



DATA SPACES SUPPORT CENTRE

Methodology for the identification of the Blueprint building blocks

Authors: Tuomo Tuikka, Christoph Strnadl, Clara Pezuela, Bert Peeters, Claire Stolwijk, Sumiya Sultana, Jarno Halme, Matteo Frigeri

DATA SPACES SUPPORT CENTRE

Methodology for the identification of the Blueprint building blocks

Document Summary Information

Grant Agreement No	101083412	Acronym	DSSC
Full Title	Data Spaces Support Centre		
Start Date	01/10/2022	Duration	42 months
Project URL	https://dssc.eu		
Deliverable	D4.1		
Work Package	WP4		
Contractual due date	31.3.2025	Actual submission date	31.3.2025
Nature		Dissemination Level	SEN
Lead Beneficiary	VTT		
Responsible Author	Tuomo Tuikka		
Contributions from	VTT, GAIA-X, FIWARE, KU Leuven, TNO, SITRA, IDSA, MyData, BDVA, Fraunhofer, University of Galway		



This project has received funding from the European Union Digital Europe Programme under Grant Agreement № 101083412.

More information available at <https://dssc.eu/>

Revision history (including peer reviewing & quality control)

Version	Date	% Complete	Description	Contributor(s) (Organisation)
0.8	31.1.2025	85	1 st Review release	VTT, GAIA-X, FIWARE, KU Leuven, TNO, SITRA, IDSA, MyData, BDVA, Fraunhofer, University of Galway
1.0	25.2.2025	100	Acceptance review	Project Steering Board
1.1	16.5.2025	100	Minor modifications based on the review comments	VTT

Copyright message

The work described in this document has been conducted within the Data Spaces Support Centre project. This document reflects only the Data Spaces Support Centre Consortium view and the European Union is not responsible for any use that may be made of the information it contains.

This document and its content are the property of the Data Spaces Support Centre Consortium. All rights relevant to this document are determined by the applicable laws. Access to this document does not grant any right or license on the document or its contents. This document or its contents are not to be used or treated in any manner inconsistent with the rights or interests of the Data Spaces Support Centre Consortium or the Partners detriment and are not to be disclosed externally without prior written consent from the Data Spaces Support Centre Partners.

Each Data Spaces Support Centre Partner may use this document in conformity with the Data Spaces Support Centre Consortium Grant Agreement provisions.

Content

1	Introduction	1
1.1	Overall process of DSSC and identification effort	1
1.2	DSSC internal collaboration process	3
1.3	Scope of this document.....	4
1.4	List of results	5
2	Activities and content	7
2.1	Requirements	7
2.1.1	Introduction	7
2.1.2	Methodology	7
2.1.3	Outcomes	11
2.1.4	Conclusions	46
2.1.5	Results	47
2.2	Building Block identification and data space development.....	47
2.2.1	Scope and purpose.....	47
2.2.2	Methodology	48
2.2.3	Contributions to data spaces Blueprint.....	49
2.2.4	Results	54
2.3	Standards landscape	55
2.3.1	Scope and purpose.....	55
2.3.2	Standardization relevance and activities in DSSC	56
2.3.3	The collection of standards landscape in data spaces	64
2.3.4	Liaison activities with Standardization Development Organization (SDOs).....	72
2.3.5	Takeaways for data spaces blueprint	76
2.3.6	Upcoming work	76
2.3.7	Results	76
2.4	Architectures landscape.....	77
2.4.1	Scope and purpose.....	77
2.4.2	Architectures analysis	78
2.4.3	Takeaways for data spaces blueprint	79
2.4.4	Upcoming work	80
2.4.5	Results	80
2.5	Requirements of Demonstrators and test infrastructures	80
2.5.1	Introduction	80
2.5.2	Methodology	81
2.5.3	Data Space Demonstrator	82
2.5.4	Data Space Test Infrastructure.....	84
2.5.5	Upcoming work	86
2.5.6	Results	86
3	Conclusions	87
	Appendices	88

List of Figures

Figure 1: Overall process of DSSC as depicted in the Grant Agreement.....	2
Figure 2: Conceptualization of DSSC internal collaboration process and results and contribution reported in this deliverable mapped to the process.....	3
Figure 3: First version of the DSSC building block taxonomy, i.e. later on 'building blocks overview'.....	9
Figure 4: Designers, development activity and the relationship between co-creation method and building blocks.....	48
Figure 5: Building Block overview in Blueprint 2.0.....	50
Figure 6: Development processes and the relation to the building blocks presented in v1.0 of the blueprint.....	54
Figure 7: Standardization activities in DSSC	56
Figure 8. Standards vs Legislation	56
Figure 9. Usual standardization process.....	57
Figure 10. DSSC focus on standards.....	58
Figure 11. European and international standardization bodies in scope of DSSC	59
Figure 12: DSSC role in standardization process	60
Figure 13: Experience vs regulation.....	61
Figure 14. Timeline for standardization activities in DSSC.....	62
Figure 15. Screenshot of the collection (internal copy).....	65
Figure 16. Asset work plan in 2023.....	68
Figure 17: Collected standards per category.....	69
Figure 18: Collected standards per building block.....	69
Figure 19: Collected standards by the type of bodies and specifications.....	70
Figure 20: Cross-domain vs specific-domain in data spaces initiatives.....	71
Figure 21: Bidirectional information flow supported by the project liaison.....	75
Figure 22: Depicted common pattern from analysed data spaces architectures.....	79

List of Tables

Table 1: Results for D4.1 by dedicated task.....	6
Table 2: Data space initiatives	11
Table 3: Regulations and Legal challenges for Data4Skills	13
Table 4: Regulations and Legal challenges for Green Deal Data Space - GREAT	15
Table 5: Regulations and Legal challenges for Agri Data Space.....	16
Table 6: Regulations and Legal challenges for DS4SSCC	17
Table 7: Participants and roles for DATES.....	19
Table 8: Regulations and Legal challenges for DATES.....	19
Table 9: Participants and roles for int:net.....	20
Table 10: Regulations and Legal challenges for int:net.....	21
Table 11: Participants and roles for Europeana.....	22
Table 12: Regulations and Legal challenges for Europeana	22
Table 13: Regulations and Legal challenges for LDS	23
Table 14: Data Interoperability Requirements – Data4Skills	24
Table 15: Data Sovereignty and Trust – Data4Skills.....	25

