villani (MITUR) Belén Martinez (TECN) Sonia Bilbao (TECN) Jesús Herrero (TECN) Mariano Blaya-Andreu (IDSA)
Reviewers (Name and Short Org. Name)	Valentin Sanchez (TECN) Martin Soutschek (OUTD) Umutcan Serles (STI, contractor) Dolores Ordóñez (ANYSOL)
Status	Draft ☒ Final ☐

Deliverable Type	Report ☒ Data ☐ Demonstration ☐ Data ☐
Dissemination Level	Public (PU) ☒ Sensitive (SEN) ☐ EU Classified (EUC) ☐ *PU: fully open (automatically posted online on the Project Results platforms)* *SEN: limited under the conditions of the Grant Agreement* *EUC: WU-RESTRICTED, EU-CONFIDENTIAL., EU-SECRET under Decision 2015/444.*
Date Approved by Coordinator	31/07/2023

DOCUMENT REVIEW HISTORY

Date	Version	Reviewer Name (Short Organisation Name)	Summary of main changes
10.07.2023	V0.1	Valentin Sanchez (TECN)	Contribution to technical standards.
13.07.2023	V0.2	Martin Soutschek (OUTD)	References to the DCAT metadata standard and data exchange APIs in 2.3 and 2.8 added.
24.07.203	V0.3	Umutcan Serles (STI, contractor)	Semantic data model specifications and data format references to VoID and ODTA in 2.2. and 2.8 added.
26.07.2023	V0.4	Dolores Ordóñez (ANYSOL)	Point 3: contributions to the different sections. Relation with other WPs.
31/07/2023	V1	Dolores Ordóñez (ANYSOL)	Revision and Acceptance v0.4

FINAL ACCEPTANCE BY COORDINATOR:

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

LEGAL DISCLAIMER

INDEX

1. EXECUTIVE SUMMARY 8
2 INTRODUCTION 10
2.1 Project Summary 10
2.2 Technical Standards supporting Interoperability 11
3 ESTABLISHED TECHNICAL AND CONCEPTUAL STANDARDS FACILITATING DATA SPACE INTEROPERABILITY 14
3.1 General Remarks 14
3.2 Standards regarding Data Models & Formats 15
3.3 Data Exchange APIs 17
3.4 Provenance and Traceability 18
3.5 Identity Management 20
3.6 Access & Usage Policies and Control 22
3.7 Trust 23
3.8 Data, Services and Offerings descriptions 25
3.9 Marketplaces & Usage Accounting 26
3.10 Publication and Discovery 28
4 ESTABLISHED STANDARDS FACILITATING DATA SPACE GOVERNANCE 31
4.1 Overarching Cooperation Agreements 31
4.2 Operational Governance 33
4.3 Continuity Model 34
4.4 Business Agreements 35
5 STANDARDS SUPPORTING OTHER OPEN DEI BUILDING BLOCKS 36
5.1 Other Technical Building Blocks I: System Adaptation 36
5.2 Other Technical Building Blocks II: Data Analytics Engine 37
5.3 Other Technical Building Blocks III: Data Visualization 39
5.4 Other Technical Building Blocks IV: Workflow Management Engine . 41
6 CLOSING REMARKS 42

List of Abbreviations and Acronyms	
AISBL	Association internationale sans but lucratif (legal form of GAIA-X)
API	Application Programming Interface
BAE	Business Application Ecosystem
C2D	Compute-to-data
C2E	Compute-to-edge
CRS	Central Reservation System
CSA	Coordination and Support Action
CSV	Comma separated value
DSBA	Data Spaces Business Alliance
DGA	Data Governance Act
DID	Decentralized identifier
DIH	Digital Innovation Hub
DMO	Destination Management Organization
DSA	Data Space Authority
DSSC	Data Spaces Support Center
DATES	Data Space for Tourism
EC	European Commission
eID	Electronic identity
eIDAS	electronic IDentification, Authentication and trust Services
EU	European Union
FAIR principles	Findable, Accessible, Interoperable and Re-usable
FIWARE	Organization providing a Data Space 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
JSON-LD	JavaScript Object Notation - Linked Data
KPI	Key Performance Indicator
MyData	Human-centric approach to personal data management
NGSI	Next Generation Service Interfaces
NGSI-LD	Next Generation Service Interfaces – Linked Data
ODTA	Open Data Tourism Alliance
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

PMS	Property Management System
PoC	Proof of Concept
POI	Point of Interest
REST	Representational state transfer
SimpI	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
WP	Work Package
XFSC	Cross Federation Services Components (repository on GitLab, where the GXFS components are now hosted)
XSLT	Extensible Stylesheet Language Transformations

List of Figures

Figure 1: Data Space Building Block Schema as originally published by Open DEI 12
Figure 2: OpenDEI Building Block Schema as updated by DSSC (state July 2023) 13
Figure 3 Illustration of the different levels of governance, proposed by aNewGovernance 32
Figure 4 Magic Quadrant for Analytics and Business Intelligence Platforms 40

1. EXECUTIVE SUMMARY

Data Spaces are a key instrument to promote data sharing by accelerating the data economy in different verticals. Tourism, being one of the most powerful industries in Europe, because its countries are preferred worldwide tourism destinations, will increase its competitiveness thanks to the creation of a continent-spanning European Tourism Data Space. DATES will work intensively with the whole tourism ecosystem to set the fundament of the European Data Space of tourism based on the principles of data sovereignty, trust, resilience, and sustainability.

The design of such a Data Space has many organisational and technical aspects, making a Data Space a complex entity, and its creation is still a non-trivial challenge. In order to cope with these challenges a body of high-level theoretical models has been created up to now by various projects and organisations. The method was to identify a system of so-called "building blocks" dealing with all key aspects of Data Space design and operation. The most accepted concept is that of the Open DEI initiative. In spring 2023 the Data Space Support Center (DSSC) started its work to provide recommendations for a common data space architecture and technical standard elements. It basically adopted the Open DEI building block schema, but slightly modified the terminology as described in section 1.2.

Within DATES this updated schema by DSSC has been used as the framework for identifying relevant generic and sector-specific elements for the future tourism Data Space. The assessment of these important building blocks and their integration within preferable architecture models was the objective of other task groups within DATES and their findings are described in the deliverables "D2.3 Identification of data typology and priority list of datasets, potential use cases and common building blocks with other Dataspaces" and "D3.1 Reference architecture", respectively.

This task T3.2 "Common standards and interoperability" extends now the previous work, insofar it addresses concrete standardized elements (software, protocols, concepts), which might be used for the implementation of the Tourism Data Space. Obviously, using such standardised elements are key for easy connecting all participants of a specific Data Space (so-called intra-Dataspace interoperability), but moreover is an inevitable prerequisite for future connections with other Data Spaces of the Tourism sector and with those of other sectors (inter-Dataspace interoperability). However, one core technical component of each data space is the "connector", which each member of a data space must implement in its private environment. The recommendation regarding this key element is in the scope of the Task 3.4 and will be reported in D3.3 "Tourism Data Space detailed design and blueprint specification, including data connectors".

To best meet the interoperability requirement, the investigation for eligible standards has been carried out in close collaboration with the Data Space Support Center (DSSC), which has a similar target, because it is supposed to provide a generic blueprint for the creation of Data Spaces in general by end of this year at latest. As mentioned above, DSSC's starting

point was again the building block schema of OpenDEI, which they slightly modified, meanwhile. Within this task it was thus focussed on the identification of Tourism specific standards for the building blocks, and moreover, **on the best practises implemented by mature data sharing initiatives**, as retrieved from the numerous interviews executed by the team of WP 2. Therefore, in the chapters where the standards eligible for being used to implement the respective building block are described, we highlight especially those items, which are obviously recommendable, because they are outstanding generic standards on one hand, and secondly, those which are proven best practices of existing Data Spaces.

2 INTRODUCTION

2.1 Project Summary

To turn the vision of a **European Tourism Data Space** into reality, DATES brings together key actors of the tourism and data ecosystems and their wide public, business and research partner networks. Convinced that tourism data are crucial for the **data economy in a European Single Digital Market**, DATES will develop a measurable contribution for the digital transformation of the services sector that can significantly strengthen the European competitiveness.

The project outcomes provide a basis for **governance and policymaking** to foster innovation powered by tourism data and will enable society to make Europe the most desired sustainable space for living. Starting from a mapping of the EU and non-EU tourism data landscape (with a focus on the European environment), and leveraging on initiatives on data sharing, DATES will reach out to relevant stakeholders to collaboratively develop a shared **strategy roadmap** for building a sustainable tourism data space. The process will define clear objectives and key results to inspire, support and motivate all stakeholders to contribute and use high quality tourism data as a basis for innovation.