Table 16: Data Value Creation – Data4Skills	25
Table 17: Data Interoperability Requirements Green Deal Data Space - GREAT	25
Table 18: Data Sovereignty and Trust Requirements – Green Deal Data Space - GREAT	26
Table 19: Data Value Creation – Green Deal Data Space - GREAT	26
Table 20: Data Interoperability Requirements - AgriDataSpace	27
Table 21: Data Sovereignty and Trust Requirements - AgriDataSpace	28
Table 22: Data Value Creation - AgriDataSpace	29
Table 23: Data Interoperability Requirements – DS4DSSCC.....	30
Table 24: Data Sovereignty and Trust Requirements – DS4DSSCC.....	31
Table 25: Data Value Creation – DS4DSSCC.....	32
Table 26: Data Interoperability Requirements - DATES.....	32
Table 27: Data Sovereignty and Trust Requirements - DATES	33
Table 28: Data Value Creation - DATES.....	34
Table 29: Data Interoperability Requirements – int:net.....	34
Table 30: Data Sovereignty and Trust Requirements – int:net.....	35
Table 31: Data Value Creation – int:net.....	35
Table 32: Data Interoperability Requirements - Europeana.....	37
Table 33: Data Sovereignty and Trust Requirements - Europeana.....	38
Table 34: Data Value Creation - Europeana.....	39
Table 35: Data Interoperability Requirements - LDS	39
Table 36: Data Sovereignty and Trust Requirements - LDS.....	40
Table 37: Data Value Creation - LDS.....	41
Table 38: Gap Analysis Data Space DS4Skills	44
Table 39: Gap Analysis Data Space Green Deal - GREAT.....	44
Table 40: Gap Analysis Data Space AgriData Space.....	44
Table 41: Gap Analysis Data Space DS4DSSCC.....	45
Table 42: Gap Analysis Data Space DATES.....	45
Table 43: Gap Analysis Data Space Int:net.....	45
Table 44: Gap Analysis Data Space Europeana	46
Table 45: Gap Analysis Data Space LDS	46
Table 46: Summary of results of requirements collection.....	47
Table 47: Summary of results of building block identification	55
Table 48: DSSC partner activities in standardisation summary	64
Table 49. Solution Maturity model for BBs.....	67
Table 50. Most endorsed standards by data spaces initiatives contributing to the collection	72
Table 51: Relevant standardisation committees.....	73
Table 52: DSSC contribution in certain standardisation bodies.....	74
Table 53: Summary of results of standardisation.....	77
Table 54: Summary of results of architecture landscape	80
Table 55. Methodology for identifying test infrastructures and demonstrators	81
Table 56: Summary of results of demonstrators and test infrastructures	86

Glossary of terms and abbreviations used

Abbreviation / Term	Description
DSSC	Data Spaces Support Centre
SDO	Standardization Development Organization
EC	European Commission
BB	Building Block
GA	Grant Agreement
BP	Business Process
BLOFT	Business, Legal, Operational, Functional, Technological framework
AI	Artificial Intelligence
SAB	Security Advisory Board
CSA	Coordination and Support Action
DS4SSCC	Data Space for Smart and Sustainable Cities and Communities
LDS	Language Data Space
EU	European Union
GDPR	General Data Protection Regulation
DGA	Data Governance Act (EU Regulation 2022/868 of 30 May 2022 on European data governance)
DA	Data Act (EU Regulation 2023/2854 of 13 December 2023 on harmonised rules on fair access to and use of data)
IDS-RAM	International Data Spaces – Reference Architecture Model
IDSA	International Data Space Association
ICT	Information and Communications Technology
EDIB	European Data Innovation Board
ETSI	European Telecommunications Standards Institute
CEN	European Committee for Standardization
CENELEC	European Committee for Electrotechnical Standardization
FAIR	findability, accessibility, interoperability, and reusability
API	Application Programming interface
EBDVF	European Big Data Value Forum
EG	Expert Group
TG	Thematic Group
TMForum	Telemanagement Forum
DS4.0	Manufacturing Data Space
NGSI-LD	Next Generation Service Interfaces for Linked Data



JSON-LD	JavaScript Object Notation for Linked Data
LDS	Language Data Space
DID	Decentralized Identifier
WG	Working Group
BDVA	Big Data Value Association
DCAT	Data Catalog Vocabulary
RDF	Resource Description Framework
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
PSB	Project Steering Board
PMB	Project Management Board
VC	Verifiable Credentials
TTG	Technical Thematic Group
CoP	Community of Practice
IoT	Internet of Things

1 Introduction

The effort of identifying existing and available instruments and resources for data space development is an art of collaboration with many stakeholders in the community. By stakeholders we mean of course initiatives that are deploying and have the mandate to build the Common European Data Spaces, but then again also previous research projects, advancement in the industry, contribution of associations, European Commission, standardization organizations etc. practically a vast number of sources which have already done tremendous amount of work to establish the basis for technology development. Therefore, the methods of the work chosen to do this endeavour of identification must support multidisciplinary, multisectoral, cross-topic approach allowing to harness the power of the community.

It is not enough, however, to only collect information, but the information must be analysed and understood so that it can be used for a purpose of a customer, i.e. the user of this information. The customer can be external, outside the Data Spaces Support Centre – DSSC project or internal, e.g. other work package of the project. Therefore, the result of this information collection and analysis must serve the end user, who is developing instruments for the data space development or feverishly developing a data space. In other words, the result must lay ground so that it facilitates further either the actual activity of data space development or the next stages progress of tools and instruments, both conceptual and technical, for data space developers.

With this spirit, this deliverable, methodology for the identification of the blueprint building blocks, collects, and delineates a selected set of inputs and results of four identification tasks of Data Spaces Support Centre Work Package 4, which focuses on Data Spaces Identification and Synergies. These tasks are T4.1 Common requirements and common building blocks for data spaces, T4.3 Common standards and open-source technologies for data spaces, T4.5 Collaborative business governance and data governance models for running data spaces, T4.6 legal issues and market barriers. Especially, these tasks contribute to the identification of candidate building blocks but also support in development of corresponding content of data spaces building blocks. The outcomes described in this document form the grounds for the blueprint building blocks published in the DSSC internet page.

1.1 Overall process of DSSC and identification effort

The most important point for identification effort is to collect relevant information and translate that into meaningful observations and conceptual understanding. Thus, to start with, DSSC draws from previous research project work and from related data space development. For example, OpenDEI project provided the first framing of data spaces building blocks, i.e. a representation of technical building blocks, and definition on them¹. Another example could be the development of verifiable credentials, means to identify data space participants. This is an example from the genre of technology development, which is the basis for future implementation of data spaces.

¹ [Design Principles for Data Spaces | Position Paper](#)

The current set of building blocks include many technological options and can be found in the building blocks of Blueprint². This document refers to multiple versions of the blueprint, namely [blueprint 0.5](#), [blueprint 1.0](#) and [blueprint 1.5](#). Blueprint 2.0 will be published in parallel with this document as the [latest version of the blueprint](#).

Progress in data spaces community brings another aspect to the identification effort. It may be that some solutions are still not matured and are not really chosen as a common way of working. There may be several candidate approaches, but which one to choose as the right choice may not be clear. Our assumption has been that, as the territory has not yet been certain, good choices are collected equally and listed for evaluation, either for next steps at DSSC process or for the community. Constant progress requires also updates in the materials.

This said, emphasis on the identification work has been in the early phases of the project and thus serving later starting work packages which use the results. To be more precise, this work has been following the heartbeat of DSSC with a milestone (Month 12) synchronizing the identification of candidate building blocks with the first version 0.5 of the Blueprint. Figure 1 depicts the overall process as a timeline and presents milestone of month 12 positioning the considerable identification work effort in the first year of the project.

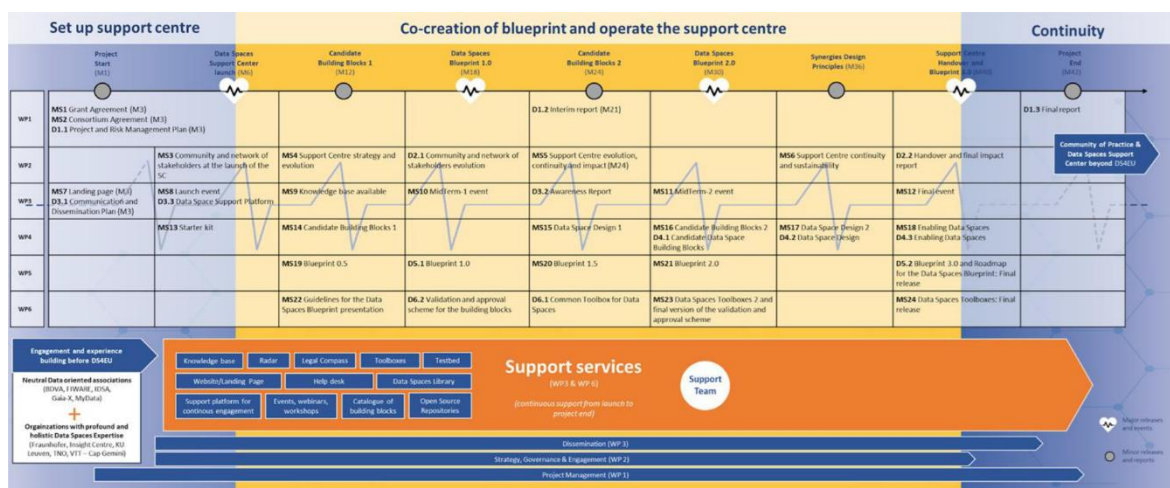


Figure 1: Overall process of DSSC as depicted in the Grant Agreement.

Heartbeat expresses how different milestone results and work are inter-related and the checkpoints for this interconnection.

WP4 internals looked like this: WP4 started immediately after DSSC project start in October 2022 (Figure 1). Two milestones were set during M1-M18: Starter kit (M6), and **Candidate building blocks 1 (M12)**. The milestone for Candidate building blocks 1 (M12) was to produce a set of the candidate building blocks for the Blueprint and collection of Standards related to data space development. Furthermore, a set of requirements for data spaces were collected (M12) from community of practice for DSSC to serve the first gap analysis in WP5 for building block identification (M18). An internal year report was made to support the deliverable process (M12). The next step is this document which gathers both the results until Month 24, and effort which has been made until Month 30, for the deadline of this deliverable. Some of the materials

² [Data Spaces Support Centre](#)

have been released earlier, these are listed with links in the result list and information they have been published and if not published otherwise, they are included within this deliverable as a D4.1 report.

Especially progress on candidate building blocks occurred during work spanning from month 12 to month 24, and Work Package 4 has been facilitating the refinement of the candidate building blocks focusing mainly on the Blueprint asset. The originally planned schedule of WP4 tasks supports finalization of Blueprint 2.0 (March 2025) and the building blocks. The expectation is that the building block overview setting will not have dramatic high-level changes anymore.

It was also planned that identification tasks (T4.1 requirements, T4.3 standards) would be close to finalization at the date of D4.1 submission (March 2025). As the project progressed, needs beyond the original DSSC plan have emerged and extension for standardization task has been planned until the end of the project. The rationale for extension is that due to timelines of data spaces deployment initiatives and standardization activities flexibility is required for optimal information gathering. Any extension reporting can happen by using forthcoming deliverables which will be provided according to the DSSC plan.

1.2 DSSC internal collaboration process

At an early stage of the project planning, it was understood that the different work packages and tasks need to collaborate. Figure 2 presents how the work between different work packages can be conceptualized as a project internal collaboration process. Main stages of the process are community building, distillation, synthesis, dissemination and guidance and support. For all the stages the figure 2 shows what DSSC does and what it offers. A “direct result” means which work is reported in this document. A contribution, however, is in the borderline of work packages and assets produced. Candidate building blocks were produced for the use in Blueprint, and they have been elaborated with the contribution from WP4.

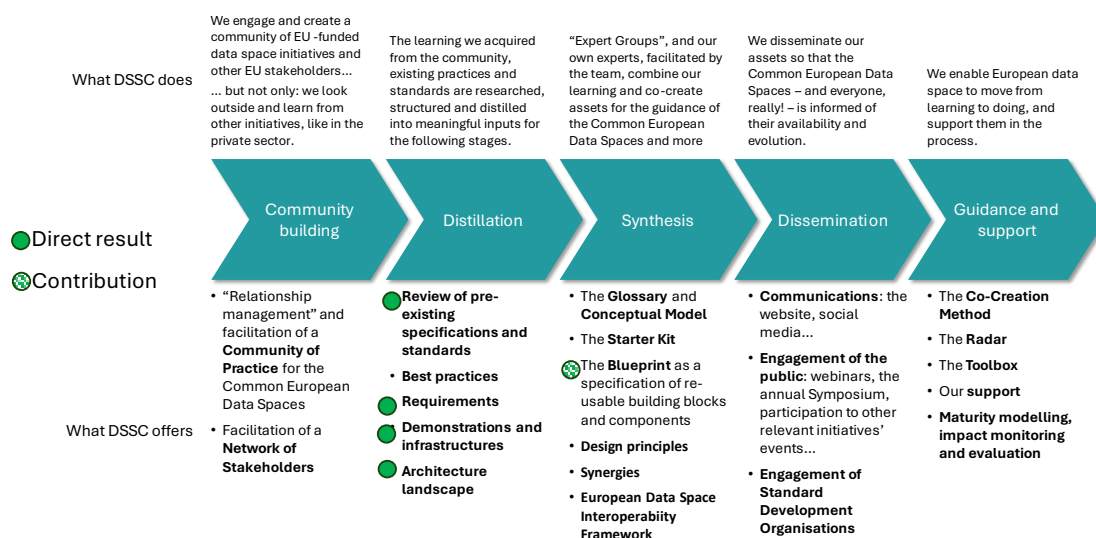


Figure 2: Conceptualization of DSSC internal collaboration process and results and contribution reported in this deliverable mapped to the process

From work package perspective, WP2 is responsible of community building, and WP4 is the one identifying collecting, distilling, and synthesising information to a certain extent to support synthesis, dissemination, and guidance, occurring in other work packages. WP2 and WP4 collaborated in the knowledge creation together with stakeholders.