DATES will provide recommendations for governance and digital business models, and it will highlight how benefits for society can be created. Key success factors will be identified, and it will be outlined how a tourism data space can create added benefits for the tourism industry and all sectors that tourism is interlinked with. In addition to providing a comprehensive inventory of existing platforms sharing relevant data, **blueprints for addressing technical and organisational challenges will be created** to spark and fuel use of interoperability standards and participation in a tourism data space to foster the digital transformation of SMEs in tourism and relevant cross-sector industries. DATES will be a kick-start for implementation of a European Tourism Data Space by involving all players on the supply and demand side through the strong ecosystem representatives of the DATES.

The following actions will give an overview over **established technical standards**, which are eligible to be recommended in the future **blueprint** to be used in the realization of the respective technical building blocks. Within the objectives of WP3, listed on page 35 of the application form part B, the following items are especially addressed in D3.2:

- Identify common standards, including semantic standards and interoperability protocols – both tourism-specific and crosscutting.
  These are listed in the chapters 2.2 to 4.4 together with hints about which established data sharing platform uses them in its environment.
- Identify common toolboxes that could be used across data spaces.
- Identify existing APIs that are relevant for a Tourism Data Space.

2.2 Technical Standards supporting Interoperability

The interconnection of data and the implementation of innovative applications on top of them is considered to be one pillar of improving the **well-being of the inhabitants of Europe, as well as the competitiveness of European society in comparison with other regions in the world**. The technological concept selected to achieve this, is that of the Data Spaces. Legacy data exchange modes, like peer-to-peer connections or data lakes, where many data providers pool their data on one side and many data consumers take out selected subsets of that data, have the disadvantage that the centralized data storage and moreover the custom-made connections are costly, and poorly scalable, even in the times of cloud computing. Data providers give up the control over their data when they leave their premises. In contrast, **Data Spaces are decentralized networks of interconnected data providers and data consumers**. They allow especially Many2Many data exchange connections. Data Providers are supposed to keep control of their data, they may be exchanged with authenticated data consumers on the base of use specific contracts, and agreed transactions which follow should be monitored to provide usage control at best for the data provider, and provenance tracking for the consumer.

Each activity within a Data Space depends on **commonly shared and standardised concepts and structures**, which only allow for seamless exchange of information necessary for operation of the Data Space itself, and for the exchange of business data between the partners in a Data Space. Moreover, as it is inevitable that a sector specific Data Space like the Tourism Data Space will soon interact with related sector specific Data Spaces (e.g. mobility, cultural heritage, environment, health, smart cities), technical and organisational standards should (or even must) be chosen which **facilitate the interoperability within the Tourism Data Space, but have a high probability at the same time to be chosen also by other data space projects**.

This study therefore deals with the functional and methodological aspects of interoperability standards of relevance for the Tourism Data Space, by capitalizing on leading expertise across EU and at local level. In particular, leveraging on the assessment within WP2 to evaluate existing Data Spaces and the respective features, including data ecosystems, open geo-spatial standards, ontologies, data services and reference models. Based on that **it highlights best practices** (collected from interviews with advanced data space initiatives) and common standards (by analysing the proposals published by EU working groups and standardisation organisations, respectively) that overcome critical interoperability challenges. Based on these insights, those common standards and interoperability protocols are listed, which are essential for the development and harmonization of new data services for the tourism industry and those industries connected to it.

The proposed standard concepts and technologies are grouped, as mentioned, according to the building block schema set up by the Open DEI project (see figure below). The same categorization has been adopted in the beginning by the DSSC for its collection of standards in preparation of the DSSC Blueprint; the Open DEI concept is basically still valid

despite some modifications were introduced by DSSC, meanwhile. However, these building blocks are quite high-level categories, summarizing more or less homogenous groups of capabilities. Sometimes the finer grained capabilities are used in the text to better explain the applicability of some necessary standard, where appropriate.

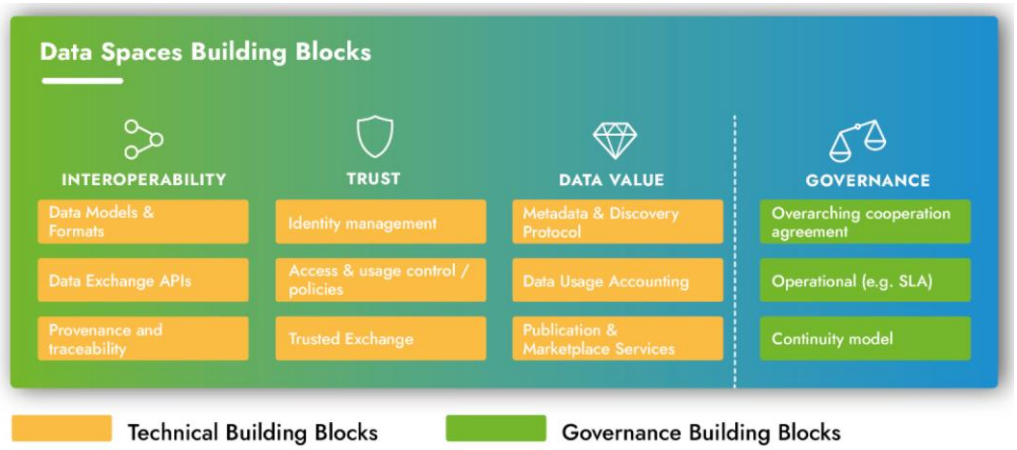

Figure 1: Data Space Building Block Schema as originally published by Open DEI

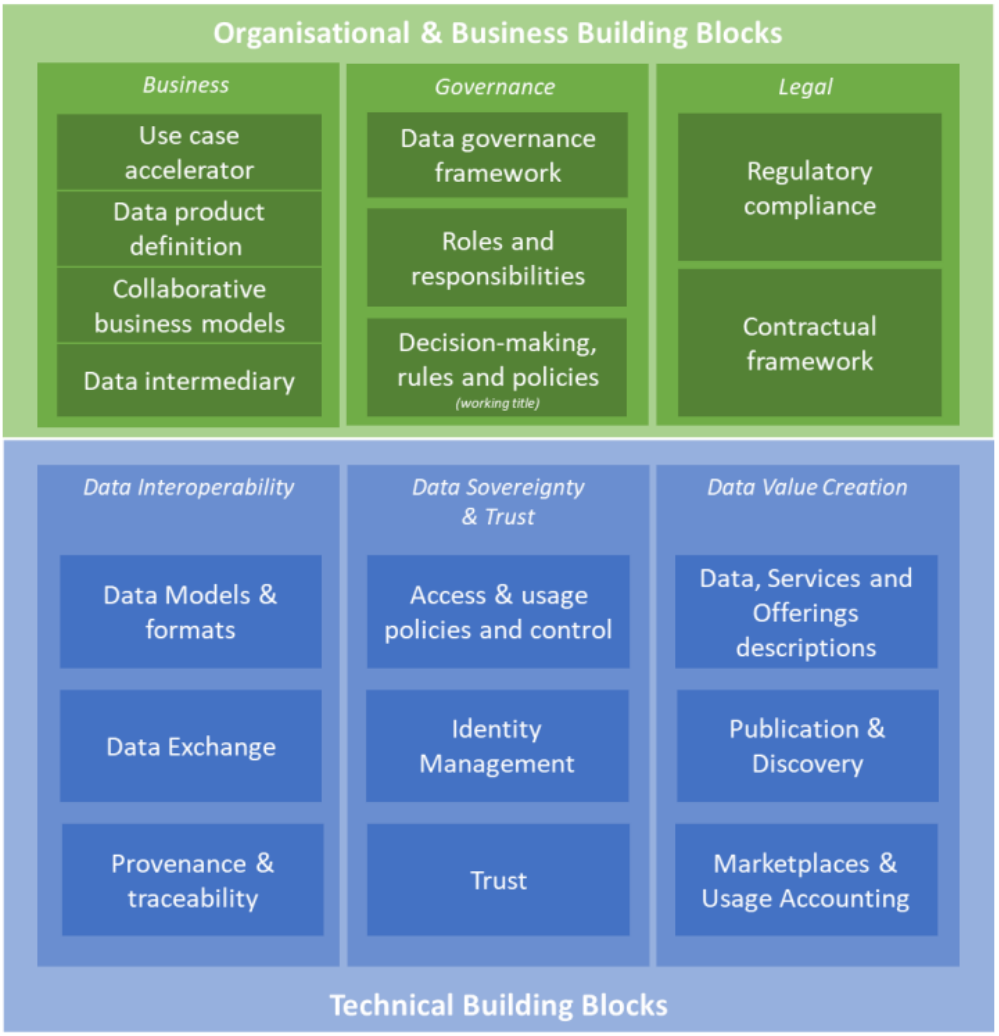

Figure 2: OpenDEI Building Block Schema as updated by DSSC (state July 2023)

3 ESTABLISHED TECHNICAL AND CONCEPTUAL STANDARDS FACILITATING DATA SPACE INTEROPERABILITY

3.1 General Remarks

Parallel to the work done in DATES, the Data Spaces Support Center (DSSC) started a similar study with the objective to identify potential technical and conceptual standards for the implementation of the capabilities required by the OpenDEI building blocks. Since the DSSC is supposed to recommend standard components of future data spaces, DATES did not want to set up an alternative list of such elements. The standards highlighted in the following sections for use in the tourism dataspace is therefore basically the list drafted by DSSC in spring 2023, enriched with findings retrieved by the DATES team during its own research and from the interviews with existing data space initiatives. Particular focus for such amendment was on identifying tourism specific standards, because sector specific details were not in scope of the DSSC study.

The general remarks section may be the right place to discuss the "common toolboxes that could be used across data spaces", whose identification is one sub-objective of this deliverable. Such "toolboxes" exist meanwhile, indeed. However, it is DSSC's objective, to analyze them, together with other independent tools which we list below, and which are also collected by DSSC, and compile thereof a blueprint describing a consolidated set of interoperable tools, that might then serve as an European standard for the setup and operation of data spaces.

The modules (=building blocks) of those "toolboxes" appear frequently in the following chapters, but it seems appropriate to highlight them once as more or less complete systems which provide many elements to run a data space with. There are several widely established systems:

1. the IDS toolsuite[1], specified by the International Data Space Association (IDSA) and implemented by its partners;
2. Originally a derivative thereof are the Eclipse Dataspace Components (EDC)[2]
3. The GAIA-X Federated Services[3] (GXFS, the repository will be migrated to XFSC on GitLab[4])
4. FIWARE[5]
5. iShare[6]

[1] https://github.com/International-Data-Spaces-Association
[2] https://github.com/eclipse-edc
[3] https://gitlab.com/gaia-x/data-infrastructure-federation-services
[4] https://gitlab.eclipse.org/eclipse/xfsc
[5] https://www.fiware.org/
[6] https://ishare.eu/

6. The building blocks of Digital Europe[7]
7. Italian Tourism Data Hub (TDH22)[8]

The technologies of the first three at least are currently strongly converging and become more and more interoperable. For Example, EONA-X uses in its pilot data space elements of the EDC bundle, especially the EDC connector, and the GXFS Clearing House[9]. Moreover, the EDC connector has recently been adapted to support the IDS protocols. IDS[10] and EDC[11] provide meanwhile so called "Minimum Viable Dataspaces" (MVDS). These are specifications, guidelines and open source software to build a core infrastructure for data sharing. The FIWARE and iShare solutions are also on a good way to become interoperable with the GAIA-X/EDC/IDS solutions. A bit apart from the technology of the "GAIA-X mainstream" is the solution of Digital Europe, because it is much older than the data space concept. Their focus is on connecting public offices and citizens (thus G2C/C2G), its main limits come from the still insufficient support of business participants. However, the modular structure of their system allows the usage of some elements also in an GAIA-X/EDC/IDS environment, especially the eID module might be a strong candidate to be used in establishing a trustful IAM in the future tourism data space.
The Italian tourism data sharing platform "Tourism Data Hub" has been launched in 2022. It is indeed the only platform which is currently exclusively dedicated to the tourism sector. The platform is open, scalable, and interoperable enabling to connect the tourism destinations with related offerings and contents of interest to tourists. Interoperability between TDH and the ecosystem is enabled by "TDH022", a standard communication protocol for the interchange of data and contents via APIs, organized into ontologies and taxonomies, that connects the TDH with public organizations and private brands/partners. The standard has been built based on interoperability guidelines defined by AGID and in line with the European Interoperability Framework. Compatibility with GAIA-X/EDC/IDS solutions is assured by leveraging on existing APIs or integrating with new ones.

3.2 Standards regarding Data Models & Formats

The Open DEI building block "Data Models & Formats" addresses a key requirement for the Tourism Data Space. Whereas many of the building blocks address features which are essentially sector agnostic, especially the block dealing with data models is the one with the most tourism-specific character. An ontology, in the context of data models, is a semantic network that represents the knowledge associated with a specific domain (the Tourism one, in this case), being able to organize, model and manage information and knowledge, so to share them and reuse in different domains. Its base is a coherent taxonomy of its elements.

[7] https://ec.europa.eu/digital-building-blocks/wikis/display/DIGITAL/Digital+Homepage
[8] https://docs.italia.it/italia/mitur/gl-tourism-digital-hub-interoperabilita-docs/it/main/index.html
[9] See a demo video on YouTube
[10] https://github.com/International-Data-Spaces-Association/idsa/tree/main/how-to-build-data-spaces
[11] https://github.com/eclipse-edc/MinimumViableDataspace

That is a hierarchical form of organizing the information, so as to spot the relationship between data. In the future ETDS, it will be mandatory to have a well-organized taxonomy that enables the touristic organization to understand the diverse relationships between all owned data. Thus, standardised data models need to be applied wherever in a Data Space information has to be exchanged. Obviously, this holds in particular for the touristic business data, which are exchanged within transactions between partners of a Data Space.[12]

The base for most ontologies is the typology for structured, machine-readable data developed by "schema.org"[13]. The Schema.org vocabularies are designed in a way enabling the use with different encodings, including RDFa, Microdata and JSON-LD. Extensions especially for the tourism sector exist meanwhile in various variants. A recommendable tourism specific ontology has been derived by the Open Data Tourism Alliance[14], a multinational consortium based in German-speaking countries that is currently integrating with the models of organizations from other European nations (they mention Sweden, Great Britain, Belgium and France). ODTA consists of more than a dozen domain experts and produces schemas called Domain Specifications that restrict and extend schema.org to create metadata for semantically describing tourism-related data. This restriction and extension is done for two reasons: (1) schema.org covers many domains, and is not always conceptually accurate (e.g., waterfalls can have phone numbers), ODTA selects tourism-related schema.org types and restricts their properties to a useful subset. (2) the coverage of schema.org is wide, however relatively shallow. This shortcoming is mitigated by extending schema.org with new types and properties that cover the tourism domain in a deeper manner. The domain specifications are published as SHACL shapes, to increase interoperability with other semantic technologies.

Another advanced ontology for the tourism sector has been developed by the Italian "Tourism Digital Hub"[15] (TDH022, sponsored by the Italian Ministry of Tourism). It is also based on the schema.org typology. In TDH022 the ontologies identified were related to the dimensions associated with tourists as well as the dimensions associated with the tourism operators. By being able to correctly represent them, the TDH022 has interoperated with all the partners' ecosystems. A third ontology proposal again based on schema.org has been released by the French government with the "DATAtourisme ontology" [16]. An

[12] This holds also for the metadata of the data sets, partners, and other components/apps, respectively. However their "nature" is quite different to the payload data of transfer actions, since they are needed as identifiers within the operation of the Data Space itself. The formats and handling of metadata will be covered within the discussions of the building blocks "Identity Management" and "Data, Services and Offerings Descriptions", respectively.

[13] See https://schema.org/

[14] See https://open-data-germany.org/en/technology/

[15] https://www.ministeroturismo.gov.it/tourism-digital-hub/

[16] https://www.datatourisme.fr/ontology/core/index-en.html

independent approach has been taken instead by Fiware with its "Smart Destination" model, which is a subset of the output of the Smart Data Model initiative.[17]

During our investigations we discovered still two other Data Models which have strong relationship to the tourism sector:

One published by the Open Travel Alliance (OTA), and the "Hotel Data and Destination Data Standards" by AlpineBits. Both providers offer APIs to handle data structured according to their models (see next section about Data Exchange APIs). Besides that, one generic concept was mentioned several times in the interviews, namely the "SKOS Simple Knowledge Organization System" published under the umbrella of the W3C consortium[18]. SKOS provides a standard way to represent knowledge organization systems using the Resource Description Framework (RDF). Encoding this information in RDF allows it to be passed between computer applications in an interoperable way. Using RDF also allows knowledge organization systems to be used in distributed, decentralised metadata applications. Decentralised metadata is becoming a typical scenario, where service providers want to add value to metadata harvested from multiple sources.

Tourism-related semantic metadata standardization efforts are also being conducted by ISO. The ISO Technical Committee ISO/TC 228 is working on the development of a semantic metadata standard for tourism-related data. The standard development is in the preparatory stage as of July 2023[19].

The most frequent protocols which are used to implement the models are JSON-LD and XML Schema.

According to the IDS-RAM an ontology applied in a Data Space will need a "Vocabulary Provider" building block to be made available to the data providers and data consumers, the implemented feature which supports that is the 'Vocabulary Hub'.

3.3 Data Exchange APIs

Data Spaces are decentralised networks that connect data providers and consumers. Connections are implemented through data exchange APIs that focus on query functions rather than requiring copying of large amounts of data. Data spaces ensure that data providers retain more control over their data and that no central infrastructure is required. The Open DEI building block "Data Exchange APIs", covers general interfaces used for data transfer in data spaces of all sectors.