To give an example, in the first year 2023, WP2 and WP4 organized Monday discussions to coordinate between the work packages and on what different tasks we were aiming at, how to engage various stakeholders. WP2 representatives were participating the meetings to share and discuss the plans on events to take place with stakeholders so that these plans could be aligned with the needs and plans of the experts in WP4. Identification of data spaces requirements, standards, toolboxes, architectures for example, have first and foremost been identified in conjunction with the progress of community of practice.

All in all, a special characteristic of the work package 4, as an identification effort, has been that it produces information and many materials which support other work steps. Thus, the results are intertwined with content development for the Blueprint and Toolbox but especially in the case of standards landscape, distributed to the wide audience.

1.3 Scope of this document

This document scope is in the work done on identifying and defining candidate building blocks, their content, related requirements, standards, and architectures. In results list in chapter 1.4 it reports what materials have been produced and where they can be found.

Chapter 1.1 first presents the DSSC internal collaboration identification effort, and how identification lays between other efforts, especially addressing communication with community of practice (WP2), i.e. European common data spaces, and development of Blueprint and Building Blocks (WP5). These efforts are planned to collaborate seamlessly.

In the chapters on Activities and content, we delve in more detail into what has been done in the selected areas, for example requirements, where we explain the process, and what are the results. Subsequently, how we identified candidate building blocks overview, work which was started in February 2023, shortly explaining the process, and use of the overview in the Blueprint in September 2023.

From the beginning of the project one of the core efforts has been to identify what kind of standards exist or are under development for the use in data spaces. The collection of standards, called standards landscape, is presented, and it is also available as a separate asset at the DSSC website. Such a long list of standards serves as the basis for further study and identification of any gaps the standardization should cover for successful data space interoperability. Architecture landscape gives an outlook to current selection within data space initiatives. Finally, demonstrations and test infrastructures provide definitions of the concepts.

In the end a short summary discusses about the content and planned next steps.

1.4 List of results

The following table lists WP4 tasks, document i.e., resulting material, and its reference. Some of the reports are already published and some are internal and thus not yet released but included in this document.

Task	Document	External reference	Document internal reference
T4.1	Requirements		Chapter 2.1
T4.1	Identified Building Blocks Overview	Blueprint 0.5	Chapters 2.1 and 2.2
T4.5	Identification of and contribution to organisational, business, governance building blocks	Blueprint v0.5 , v1.5 , v2.0	Chapter 2.2.3
T4.6	Identification of legal building blocks	Legal building blocks in blueprint	Chapter 492.2.3.1
T4.6	DSSC role in addressing legal challenges	Blogpost	-
T4.3	Data Space Standards and Technologies Landscape	Report	Chapter 2.3
T4.3	Dashboard for all potential SDOs		Summary included in chapter 2.3.4
T4.3	List of analysed SDOs		Appendix 2
T4.3	Standardization plan	A separate document delivered to EC by request, not in the original project plan	-
T4.3	Liaison person roles	Appendix 1: Liaison person roles	Appendix 1
T4.3	Architecture landscape		Chapter 2.4
T4.1	Requirements for Demonstrators and Test infrastructures		Chapter 2.5

Table 1: Results for D4.1 by dedicated task

2 Activities and content

2.1 Requirements

2.1.1 Introduction

As part of work package 4, task 4.1 (T4.1) identified requirements and building blocks in several dimensions such as business, functional, operational, governance, technical, security, legal and ethics, including trustworthiness and AI-related aspects, when applicable. These findings contributed to the creation, definition, and continuing development of several versions of the building blocks. More specifically, this part of task 4.1 contributed to the following parts of the DSSC overall project:

- ❑ **Deliverable:** D4.1 Candidate of Building Blocks (M30)
- ❑ **Milestones:**
 - MS12-Candidate Building Blocks 1
 - MS14-Candidate Building Blocks 2
- ❑ **Building Blocks version v1.0, 1.5 and v2.0**

The result comprises a collection of requirements, both technical and non-technical, from the CoP initiatives. Various methods (e.g. survey with key questions, various cycles of interviews and revision of key outputs and deliverables) have been utilized to gather inputs in terms of the initial BLOFT and OpenDEI framework to initially structure the information about these main pillars. Additionally, the task collected main challenges/gaps (see gap analysis in chapter 2.1.2.5) encountered by these initiatives and other relevant horizontal aspects such as means to ensure trustworthy AI and ethical concerns as well as cyber security measures for designing the data spaces in their respective domains or cross-domains.

The collection of requirements and most used building blocks started in March 2023 with continuous iterations with the CoP initiatives to update the information until September 2024. The information was made available internally in a dedicated space in Confluence linked to the DSSC support platform.

As the actual output of task 4.1 has been duly considered and absorbed in the overall structure and scope of the building blocks and their individual descriptions, this section 2.1 restricts itself to a summary of important activities, the methodology, and noteworthy particular outcomes which may be relevant on a stand-alone basis as well.

2.1.2 Methodology

2.1.2.1 Introduction

The methodical strategy crafted to attain both short-term and medium-to-long-term goals encompasses various mechanisms outlined in the below sections. From the onset of the initiatives, the crucial steps of identifying and gathering requirements, as well as establishing foundational building blocks, have been closely intertwined with collaborative efforts among consortium partners and the Community of Practice (CoP).

The methodical strategy is explained in the following four stages:

- Stage 1: Building Block Taxonomy Paper
- Stage 2: Collection of requirements from CoP data space initiatives
- Stage 3: Review of blueprints and other relevant material delivered by the CoP initiatives
- Stage 4: Gap Analysis

In order to provide a high-quality gap analysis as the final outcome, the four-stage process was enriched with a feedback cycle from later stages to earlier stages (e.g., to obtain additional pieces of information, for clarification purposes), where appropriate.

2.1.2.2 Stage 1: Building Block taxonomy paper

The building block taxonomy paper provides a classification scheme to be used to describe, analyse and organize data spaces according to basic units or components, i.e. building blocks. The development of the Building Block (BB) taxonomy³ paper, comprising a collection of potential building blocks, commenced in March 2023. This paper served as the primary foundation and framework, delineating the necessary building blocks for advancing data space development and fostering interoperability.