Since there are few sector-specific requirements, choosing a small set of data exchange APIs for all data spaces helps to ensure that interactions with other data spaces of related sectors can reuse these APIs. The DSSC lists the European Next Generation Service

[17] See https://github.com/smart-data-models/SmartDestination and https://smartdatamodels.org/
[18] See https://www.w3.org/2004/02/skos/
[19] https://www.iso.org/standard/86307.html

Interfaces (NGSI-LD) standard developed by the European Telecommunications Standardization Institute (ETSI) as a good choice. It was initiated and is used by the FIWARE open-source community. In addition, NGSI-LD is available as OpenAPI Specification.

The OpenTravel Alliance (OTA) is an industry led organisation founded in 1999 that aims to meet the needs of the travel industry by developing open standards for the exchange of travel data. OTA standards include data exchange APIs for managing the booking of airline tickets, rental cars, hotel rooms, cruise line trips and other travel and transport related products. The standards are widely in use and are adopted by many large companies making them an obvious candidate for Data Exchange APIs for tourism data space. However, OTA processes and protocols have their base in travel standards going back to the 1960s. The OTA and the OpenAPI Initiative Cooperation are working together on API conventions and standards which will move it closer to other data exchange APIs, such as the ones offered by FIWARE.

Besides global initiatives such as OTA, regional focussed groups like the AlpineBits Alliance, and the Tourism Tech Alliance (TTA) have emerged to address specific needs and requirements that are not yet covered with relevant de facto and de jure standards.

Also TDH022 defined and implemented a set of API Protocols, to guarantee interoperability between all different actors inside and outside the TDH. This is seen as a necessary building block to be able to interoperate with partners that own different technology stacks. So, in TDH022 it has been decided that the same API had to be published in several standards.

3.4 Provenance and Traceability

Provenance, in general, is defined as “information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness.”[20] In the more specific terms of the OpenDEI definitions, the means are meant, which allow for “tracing and tracking in the process of data provision and data consumption/use”. It serves predominantly to document the origin and the factual destination of data. This is obviously important for other building blocks, like usage control and accounting/clearing. For example, proper provenance tracking may enable a data owner, which handed over data to a data provider, to get feedback about to which amount of data has been delivered to which data consumers. On the other hand, a data consumer should be enabled to get some information about the origin of the data, e.g information on the timeliness of data might be of interest in many use cases. The metadata required to characterize the attributes traced, need to be defined “globally” in order to ensure interoperability within a given data space, and between different data spaces. The concrete implementation of this concept has to be achieved by the participants

[20] https://www.w3.org/TR/2013/REC-prov-dm-20130430/#dfn-provenance

which exchange data. They have to fill the relevant metadata with values, and they have to be able to read such.

Surprisingly, there is only one standard protocol published so far to support this building block, and only two initiatives implemented tools to meet this requirement.

Since 2013 the W3C Consortium maintains the data model PROV-DM which defines a standard specification for the collection and processing of information in the context of data sharing, as referenced above. Besides some chapter which the status of "Notes", four components have reached the level of "Recommendation":

> PROV-O (Recommendation), the PROV ontology, an OWL2 ontology allowing the mapping of the PROV data model to RDF [PROV-O];
>
> PROV-DM (Recommendation), the PROV data model for provenance (this document);
>
> PROV-N (Recommendation), a notation for provenance aimed at human consumption [PROV-N];
>
> PROV-CONSTRAINTS (Recommendation), a set of constraints applying to the PROV data model [PROV-CONSTRAINTS];

Although this seems to be an advanced concept, there is no evidence that it has been applied by stakeholders in the tourism sector.

Two of the surveyed/interviewed stakeholders [Themis-X and MDS] referred to products developed by the GAIA-X Federated Services (GXFS) when asked for their capabilities to trace data and actions in their systems.

The "Continuous Automated Monitoring" (CAM)[21] component is a core service within the Gaia-X Federation Service. Its main goal is to provide transparency to the users of Gaia-X about the compliance of individual services offered in the Gaia-X Catalogue. The basis for this compliance are certain requirements and rules that Gaia-X itself has imposed on its system, for example requirements coming from the field of security, such as encryption, data privacy or interoperability. Often, existing standards such as the BSI C5 or EUCS are used as a point of reference. The purpose of the CAM service is to automatically gather evidence that can indicate the compliance or non-compliance of a certain Gaia-X service as a whole or by a concrete instantiation of a particular service by a user. This is especially necessary if the offered services are dealing with sensitive data and thus need to have a higher assurance level, e.g., compared to the assurance level "high" in the EU Cybersecurity Certification.[22]

[21] https://gitlab.eclipse.org/eclipse/xfsc/cam

[22] https://www.gxfs.eu/download/1731/

Secondly, the purpose of the "Data Exchange Logging Service" (DELS)[23] is to provide evidence that data has been (a) submitted and (b) received and (c) rules and obligations (Data Usage Policies) were enforced or violated within the given ecosystem (by default, it is assumed to be a GAIA-X ecosystem). This supports the clearing of operational issues, but also eventually the clearing of fraudulent transactions. Additionally, the data consumer can track and provide evidence on the enforcement or violation of data usage policies. Theoretically, the logs could be used also for clearing and billing.

Another concrete implementation of this building block has been documented recently within the OpenDEI "state of the art" report. There the data use traceability module offered by the "SmashHit"[24] consortium was introduced. It serves two purposes: (1) to trace data flows through fingerprinting/watermarking, and (2) the identification of data leakages. The module addresses the need of data providers, and data processors to get a solution which enables the identification of data leakages or misuse. Fingerprinting and watermark technologies are designed to identify the last data source within the ecosystem. The fingerprint or watermark will provide information on where for example leaked or misused data was last located or forwarded in order to avoid the blaming of innocent participants in a consent chain. The traceability module, which needs to be deployed to every data provider/processor, supplies an interface for logging data transfers and verifications. smashHit's data traceability infrastructure also tackles various privacy concerns of data owners and supports GDPR compliance.

Neither of the two GXFS products nor the SmashHit concept refer to the PROV-DM schema. Therefore, it cannot be excluded that the PROV concepts are not suited to serve the newer Data Space concept.

3.5 Identity Management

Identity management is a key building block for data spaces, as it enables the secure and trustworthy exchange of data among different actors. Identity management refers to the processes and technologies that establish, maintain, and terminate the identities of data providers, consumers, and intermediaries, as well as the data assets themselves. Identity management also ensures that the access rights and policies associated with each identity are enforced and verified. By implementing identity management solutions, data spaces can enhance their security, privacy, compliance, and interoperability.

Find below the list of relevant standards in the scope of Identity Management.

Description	Publisher	Link to further information
Authentication & Authorization	GXFS	Authentication & Authorization

[23] https://www.gxfs.eu/data-exchange-logging-service/ and https://gitlab.eclipse.org/eclipse/xfsc/del

[24] https://smashhit.eu/enabling-data-use-traceability-with-smashhit/

Decentralized Identifiers	W3C	Decentralized Identifiers
eIDAS2	EU	EU post on eIDAS2
Electronic Signatures and Infrastructures	ETSI	ESI
EUDI Wallet Architecture & Reference Framework	EU	EUDI Wallet Architecture & Reference Framework
GXFSv2 Identity and Access Management	GAIA-X	GXFSv2 Identity and Access Management
IAM - 001 Identity of Member Companies	Catena-X	IAM - 001 Identity of Member Companies
IAM - 002 Identity of Employees and Technical Users	Catena-X	IAM - 002 Identity of Employees and Technical Users
IAM - 003 IAM & Access Control Paradigm	Catena-X	IAM - 003 IAM & Access Control Paradigm
OAuth 2.0 Authorization Server Metadata	Internet Engineering Task Force (IETF)	OAuth 2.0 Authorization Server Meta Data
OpenID Connect	OpenID foundation	OpenID Connect
OpenID Connect for Verifiable Presentations (OIDC4VP)	OpenID foundation	OpenID Connect for Verifiable Presentations (OIDC4VP)
OpenID for Verifiable Credential Issuance	OpenID foundation	OpenID for Verifiable Credential Issuance
Organizational Credential Manager	GXFS	Organization Credential Manager
Personal Credential Manager	GXFS	Personal Credential Manager
SAML 2.0	Internet Engineering Task Force (IETF)	SAML 2.0
Self-Issued OpenID Provider v2	OpenID foundation	Self-Issued OpenID Provider v2
Verifiable Credentials Data Model	W3C	Verifiable Credentials Data Model

Recommendations based on the desk research and the interviews performed to other data spaces initiatives are as follows:

- Verifiable Credentials Data Model: Verifiable Credentials Data Model v1.1 (w3.org)

- DID resolution: Decentralized Identifier Resolution (DID Resolution) v0.3 (w3c-ccg.github.io)
- eIDAS/eID: eIDAS Regulation | Shaping Europe's digital future (europa.eu)

3.6 Access & Usage Policies and Control

This function ensures that data access and usage policies, defined as part of the terms and conditions established when data resources or services are published or negotiated between providers and consumers, are enforced. Data providers typically implement data access control mechanisms to prevent misuse of resources, while data consumers typically implement data usage control mechanisms to prevent misuse of data. In complex data value chains, prosumers combine both mechanisms. Access control and usage control depend on identification and authentication.

Data access control mechanisms can include measures such as user authentication, authorization, and encryption. These measures help ensure that only authorized users can access the data and that the data is protected during transmission. Data usage control mechanisms can include measures such as digital rights management, watermarking, and auditing. These measures help ensure that the data is used in accordance with the terms and conditions established by the provider.

In summary, this function plays a crucial role in ensuring that data is used responsibly and in accordance with established policies. By implementing robust access control and usage control mechanisms, providers and consumers can work together to prevent misuse of data and protect the rights of all parties involved.

Description	Publisher	Link to further information
Authentication & Authorization	GXFS	Authentication & Authorization
Dynamic Attribute Provisioning Service (DAPS)	IDSA	DAPS
Extensible Access Control Markup Language (XACML)	OASIS Open	Extensible Access Control Markup Language (XACML)
GXFSv2 Identity and Access Management	GAIA-X	GXFSv2 Identity and Access Management
IAM - 003 IAM & Access Control Paradigm	Catena-X	IAM - 003 IAM & Access Control Paradigm
IAM - 006 Company Role by the Connector	Catena-X	IAM - 006 Company Role by the Connector
IDS Connector	IDSA	International-Data-Spaces-Association/IDS-RAM_4_0

ODRL Information Model	W3C	ODRL Information Model 2.2
SAML 2.0	Internet Engineering Task Force (IETF)	SAML 2.0
SOV-001 Eclipse Data Space Connector (EDC)	Catena-X	SOV-001 Eclipse Data Space Connector (EDC)
User-Managed Access (UMA) 2.0		User-Managed Access (UMA) 2.0

Recommendations based on the desk research and the interviews performed to other data spaces initiatives are as follows:

- IDS connector: IDS Components - International Data Spaces
- Eclipse Data Space Connector: Eclipse Dataspace Components | projects.eclipse.org
- Dynamic Attribute Provisioning Service (DAPS): IDS-G/Components/IdentityProvider/DAPS/README.md at main · International-Data-Spaces-Association/IDS-G · GitHub

3.7 Trust

Trust is a fundamental issue for data space deployment, and it is built from the moment of onboarding, materializes when transferring the data and is confirmed by checking the monitoring and auditing capabilities of the data space.

This section focuses on the onboarding process and how it contributes to generate trust among the data space participants.

Both GAIA-X and IDS define an onboarding process that in the end provides electronic proofs of the characteristics of the participants, services and resources taken part of the data space. This onboarding process is based on the so-called (according to GAIA-X terminology) “Trust Framework” which is defined by the Data Space Authority.

According to GAIA-X and DSSC, the purpose of developing a Trust Framework is to define and enable the enforcement of a common set of rules that every participant and service in an ecosystem must be verifiable against.

Furthermore, the definition of a Trust Framework at a higher level than the data space level results in a common governance that provides the basic level of interoperability across individual data spaces, while letting the users in full control of their choices.

According to DSSC, a fundamental goal in defining a Trust Framework is to provide transparency, by measuring and comparing the legal, technical, and operational autonomy levels of services with regard to service composability, service characteristics, compliance to

existing standards, and portability and interoperability capabilities – covering licenses, workloads, and data.

In the case of IDS, the overall 'Onboarding' process requires of two preparational steps required for an organization to act as Data Provider or Data Consumer in the International Data Spaces:

1. Registration and certification of the organization.
2. Acquiring a certified IDS connector.

As a result of the onboarding process, both the organization and the IDS connector receive a X509 digital certificate approved by the IDS certification body.

The whole IDS certification process is described in the following web site: https://internationaldataspaces.org/offers/certification.

Based on those prerequisites, an organization can instantiate an arbitrary number of IDS connector instances with the following steps:

1. Provisioning and configuring the connector.
2. Availability setup

GAIA-X Trust Framework comprises the definition of both high-level and low-level requirements, their implementation, and the related extension models. The rules in the Trust Framework apply to Gaia-X credentials, which exist for all the entities that are part of the Gaia-X Conceptual model (mainly participants, service offerings, resources).

From the standards point of view, the operationalisation of the Gaia-X Trust Framework is based on Linked Data stored in W3C Verifiable Credentials[25] documents. Each Verifiable Credential can contain several types of claims and is signed using a JSON Web Key. Each type of claim is defined by a specific controlled vocabulary and a list of Gaia-X approved Trust Anchors.

The Gaia-X Trust Anchors are defined as bodies, parties, i.e. CABs or technical means accredited by the bodies of the Gaia-X Association to be trustworthy anchors in the cryptographic chain of keypairs used to digitally sign statements or claims about an object. The possibility to retrieve trusted information in a machine-readable format is a prerequisite for federating trusted statements within the Gaia-X Ecosystem and developing mechanisms to re-assess the validity of claims.

More in detail, the implementation of the Gaia-X Trust Framework ensures the minimum level of compatibility in terms of:

- serialization format and syntax.

[25] Verifiable Credentials Data Model v1.1 (w3.org)

- cryptographic signature validation and validation of the keypair associated identity.
- attribute value consistency.
- attribute veracity verification

Two types of standards must be considered in the context of trust generation within the onboarding processes:

- Standards to describe the entities participating in the data spaces and its characteristics.
  - **GAIA-X Self-description**[26] models: Participant[27], Service[28], Resource[29]
  - **DID** resolution: Decentralized Identifier Resolution (DID Resolution) v0.3 (w3c-ccg.github.io)
  - **DCAT**: Data Catalog Vocabulary: DCAT version 3[30] (for more information see next section)
  - **JSON-LD**[31]: linked data format
  - **SHACL**: Shapes Constraint Language: a language for validating RDF graphs against a set of conditions
- Standards to provide electronic proofs about the information included in the descriptions.
  - W3C Verifiable Credentials

3.8 Data, Services and Offerings descriptions

Data Spaces infrastructure is decentralised, and metadata and discovery protocols are used to publish data resources and services so that data consumers can discover available data providers. Data providers publish self-descriptions in centralised data catalogues. The descriptions may be domain-agnostic to enable communication between data spaces, or domain-specific to meet special requirements needed for communication within data spaces.

The W3C Data Catalog Vocabulary (DCAT) is a widely used standard that specifies a vocabulary for describing data sets, their distribution, data access, licensing and

[26] SSI_Self_Description_EN_V3.pdf (gaia-x.eu)
[27] Participant - Gaia-X Trust Framework - main version (fb420580)
[28] Service & Subclasses - Gaia-X Trust Framework - main version (fb420580)
[29] Resource & Subclasses - Gaia-X Trust Framework - main version (fb420580)
[30] Data Catalog Vocabulary (DCAT) - Version 3 (w3.org)
[31] JSON-LD - JSON for Linking Data

provenance. The W3C is already working on a new version 3, which will revise the DCAT standard and provide support for describing data set series and versioning of resources. DCAT is a RDF vocabulary but makes extensive use of terms from other vocabularies such as Dublin Core. The DCAT vocabulary has been extended as DCAT profiles for application in catalogues in different domains. The DCAT Application Profile for data portals in Europe (DCAT-AP) is a specification based on W3C's Data Catalogue vocabulary (DCAT). The DCAT-AP is further extended by country specific vocabularies such as DCAT-AP-IT (Italy), DCAT-AP-NO (Norway), DCAT-BE (Belgium), DCAT-AP-SE (Sweden) and DCAT-AP.de (Germany).

To date, no metadata and discovery protocols and no DCAT profiles have been defined specifically for the tourism sector. Although there are discussions about whether DCAT profiles should be eschewed in favour of requests for changes to the DCAT standard, and how the DCAT standard might evolve to meet all the requirements for data space data catalogue frames, there is probably no alternative approach or standard with comparable characteristics and widespread use.

For the exchange of statistical data and metadata between international organisations, the Statistical Data and Metadata eXchange (SDMX) standard can be used to modernise the mechanisms and processes related to sharing statistical tourism data.

Vocabulary of Interlinked Datasets (VoID) is an RDFS ontology to describe RDF datasets. It was developed by a W3C interest group in 2011. It allows description of RDF datasets from different aspects such as general metadata about the dataset based on DCAT, access metadata (about SPARQL endpoints, RDF dumps), structural metadata (e.g., example triples) and link metadata (to which RDF datasets the dataset is linked).

Metadata Models of component descriptions of Data Spaces are often derived from the W3C Resource Description Framework (RDF) Schema and are described according to Web Ontology Language (OWL) standards.

3.9 Marketplaces & Usage Accounting

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

According to the IDS Glossary[32], the implementation of the data usage accounting building block leads to the set-up of a Clearing house feature: "[Clearing House] Intermediary providing clearing and settlement services for all financial and data exchange transactions within the International Data Spaces."