Initially, leveraging the existing OpenDEI framework, particularly for technical building blocks, the paper underwent circulation within DSSC and among experts starting from May 2023, primarily to provide a structural basis for subsequent further developments.

The initial draft was shared with the Community of Practice (CoP) in early July 2023, followed by soliciting feedback from both the DSSC Architectural Board on multiple occasions and the CoP in August. Throughout the two review rounds, the Security Advisory Board (SAB) raised no objections to the material's release.

The BB taxonomy paper offered a comprehensive overview of all the proposed candidate building blocks within the DSSC, explaining their nature and function. It was determined that the further analysis was required for the Governance building blocks. The contents of the paper were merged into Blueprint and released via web site.

The next figure shows the building block taxonomy produced by task 4.1, renamed in the release as building block overview.

³ 'Building block taxonomy' was renamed as 'Building block overview' after Blueprint 0.5

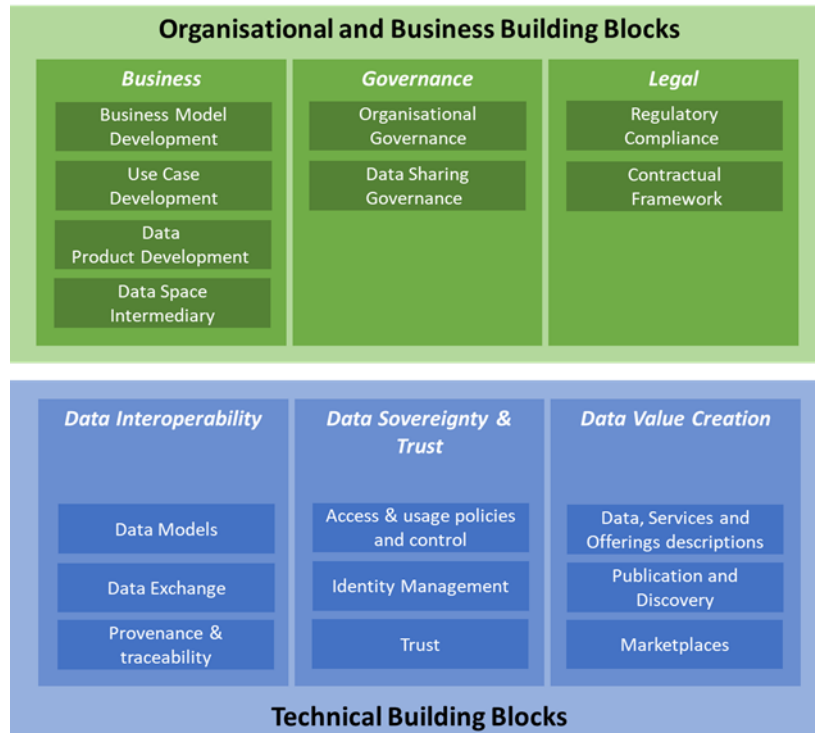


Figure 3: First version of the DSSC building block taxonomy, i.e. later on 'building blocks overview'.

2.1.2.3 Stage 2: Collection of requirements from CoP

The stage 2 task involved in an elaborated questionnaire/guide interview based on the initial DSSC and related developments that are as follows:

- DSSC Starter kit
- [DSBA Technical Convergence](#)
- Building Block taxonomy paper

The purpose of this questionnaire was to collect relevant information from the CoP initiatives for identifying commonly used building blocks and related requirements, both technical and non-technical.

The questionnaire contained background information, purpose, objectives and a section to be filled out the initial DSSC building block taxonomy including core questions to guide the participant. Additionally, there were questions related to other relevant aspects such as ethical concerns, trustworthy AI and a section for further suggestions.

The objective of the questionnaire was to identify the following criteria:

- Tech and non-tech requirements.
- Most used building blocks within sectoral data spaces and across sectoral data spaces.
- Gaps and needs.
- If applicable, identification of ethical considerations, trustworthiness of AI related solutions and applications including cyber-security aspects.

The interaction with the CoP initiatives occurred in close collaboration with WP2, specifically through the Relationship Managers, who facilitated seamless communication. The interviews and analysis of information provided by the data spaces were conducted by T4.1 partners. This process involved multiple methods to gather extensive inputs and detailed information. It included iterative cycles of interviews, analysis, and revision of questionnaires, as well as analysis of key outputs and deliverables from the data space initiatives. Several rounds of updates were made according to the progress of the data spaces to ensure accuracy and relevance.

The Data space initiatives which actively contributed to the surveys/questionnaire are listed below:

- DS4Skills: Data space for skills
- Green Deal Data Space: GREAT project
- AgriDataSpace: European data space for agriculture
- DS4SSCC: Data space for Smart and Sustainable Cities and Communities
- DATES: Data space for tourism
- int:net: int:net is a CSA which supports 5 energy related projects: Omega-X, ENERSHARE, EDDIE, Data Cellar, Synergies
- Europeana: Common European data space for cultural heritage
- European Language Data Space (LDS)

Note: Data 4.0 Space participated in an early interview, however insufficient details were provided to complete the questionnaire. Tourism data space and PrepDSpace4Mobility did not participate in the survey/questionnaire but requested the DSSC to review their blueprints or relevant deliverables before finalization to assess alignment with the DSSC project.

2.1.2.4 Stage 3: Review of blueprints and other relevant material delivered by the CoP initiatives

Some of the data space initiatives from the DSSC CoP requested the support of the DSSC to review their (draft) blueprint and other relevant material to analyse the alignment with the DSSC initial blueprint and building block taxonomy. This exercise was also an opportunity for DSSC to gather better insights of the CSAs blueprints and roadmaps for upcoming deployments as well as identifying potential gaps in the DSSC initial taxonomy.

The approach taken to effectively conduct these reviews involved the following steps:

- Creation of specific templates aligned with the table of contents of the report provided by the corresponding CSA.
- Development of a generic template to compare the DSSC building blocks taxonomy with the proposed taxonomy of each CSA. This facilitated the identification of synergies, missing specifications, and/or potential building blocks.

As such, the following projects were analysed by T4.1 partners:

- **PrepDSpace4Mobility:** D3.1 – Roadmap for future European Mobility Data Space.
- **DATES:** Blueprint for a European Tourism Data Space.
- **DS4SSCC:** D3.2 – Architecture Model as part of Technical Blueprint.

The reviews and the dedicated sessions with the data space projects took place between August and September 2023 to exchange views and provide feedback.

2.1.2.5 Stage 4: Gap Analysis

The final activity within the T4.1 during the 18 months period was performing an in-depth gap analysis based on the updated list of requirements - both technical and non-technical - collected from the CoP initiatives. The analysis also contains a first version of the comments to the DSSC Blueprint 1.5 from the DG-DIGIT B2 Interoperability Architecture Solutions team.

The purpose of this exercise was twofold:

1. To identify potential extensions or missing 'components' in the DSSC Building Block taxonomy and/or potential new Building Blocks, e.g. personal data space intermediaries.
2. To provide inputs for the work related to the gap analysis corresponding to the T5.5 Gap Analysis of the *blueprint building blocks* which explicitly mentioned the requirements obtained in T4.1 from the CoP projects being part of the DSSC community among other sources of information.

Moreover, a new JIRA process, **Asset lifecycle Management**, is launched to facilitate the communication of the key findings to the Expert Groups (EG) and Architecture Board (AB) for final evaluation and necessary updates of the DSSC building blocks specifications Asset Lifecycle management

2.1.3 Outcomes

List of data space initiatives and their current phase:

Data Space	Phase
DS4Skills	Exploratory/Preparatory
Green Deal Data Space - GREAT	Exploratory/Preparatory
AgriDataSpace	Exploratory/Preparatory
DS4SSCC	Exploratory/Preparatory and Implementation
DATES	Exploratory/Preparatory
Int:net	Exploratory/Preparatory
Europeana	Exploratory/Preparatory
Language Data Space (LDS)	Exploratory/Preparatory

Table 2: Data space initiatives

All the collected requirements are captured and structured in the following pattern for each data space initiative:

- Business, Legal and Governance Requirements

- Operational, Functional and Technical Requirements
- Ethical/Cyber security measures

2.1.3.1 Business, Legal and Governance Requirements

Business, legal and governance requirements are presented for each data space initiative based on the following criteria:

- Main Requirement
- Applicable Business models
- Participants/roles in the data space
- Regulations and legal challenges

Data4Skills

Main Requirement

The value of data sharing from this Data space is by enabling, mapping, matching, and forecasting skills and competencies based on a human-centric approach.

Applicable business models

The joint innovation approach was identified, which means that customer innovation can only be realized by ecosystem members working together as no single ecosystem member has all the necessary data.

Business models:

The business models that are relevant for the development of a data space for skills are identified in the following ranges:

- B2B and B2C (the most common cases)
- C2B (business aimed at the creation of services to provide access for companies to individual profiling tools regarding skills and expertise mapping)
- B2G (business conducted between a company and a government in the field of open data service providers)
- G2C (access provided by governments to their information and services towards citizens)
- G2B (government provides services or information to a business entity via governmental portal or other IT solutions)

Less commonly used models are:

- C2G (the citizen provides the government with information or services)
- C2C (service allowing single individuals data sharing for skills comparison, learning and certification)
- B2B2C (a company accesses the customer via another business, i.e. two businesses come together to deliver their products or services to end-customers)

Participants

No available participants.

Regulations and Legal challenges - the most relevant EU regulations and legal challenges are listed in the below table:

EU Legislation (Directives and Regulations)	Gaps/Legal Challenges
<ul style="list-style-type: none"> • European Digital Strategy • European Strategy for Data • European 2030 Digital Decade • Next Generation EU Financial Policy Framework • European Declaration on Digital Rights and Principles • Regulation on a Single Market for Digital Services (Digital Services Act - DSA) • Regulation on contestable and fair markets in the digital sector (Digital Markets Act - DMA) • Data Governance Act (DGA) • Interoperable Europe Act • European Interoperability Framework (EIF) • General Data Protection Regulation (GDPR) • Regulation on Privacy and Electronic Communications (ePrivacy) • Ethics guidelines for trustworthy AI • Directive on Open Data and the reuse of Public Sector Information • Directive on Copyright in the Digital Single Market • European Skills Agenda • Pact for Skills • Digital Education Action Plan • European Framework for Digital Education Content • European Digital Competence Framework (DigComp) • European Qualifications Framework (EQF) • European Digital Skills Certificate (EDSC) • European Digital Credential Infrastructure Standards (EDCI) • European multilingual classification of Skills, Competences and Occupations (ESCO) 	<ul style="list-style-type: none"> • There is a lack of clear reference models. • Lack of a comprehensive framework which defines governance rules and agreements for data spaces management. • The main challenges encountered to continue the development and scaling are related to: <ul style="list-style-type: none"> • Financing • Governance • Implementation of legal frameworks • Participation and engagement by the different ecosystem actors • Lack of or reduced political support • Lack of common vocabularies • Lack of a unique selling point • Complexity of agreements and contracts among the different ecosystem actors

Table 3: Regulations and Legal challenges for Data4Skills

Green Deal Data Space - GREAT

Main Requirement

The value of data sharing from this Data space is about aiming on the social and environmental aspects. The data value proposition is mainly focusing on the interests of public and private sector-related entities to share data for a greater common societal goal, for example, climate protection.

Additionally, other relevant indicators are as follows:

Social Value: Empowering the consumer to better use the facilities they have at home (when to use washing machines, when to wash the car, etc).

Environmental value: Reducing material usage which in turn reduces the pollution impact.

Applicable business models

There are collaborative business models for cross-border use cases. For example, to fulfil UN convention requirements, data is needed from many data sources which might include third party transfers (e.g. example of **GOS4M**)

Requirement:

- Track the provenance and sustainability of data/data products, for example, how often products change, or the risk of a product/dataset not being maintained any more, and make the information public, so that the potential funders are aware about it.
- Multitude of data transactions based on bilateral agreements. There is not a standardized format, particularly when third countries (non-EU) are involved.

Note: Specific building blocks are being developed by several funded projects.

Participants

Below is the list of participants:

- Data providers
- Users
- Data aggregators
- Infrastructure providers*
- Service providers

*Infrastructure/service providers, for example, EGI will offer authentication services with many identity providers, for example, **edugain** with over 5000 identity providers over the world. Plenty of roles to consider around governance, how to run the infrastructure and operations (data architects, support, operators).

Regulations and Legal challenges - the most relevant EU regulations and legal challenges are listed in the below table:

EU Legislation (Directives and Regulations)	Gaps/Legal Challenges
<ul style="list-style-type: none"> • General Data Protection Regulation (GDPR) for sensitive information (e.g. earth observation) data at the farm level identifies the farmer's activities) • Data Governance Act (DGA) a GDDS likely to be a data intermediary, also involving considerable public sector data • Data Markets Act (DMA) • Data Services Act (DSA) • AI Act • Data Act (DA) 	<ul style="list-style-type: none"> • Understanding the roles and responsibilities (Need for a standardized approach) • Keeping up with new regulations (e.g. DGA, DA, DSA, DMA and AI Act) • Interplay between legal initiatives • Complying with personal data protection • Establishing trust between partners (Need for a standard trust framework) • Antitrust law concerns (strong need to pool data from multiple players which may trigger competition issues.)