The DSBA technology convergence study suggests a Decentralized Open Marketplace Ecosystem (DOME) based on TM Forum Open API specifications. A DOME is thought to

32 https://github.com/International-Data-Spaces-Association/IDS-G/blob/main/Glossary/README.md#clearing-house

support the concept of a product catalog, which comprises different elements including the description of the usage conditions of the offered data or service products. Each time a resource or service associated with a product is used, a log entry is created, which is read by an accounting module to analyse the transaction and trigger consequences (e.g. calculate how much can be charged to customers and paid to providers).

The OpenDEI "State of the Art" report focuses on the FIWARE Business Application Ecosystem (FIWARE BAE) as a mature solution that provides Accounting functions which meet the requirements formulated by DSBA.

The FIWARE BAE is a collection of different API modules which collaborate to provide the complete commercial functionality. In detail it comprises in the following components[33]:

- 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.
- Logic Proxy

The Authentication module deserves particular attention, since accounting and clearing strongly depend on the identification of trusted parties: it must always be possible to identify the actors of each transaction. Here is a clear interdependence with the Trust: Identity Management Building Block.

Other possible implementations of the Usage Accounting Building Block are provided by IDS, and by the GAIA-X Federated Services (GXFS, soon in XFSC) project.

The clearing house in the Smart Connected Supplier Network (SCSN) has been set up using the IDS technology. It is indeed used to send purchase-to-pay information in a business-to-business scenario. This information can be highly confidential, and it is often mission critical for the day-to-day business of the participants involved. If a conflict arises the clearing house logs are used as a trusted neutral entity to clarify the issue by comparing the fingerprint of the messages and identifying the error.

The GXFS Clearing House feature is not yet implemented in any Data Space initiative (it seems to be used by Themis-X and is under investigation in EONA-X). Moreover, the GXFS project finishes by the end of 2023, its assets developed so far will be maintained in future

[33] See https://github.com/FIWARE-TMForum/Business-API-Ecosystem

on GitLab in the repository of the Cross Federation Services Components (XFSC)[34]. There is a certain risk that no seamless continuation of the work will be achieved, thus support for these modules may be lacking from a mid-term perspective.

3.10 Publication and Discovery

Once individual self-descriptions have been created, it should be possible to publish them in a catalogue and enable other participants of the data space to find them. This is the scope of the building block for Publication & Discovery.

This building block facilitates the publication, distribution, and commercialization of data within the Data Space ecosystem and enables data providers to share their data assets and data consumers to access and utilize them effectively.

Thus, the data space for tourism, by implementing the Publication & Discovery building block, could promote data sharing, and could encourage collaboration between data providers and consumers, and finally could create economic opportunities around data assets. It would also enable data providers to reach a wider audience, monetize their data, and foster innovation through the utilization of valuable data resources by data consumers.

The publication and marketplace services building block should take into consideration the following elements: API, Clearing House, App Store, Data Contract Service.

The Data Space may offer APIs (Application Programming Interfaces) that enable programmatic access to the data. In particular, the TM Forum[35] **Open APIs** are a suite of application programming interfaces that:

I) Enable services to be managed end-to-end throughout their lifecycle.

II) Work in an environment where multiple partners are involved in service delivery. The APIs allow simple and coherent management of any given service. Their rapid implementation is supported by key industry-strength design patterns. TM Forum's Open APIs are based on representational state transfer (REST). They are technology agnostic and can be used in any digital service scenario, including Data Spaces.

OpenAPI is among the most well-known and widely used. While OpenAPI can be considered the industry standard, in the end, companies often pick the format that best suits their business needs.

In this context, it is interesting to highlight that the experience of the Italian Minister of Tourism within its Tourism Digital Hub shows that Publication and Marketplace Services can be a cornerstone building block of an interoperability service. The main focus of the digital

[34] See https://gitlab.eclipse.org/eclipse/xfsc

[35] Open APIs - TM Forum

ecosystem, called "TDH22" which stands for "Tourism Digital Hub '22", is to gather tourism data from various organizations and to build from it a full data lake from which extract both relevant insights and front-end data to be shown in the Tourism Digital Hub channels: sites, mobile apps, digital email marketing, chatbots, call centres.

While **Open API Table** project created and maintained by TM Forum is a great and wide experience of interoperability between open services that needs to share not only data, but also ways to operate with data, TDH22 ecosystem's priority has been to maintain high quality standards on data shared between the ecosystem and the necessity of reviewing such data. In fact, it is important in a tourist, public, government-led ecosystem to gather access to such APIs in a controlled API table. Thus, the decision taken on TDH22 has been to give access to APIs only to subscribers of the service so as to avoid any disruption in the quality of the data collected. It has also to be considered that TDH22 APIs exposed do not offer data owned by the Ministry of Tourism; instead, they allow external subscribers to provide data to be published in the ecosystem channels. Data Space should consider this when dealing with front-end data that must be published only when it meets certain high-standard requirements. A **Complete API Catalogue** is one of the most important building blocks, as it contains all the developed APIs. Such Catalogue is comprehensive of all the APIs in the ecosystem, fully comprehensible by the adherents of the ecosystem by providing specifications, examples, code snippets. According to the Open API building block, the access is granted only after the completion of a certification path.

Still regarding the API ecosystem, another important project that can be considered for the construction of the data space, is the business API ecosystem part of FIWARE, made in collaboration with the TM Forum. The Business API Ecosystem is a joint component made up of the FIWARE Business Framework and a set of APIs (and its reference implementations) provided by the TM Forum. This component allows the monetization of different kind of assets (both digital and physical) during the whole service life cycle, from offering creation to its charging, accounting and revenue settlement and sharing.

Another element of the publication and marketplace services building block is the clearing house service. A clearing house service acts as a central hub that promotes data integration, quality assurance, security, interoperability, streamlined transactions, and governance within a data space. It enhances the trustworthiness, reliability, and value of the data ecosystem by providing essential services and infrastructure for data management and exchange. The **IDS Clearing House Service**[36] is a prototype implementation of the Clearing House component of the Industrial Data Space. Data in the Clearing House is stored encrypted and practically immutable. A single valid receipt in possession of any connector is enough to detect any change to data up to the time indicated in the receipt. **Clearing House** building block as developed by IDS considers that any exchange in data between a Data Provider and a Data Consumer should be mediated by a trustworthy layer capable of restricting data according to the business rules defined by both parties. In the case of TDH,

[36] Clearing House | Dataspace Connector

the TDH022 module acts as a mandatory middleware layer, providing mainly technical clearing functions.

Additionally, the **App store** represents an additional element that can be considered for a Data Space for tourism. The International Data Spaces App Store is a secure platform for the distribution of data apps. It offers various search options, for example by functional or non-functional characteristics, the pricing model, certification status or community reviews. An International Data Spaces App Store consists of registration for available data apps in this App Store. Therefore, an App Store supports operations for the registration, publication, maintenance and querying of data apps as well as operations for providing a data app for a connector. These basic operations can also be supplemented by additional services, such as billing or support services.

The Data Space must provide a formal data transaction initiation handshake between the data provider and the data consumer. As regards to the data transaction a relevant example is the "**Data Contract Service**" that enables data transactions in a secure, trusted, and auditable way within the **Gaia-X ecosystem**. It offers interfaces for the negotiation of data contracts detailing the agreed terms (Data Asset Usage Policy) for planned data exchange. The service is not meant to handle the transaction of data (that is described in the negotiated data contracts). Basically, the service allows sending offers and counteroffers resulting in agreement or rejection.

A second meanwhile frequently used contract service building block is the contract closing feature of the **Eclipse Dataspace Components**. Due to common interfaces and mapping of existing standards, the EDC adds capabilities of contract negotiating and policy handling in an interoperable manner.

4 ESTABLISHED STANDARDS FACILITATING DATA SPACE GOVERNANCE

Governance is mainly an organizational challenge for organizations. It consists essentially in a body of policies, processes and methods guiding the organization to making the right decisions to master the challenges against all aspects of its daily life, and in the interactions with external stakeholders (authorities, other data spaces, boards on overarching topics, prospects).

The elaboration of a governance framework for the future tourism data space will be the topic of task 4.2 and it will be described in the deliverable D4.2 “Governance framework of the Tourism Data Space”. In the following chapter, we would therefore like to focus on the available technical support for the governance building blocks still using the categorization of the OpenDEI document, since they addressed indeed topics that could potentially be supported by technical means. However, given the organizational character of the governance pillar, best practices regarding technical support are poor, either from the perspective of the results from our interviews or from desk research. It may be also due to the fact that there are currently no “true” data spaces existing in Tourism (besides a lot of legacy data sharing initiatives) that used standards are hardly found.

4.1 Overarching Cooperation Agreements

In the position paper of the Open DEI project the building block regarding internal agreements is defined as follows: “All data space participants need to agree on certain functional, technical, operational and legal aspects. While some agreements are reusable in a generic or sector-specific way (e. g. rule books), others are use-case specific.” Thus, a concept published by aNewGovernance, which proposes indeed “Rulebooks and Rolebooks” for different layers of scope and responsibility within a single data space environment, but also beyond that considering inter-data-space governance, might reasonably serve as a tool to support the setup of the respective governance frameworks.