<ul style="list-style-type: none"> • Open Data Directive (ODD) • Framework for a European Digital Identity (eID), a reliable identity management is needed. • Inspire Directive 	<ul style="list-style-type: none"> • Uncertainty regarding liability • Commercial sensitivity of data (Environmental data from entities might be extremely sensitive - Proceeding with data sharing, IoT/sensor data covered by DA).
--	--

Table 4: Regulations and Legal challenges for Green Deal Data Space - GREAT

Note: Generally, participants and 3rd parties access legal agreements data based on bilateral agreements.

Agri Data Space

Main Requirement

There are an extensive number of government-managed datasets and B2G relationships which are mostly based on manual inputs. Thus, there are a lot of opportunities in connecting all these data silos through data spaces, which automate interchanges, facilitate interoperability, and lower trust barriers. Additionally, following are the most relevant indicators:

Economical value: The economic value can be identified through a new data market, cutting expenses and increasing value of end products or services.

Social Value: The future European data space for agriculture is part of the sector digitalization to tackle the social and environmental challenges. The CSA is part of the Farm to Fork strategy for a fair, healthy and environment-friendly food system, at the heart of the European Green Deal.

Environmental value: The direct environmental impact of data space, regarding data storage, hardware components, etc. is very difficult to access. Indeed, the data space will be decentralized, and data will be stored and managed by each user. This architecture does not require energy consuming system. The indirect impact will come from the use that is made of these data. The consent can be a solution to engage the actors using the data to be in-line with the Farm to Fork strategy.

Common Use Cases: Risk management to face extreme weather conditions, digital twins (future plans), Production processes - complete life cycle assessment and monitoring (e.g. towards the implementation of the Common Agricultural Policy 2023-2027 which promotes a more sustainable agriculture supporting the Green Deal targets)

Applicable Business models

Besides EU funds, private funding will be analysed for supporting the implementation and maintenance of the data space. Including private investors in the development of technical building blocks (cloud storage, identity management solutions, etc.) could be a possibility.

Applicable fees can depend on the type of organization and the activity performed on the data space.

Participants

It is identified that six target groups will be interested in using the future European data space for agriculture, the below list include both private and public organizations:

- Farmers, advisors, and their associations
- Technology Providers and Data intermediaries

- Agribusiness
- Society
- Public bodies
- Academia and scientific communities

Note: The participants and their roles will be defined in the **multi-governance schema**. We expect to onboard all type of organizations of the agricultural sector, **public and private**. We can use technology and service providers to compose the future data space.

Regulations for data space and challenges – the most relevant EU regulations and legal challenges are listed in the below table:

EU Legislation (Directives and Regulations)	Gaps/Legal Challenges
<ul style="list-style-type: none"> • General Data Protection Regulation (GDPR) • E-Commerce Directive • Regulation on the free flow of non-personal data • Data Governance Act (DGA) • Data Act (DA) • Data Markets Act (DMA) • Data Services Act (DSA) • ITS Directive • European Digital Identity (EUDI) Regulation 	<ul style="list-style-type: none"> • Code of Conduct published by the Copa Cogeca dedicated to the agricultural sector and potential overlapping with EU Legislation, e.g. DGA, DA • Data exchanged usually based on bi/multi-lateral agreements between parties that are mainly business-driven and less focused on data sovereignty, consent etc. This raises concerns towards adoption and use of data-governance building blocks.

Table 5: Regulations and Legal challenges for Agri Data Space

DS4SSCC

Main Requirement

The value of data sharing of this Data Space is related to Smart city benefits, not limited to, but not either concretized by a specific use case. Following are the main aspects:

- Improvement of decision-making for both citizens and managers by providing them with accurate, up-to-date information; better alignment of city objectives, more informed decision-making, and effective problem-solving, ultimately improving the lives of citizens.
- Improved efficiency
- Sustainability
- Enabler for cities and communities to foster digitalization of their infrastructures and services.
- Alignment with EU Green Deal Objectives

In addition, in terms of relevant indicators (general level, without a specified use case):

Economical value: New business through new data-driven services; Improved cost-efficiency through optimization and improved decision-making and resource allocation (e.g. public transport dimensioning based on hotel reservation data and passenger count at the airport)

Social Value: Improved quality of life for the citizens (e.g. safety, cleanliness of the city), citizen participation

Environmental value: Less emissions, e.g. through improved traffic management solutions, improved energy efficiency (alignment with EU Green Deal Objectives).

Organizational/Governance models to support the sustainability of the data space and business models

A multi-governance model is developed in collaboration with main stakeholders, i.e. public administrations, private sector, academia and civil society, which includes a Code of Conduct to provide guidelines towards a federated European data space for smart and sustainable cities and communities.

Participants

Below is the list of participants:

- Cities and regions (data producer and user), National Governments.
- Suppliers of city solutions (service provider, data producer and user),
- Citizens/civil society organizations (data producer and user),
- Private sector/ both large corporations and SMEs/start-ups (data producers and user), including private public partnerships.
- Academic and research institutions (data producer and user)

Regulations and Legal challenge - Most relevant EU regulations for this Data Space are listed below:

EU Legislation (Directives and Regulations)	Gaps/Legal Challenges
<ul style="list-style-type: none"> • General Data Protection Regulation (GDPR) • Regulation on the free flow of non-personal data • Data Governance Act (DGA) • Digital Markets Act (DMA) • Digital Services Act (DSA) • Data Act • Interoperable Europe Act • AI Act • Copyrights Directive • ePrivacy Directive • Electronic identification and trust services for electronic transactions in the internal market (eIDAS) 910/2014 • Open Data Directive (ODD) • Platform-to-Business Regulation 	<ul style="list-style-type: none"> • Clear access control policies over data and services (supporting cities to progress from sharing of open data to data with limited access) • Multi-stakeholder governance framework fully defined.

Table 6: Regulations and Legal challenges for DS4SSCC

DATES

Main Requirement

The data space can add value to tourism by data sharing in the following different aspects:

Economical value: increase in the competitiveness of the companies (new business models, costs reduction, etc) personalization of services, implementation of sustainability principles as added value for customers, possibility to implement a model where tourists as individuals can get additional value out of their data.

Social Value: increase the quality of life, increase the quality of work conditions contribution to smart planning (of urban services) or adaptation based on seasons and level of demand, sustainability, customer-centric services, inclusiveness, compliance with EU values (privacy, transparent and explainable AI)

Environmental value: reduce the environmental impact (resources consumption, better measurement of carbon footprint, etc).