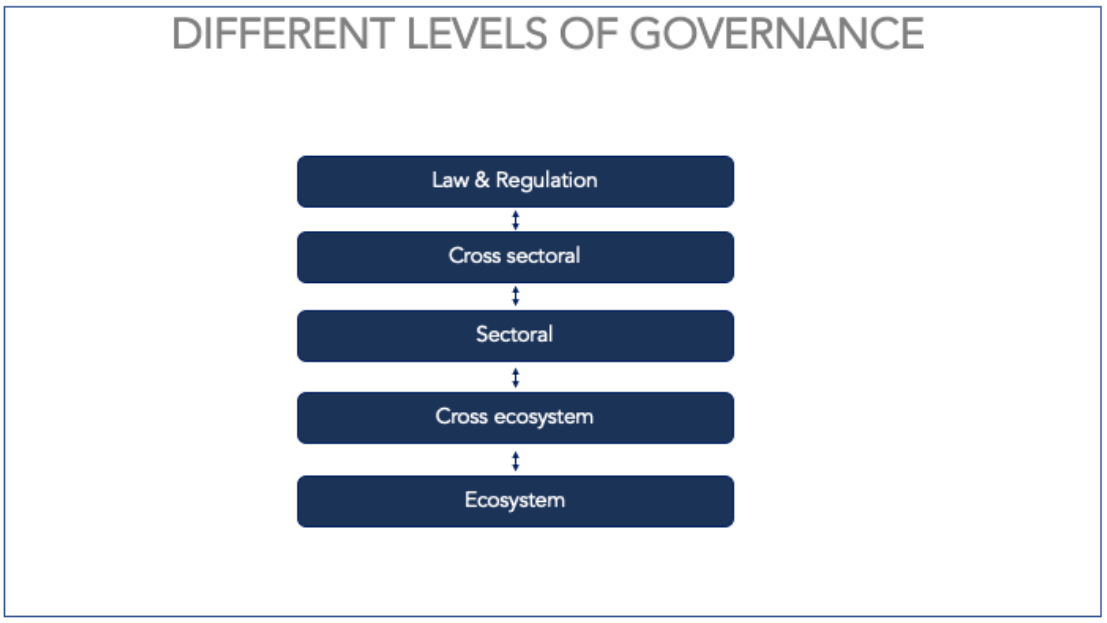

Figure 3 Illustration of the different levels of governance, proposed by aNewGovernance[37]

Generic Rulebooks for data spaces have been published by SITRA ("Rulebook for a fair data economy")[38], and by IDSA ("IDSA Rulebook Version 2")[39].

The SITRA document pretends to be "a guide for creators of fair data economy networks. Agreement templates and other tools make it easier to build and join new data networks which highlight transparency in data sharing."

It basically provides:

1. templates for legal, business, technical and administrative rules
2. control questions to validate the implemented rule
3. templates for an internal code of conduct document

The objective of the IDSA rulebook is to explain in detail which rules are mandatory and which are optional guidelines. Their governance framework includes "functional, technical, operational, and legal dimensions", it supports the governance over a data space especially under the following viewpoints:

1. Description of common services, and the definition, processes, and services of specific roles.
2. Guide to the implementation or use of technical artefacts of the IDSA.
3. Rules for the collaboration within data services

[37] https://www.anewgovernance.org/1241-2/

[38] https://www.sitra.fi/en/publications/rulebook-for-a-fair-data-economy/

[39] https://docs.internationaldataspaces.org/ids-knowledgebase/v/idsa-rulebook/front-matter/readme

4. Managing the legal basis of a data space in compliance with the regulatory environment to ensure trust and security.

A concrete instantiation of a governance framework of an operative data space can be seen on the website of the Catena-X association. See there the Catena-X charter[40] as their basic set of overarching agreements, and several documents explaining the different aspects of the Catena-X governance framework[41].

4.2 Operational Governance

Best support is available for some operational processes. Especially the management of users has already been covered by different organizations with respect to the definition of processes, and already with respect to supporting software.

A frequently used bundle are the iShare operational processes. This mini frame work consists currently of six modules:

- Member Admission: a process to onboard new members
- Member Withdrawal: a process to decommission members
- Warnings, Suspension and Exclusion: "The warnings, suspension and exclusion process describe the steps that the Scheme Owner MUST take to temporarily suspend or permanently exclude participating parties from the iSHARE network in case of non-compliance with scheme rules and guidelines, or actions with significant negative impact on the normal operation of the iSHARE network."
- Incident Management
- Release Management
- Management reporting : "The management reporting process describes the steps that parties MUST take to deliver management information about the use and working of the iSHARE network. The goal of the management reporting process is to monitor compliance to service level agreements, and to distribute info about the use of the iSHARE network."

The schema owner in the iShare terminology is approximately equivalent to the governance body of the DSSC glossary. It means the role (individuals or boards) in charge of managing the administration of the community of data space members.

A second example of practical relevance are the "Onboarding & Accreditation Workflows" published by the GAIA-X federated services project (GXFS). It is based on the GXFS

[40] This document "Articles of the association" is still accompanied by some other fundamental information on the site: https://catena-x.net/de/ueber-uns/organisation-des-vereins (Organizational principles of the association).

[41] https://catena-x.net/de/catena-x-einfuehren-umsetzen/governance-framework-fuer-datenraum-betrieb

Workflow Engine, which is designed to implement and configure onboarding and Accreditation workflows for the Gaia-X ecosystem and runs under Apache Version 2 license.

The automotive supply chain data space project Catena-X has published a set of specifications about three aspects of an onboarding process. First, the initial onboarding of a new participant; second, the registration of a data set, and a guideline to the standards which must be met by a provider of a services or business app. The specifications are quite detailed and are supposed to be implemented in an IDS/EDC based infrastructure. However, the specifications are sufficiently generic to allow for the implementation in other, somehow customizable systems, or in systems set up from scratch.

ONB-001 Registration and initial onboarding: "To become a participant of Catena-X, each applicant must go through a registration process. Registration is a mandatory requirement for all further activities within the Catena-X network. The registration process, along with other services, provide the foundation of trust for the Catena-X network. This document describes the structure of this process and defines the results of the individual process steps."

ONB-002 Minimal data provider services offering: "The purpose of this document is to describe the necessary functional building blocks of a minimal data provider service, as well as to give an overview over the interplay of necessary standards that need to be considered. The document also gives an overview over the mandatory standards that need to be integrated. Service offerings for data providers that do not fulfil the necessary standards and concepts in this document, will not work in the Catena-X network and thus cannot be provided to any customer. In the following chapters, the minimal data provider service will also be referred to as "upload tool" for better understandability."

ONB-003 Relevant standards for conformity assessments: "It lists the internal (e.g. CX Guidelines) and external (e.g. ISO) standards that a partner needs to fulfil before he/she can offer solutions (services or business apps) in the context of Catena-X. Those standards have been derived from common incoterms in the automotive industry and present the minimal basis so that software can be used within an automotive industry enterprise. Requesting certain standards from partners secures that all customers that use solutions (services or business apps) are guaranteed a certain level of professionality, security, and trust." For a Reference implementation see: https://gitlab.eclipse.org/eclipse/xfsc/oaw

4.3 Continuity Model

In the Open DEI model, Change Management is the central element of the Building Bock "Continuity Model". This is a term for all processes to prepare, support, and help organizations in executing the required changes, to continue with the organization under changing internal and external conditions. It provides methods that redirect or redefine the use of resources, business processes, or other modes of operation that significantly change the organization. In the context of a data space, again all existing building blocks, regarding processes and infrastructure can be affected by needs of change. Therefore, it is important

to implement the respective policies and procedures, describing (non-exhaustive) how to initiate a change process, how to execute stakeholder management (all relevant parties shall be included from the beginning), description of the target state, the roadmap, risks(!), decommissioning of outphased elements.

A widely accepted standard for a change management system is ISO 10020[42]. For the design of concrete activities of a change process the iSHARE Trust Framework's Release Management process could serve as a template. Continuity planning in all domains of a data space administration could follow the Plan-Do-Check-Act (PDCA) cycle, which prescribes recurring reviews of all processes and assets executing a target-performance comparison. Setting new targets (e.g. caused by the identification of innovation needs) would therefore initiate the change management process, as well as problems with any type of expected performance.

4.4 Business Agreements

According to the Open DEI state of the art document, business agreements include service level agreement (SLA), data usage and access control policies as well as specified standards. They are artifacts that regulate the business relationships between the different roles in a data-driven business ecosystem: data owner, data acquirer or data provider, data processor, data marketplace operator. For possible implementations of data usage and access control tools, see the findings described in chapter 2.6.
Support for the management of service provision can be obtained from the standard ISO/IEC 20000-1:2018 "Information technology — Service management — Part 1: Service management system requirements"[43], and of course from the other parts of ISO/IEC 20000.

Deliverables 4.3 and D4.4 will go more in deep in the different business models for the European Tourism data Space

The aspect of agreed specified standards is somehow self-explaining. Each group of parties, which want to collaborate in operating a data space, need to agree wherever possible on identified best practices and standards, for the sake of streamlining the organizational collaboration and for facilitating technical interoperability. In fact, the chapters above and the other deliverables of this CSA project are meant to give the necessary recommendations for agreements on standards. It is then a matter of good governance processes to assess the real requirements of the individual data space community and guide this community to the necessary decisions about the standards which will help to master the identified challenges.

[42] https://www.iso.org/standard/82213.html
[43] https://www.iso.org/standard/70636.html

5 STANDARDS SUPPORTING OTHER OPEN DEI BUILDING BLOCKS

The Open DEI position paper[44] mentions on page 52 some "additional technical building blocks" which "may optionally be considered to facilitate connection of additional systems to data space". They provide hints about functions which will enhance the experience of the participants of a data space by facilitating interoperability, and moreover monitoring and control of the participants´ activities.

5.1 Other Technical Building Blocks I: System Adaptation

The scope of the System Adaptation building block is to interface with the various data resources exported by the system and to perform the necessary transformation of the data formats adopted for data exchange within the data space.

The interface for data exchange depends on the nature of the system. In a tourism data space, relevant systems can be IoT resources, databases, enterprise systems (CRS, PMS, QPX) and applications, etc.

IoT protocols	CoAP (Constrained Application Protocol) MQTT (Message Queuing Telemetry Transport) DDS (Data Distribution Service)
Databases or triplestores	JDBC (Java Database Connectivity) SQL (Structured Query Language) SPARQL
Generic API protocols	RESTful services
Tourism API protocols	API conventions and standards from OpenTravel Alliance (OTA) and the OpenAPI Initiative (OAI)
Data transformation	RML (RDF Mapping Language) XML transformation languages (XSLT, XQuery) JSON transformation languages (JSLT)

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

Tourism legacy standards	AVLABL electronic data interchange message format for travel providers to distribute pricing and availability data (used widely within the UK market). REScon (Reservation Confirmation): recommended format for travel agent's systems to capture booking information directly from the tour operator. TOPAS standard for the transmission of availability data between tour operators and travel agents or intermediaries. TORIX messages created by a TTI working party of major tour operators to provide a structured alternative to the legacy Viewdata distribution channel for booking packaged holidays.

5.2 Other Technical Building Blocks II: Data Analytics Engine

A Data Analytics Engine supports execution of data analytics with regard to data shared and exchanged over the Data Space. Many Data Space use cases make this analysis possible through methods such as statistical analysis, machine learning, deep learning, and other data-mining techniques. In the context of a Data Space for Tourism, a Data Analytics Engine refers to the software or system responsible for processing and analyzing the data related to tourism. It enables extracting insights, identifying trends, and making data-driven decisions to improve tourism offerings, customer experiences, and business operations.

Depending on the nature of the data, this Building Block can take different forms, such as cloud-based analytics, streaming analytics or machine learning.

- Cloud-based analytics take advantage of cloud computing infrastructure and services to perform data analytics tasks. While there are no specific standards dedicated to cloud-based analytics, there are several widely adopted standards and frameworks that are relevant in this context:

Analytics and Data Processing standards	❖ **Hadoop and Apache Spark**, two open-source standard frameworks to provide distributed data processing capabilities for big data analytics in cloud environments. ❖ **PMML (Predictive Model Markup Language)**,

	an XML-based standard for representing predictive analytics models, allowing interoperability between different analytics platforms. ❖ **SQL (Structured Query Language)**, a widely adopted standard for querying and manipulating structured data.
Cloud Computing standards	CNCF (Cloud Native Computing Foundation) standards: ● **Kubernetes**, a standard framework to deploy and operate containerized applications. ● **CRI (Container Runtime Interface)**, a standard interface between Kubernetes and container runtimes.

- *Streaming analytics*, also known as real-time analytics or event stream processing, is the practice of analyzing and deriving insights from data streams as they are generated. There are various industry-standard frameworks and technologies commonly used in this domain:
  - **Apache Kafka** is a widely adopted standard streaming framework that provides a publish-subscribe messaging system capable of handling high-volume, real-time data streams to build streaming analytics pipelines.
  - The **Streaming SQL standard** or the **SQL-92** extensions for streaming data provide a standardized way to express real-time data processing and analytics using SQL-like queries.
  - **OpenMessaging**, an industry standard for messaging and streaming (although it doesn't specifically focus on analytics).
- *Machine learning* involves training models on data to make predictions or decisions. There are several widely recognized standards and frameworks that govern various aspects of the machine learning lifecycle:

Model Development and Deployment	❖ **TensorFlow**, an open-source machine learning framework developed by Google. ❖ **PyTorch**, an open-source machine learning library with strong focus on deep learning models. ❖ **Scikit-learn**, a popular machine learning library that provides a wide range of machine learning algorithms.
Ethical Considerations and Fairness	❖ **Microsoft Responsible AI Standard**, a framework to guide on how to build trustworthy AI systems.

	❖ **Google Responsible AI practices**, which guides the conduct of risk assessments and fairness evaluations.
Model Evaluation and Validation	❖ **CRISP-DM (Cross-Industry Standard Process for Data Mining)**, a widely adopted standard for the entire data mining process, including machine learning model development.

5.3 Other Technical Building Blocks III: Data Visualization

Data Visualization Building Blocks provide data presentation and visualisation features. These features are a key issue for any big data project and, also, for data spaces. These graphical representations can take different forms such as: charts, geographic maps, sparklines, infographics, heat maps, statistical graphs …

The goal of these visual representations is the identification of actionable insights for all relevant stakeholders to make efficient decisions.

Data visualization tools facilitate the visual representation of data and most of them incorporate self-service capabilities so that users can easily compose automated applications. This combination of capabilities is helping to broaden the view of analytics beyond the simply delivering of data sets and presentation of dashboards.

Nowadays, these visualization tools are included in more powerful Business Analytics Tools. There are plenty of this kind of tools in the market from different vendors[45]. Gartner provides every year a Magic Quadrant[46] that shows the market evolution of these tools[47]:

[45] See https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms
[46] See https://www.gartner.es/es/metodologias/magic-quadrant
[47] See https://www.gartner.com/doc/reprints?id=1-2955ETOT&ct=220215&st=sb

Figure 4 Magic Quadrant for Analytics and Business Intelligence Platforms

Based on this report, there are three market leaders: Power Bi (Microsoft), Tableau (Salesforce) and Qlik. Power BI has significant market reach through integration with Microsoft 365, Azure and Teams.

Additionally, there are smaller players in this market that are not included in this Quadrant, such as Metabase, Kibana or Grafana.

On the other hand, and in the field of standards, the "Data Visualization Style Guides" should be taken into account. These guides are standards, usually in the scope of the organization, for formatting and designing representations of information. They include, for example, types of charts, logos, brand colours, layout representation or tone of language. Special attention must be taken about accessibility facilities.

The Data Visualization Society[48] is a good initiative to compile these guides and encourage discussion about good practices.

[48] See https://www.datavisualizationsociety.org/

5.4 Other Technical Building Blocks IV: Workflow Management Engine

Interactions of data sources, data consumers, and data services must be properly orchestrated by means of structured and acknowledged workflows. Workflows can be divided into three primary types, based on their complexity:e sequential, state machine, and rules-driven.

Open WDL (Workflow Description)	https://openwdl.org/ (Community driven open-development workflow language)
BPMN (Business Process Modeling Notation)	OMG
XPDL 2.0 (XML Process Definition Language)	WfMC
WS-BPEL (Web Services Business Process Execution Language)	OASIS
TDH022 Orchestration Service	The Orchestration service building block has to be seen as a responsible manager for the provided services, by creating composite APIs that can leverage base data. TDH022 has been designed to provide orchestrated services in a seamless way so for an operator perspective, any request given to the portal maintains its atomicity.

6 CLOSING REMARKS

Taking the findings of this investigation together, it is remarkable that all of the topics of the DSSC taxonomy are supported by concrete standard components (specifications, protocols, procedures, and software), which can be used as building blocks for the implementation of the future European Tourism Data Space.

The topic, that is mainly sector-specific, is that of data models and formats. It can be expected that many data providers will at the beginning offer their data in proprietary and legacy data formats. One of the first tasks of the members of the European Data Space for Tourism might be to agree on one or a few standard models and to guide the members to adopt these models. Where it is not easy to do so, applications might be created which map the data models, and translate the data from one format to another one.

However, the results presented here are still a snapshot. The finalization of the set of recommendable building blocks is still under work in the DSSC, which is of course still collecting inputs from the other CSA projects and from various experts. Their final blueprint might look slightly different to what DATES is presenting in this deliverable.