Applicable business models

This is work in progress. There are both collaborative and individual business models. An example of individual business model could be a better understanding for a service provider (e.g hotel) of the future demand and usage of data to get a more balanced arrival of tourist (e.g instead of getting most of the revenues in summer, be able to offer services that help the hotel to reach a better level of visitors along the year).

An example of collaborative business model could be the use of data from several players of the tourism sector to anticipate the demand for services in a specific tourist destination. Better view of the usage of public transport, occupancy of hotels and other lodging, potential consumption of water and other utility services would allow the destination to make a better planning and management of resources. Data providers could benefit from the suitable planning of those services (e.g. a bus line that may be created ad-hoc to service a big number of visitors).

Participants

The list of participants and roles is based on the roles identified by DSSC, but some experts have pointed out missing roles, like those associated with the deployment and exploitation of the data space infrastructure elements or those related to guarantee the individual's control over their data via the so-called "Personal Data Intermediaries" (PDIs); others may be needed.

Participants/Roles in Data Space	Description
Partner of a data space initiative	A data ecosystem party that formally commits to contribute to a particular data space initiative.
Data space member	A data ecosystem party that has committed to the governance framework of data space and may have one or more roles in it.
Transaction participant	A data space member that directly participates in a data transaction in a data space by providing data, providing permissions related to the data or by receiving data and/or permissions to use the data.
Data space enabler	A data space member that provides a (technical or non-technical) service enabling data transactions for the transaction participants while not directly participating in that

	transaction itself. Examples of enabling services include identity provisioning, vocabulary provisioning, interconnecting, clearing, etc.
Data rights holder	A transaction participant that has the legal right to use, grant access to or share certain data.
Data provider	A transaction participant that, in the context of a specific data transaction, technically provides data to the data receivers that have a right or duty (granted by the data rights holder) to access and/or receive that data.
Data receiver	A transaction participant to whom data is or is to be technically supplied by a data provider in the context of a specific data transaction.
Data user	A transaction participant that is granted (lawful) access and the right to use data as the result of a specific data transaction. Also known as data rights receiver.
Data intermediary	A data space enabler that (technically and legally) connects one or more data space members to the data space, thereby enabling them to establish relationships and execute data transactions with other members in the data space.

Table 7: Participants and roles for DATES

Regulations for data space and challenges - the most relevant EU regulations and legal challenges are listed in the table below:

EU Legislation (Directives and Regulations)	Gaps/Legal Challenges
<ul style="list-style-type: none"> • General Data Protection Regulation (GDPR) • E-Commerce Directive • Regulation on the free flow of non-personal data • Data Governance Act (DGA) • Data Markets Act (DMA) • Data Services Act (DSA) • ITS Directive • Framework for a European Digital Identity (eID) • Other(s): Those associated with the tourism sector as such even though many are not defined at EU level 	<ul style="list-style-type: none"> • Patchwork of legal entitlements to data. • Keeping up with new regulations (e.g., DGA, DA, DSA, DMA and AI Act). • Interplay between legal initiatives. • Complying with personal data protection. • Establishing trust between partners. • Commercial sensitivity of data. • Uncertainty regarding liability.

Table 8: Regulations and Legal challenges for DATES

The value of data sharing of this Data Space is to provide reliable operations from the system operator point of view.

Economic value: benefit for consumer and operator.

Social Value: (related to environmental value) Empower the consumer to better use the facilities they have at home (when to use washing machines, when to wash the car, etc).

Environmental value: to reduce material usage which in turn reduces the pollution impact.

Common Use Cases:

- Solar renewable data sharing between PV plant operators and DSOs
- Multi-vector optimization of Local Energy Communities
- Cross-border Electromobility charging
- Advanced flexibility for Municipalities

Applicable business models

Cost sharing, Joint innovation, combined forces and shared marketplace.

Participants

Participants/Roles in Data Space	Description
Prosumer	Consumer and Producer
Energy Provider	optimize energy production and dispatch, improve ancillary services.
Others	flexibility providers, platform (service) provider, energy communities (aggregating energy flexibility).

Table 9: Participants and roles for int:net

Regulations and Legal challenges - Most relevant EU regulations for this Data Space are listed below:

EU Legislation (Directives and Regulations)	Gaps/Legal Challenges
<ul style="list-style-type: none"> • General Data Protection Regulation (GDPR) • Data Governance Act (DGA) • Data Markets Act (DMA): interconnecting markets in different countries • Data Services Act (DSA): applicable to specific operations? <p>Other legal/governance instruments:</p> <ul style="list-style-type: none"> • Directive (EU) 2019/944 on common rules for the internal market for electricity • Legal agreements for participants and 3rd parties to be fully compliant with the GDPR, for flexibility service providers which rely on data of the final users which cannot be disclosed. 	<ul style="list-style-type: none"> • Understanding the roles and responsibilities • Complying with personal data protection • Establishing trust between partners • Commercial sensitivity of data

<ul style="list-style-type: none"> • Terms and conditions to regulate data sharing between participants will be developed through contracts between the different actors. • Governance models to specify the different roles and responsibilities will be defined for specific implementations. System operator will most likely be responsible for the governance framework. 	
---	--

Table 10: Regulations and Legal challenges for int:net

Europeana: Common European data space for cultural heritage

Main Requirement

The value of data sharing of this Data Space is described below in terms of economic, social and environmental indicators.

Economic Value: The new application for education built with Bits educational challenge invites students and educators to design spaces with Mozilla Hubs integrating Europeana website content. One example for tourism are the digital tours created in the Jewish History Tours project published on the Europeana website. The tours are also available on tourism platforms. The (Scientific) innovation (e.g. new tools & service development; updates to GitHub repositories; new products or services developed) Network development (e.g. new contacts and collaboration made Training and capacity building (e.g. increased digital maturity - organizationally/individually)

Social Value: Some of the advantages that add social value:

- Engaging Europe's citizens in participatory and crowd-sourcing campaigns (e.g. GIF IT UP, Digital Storytelling Festival, Transcribathon.eu)
- Improving the quality and richness of education with digital cultural heritage (e.g. with eLearning activities and source collections)
- Inclusive access to digital cultural heritage (e.g. multilingual offer to products and services; guest-written editorials)
- Network engagement (e.g. professional and personal value of events, training and conferences; engagement of active network participation)
- Training and capacity building (e.g. increased skills/knowledge)

Environmental Value: [Climate Action community](#) (activities and resources related to climate awareness).

Applicable business models

Please see [Europeana Strategy 2020-2025](#) and [Europeana Foundation Business plan 2023](#).

Participants

Participants and roles in the Data space are mentioned in the table below:

Participants/Roles in Data Space	Description
Data space Provider	Europeana DS consortium , led by the Europeana Foundation in collaboration with a consortium of 18 partners, works to deploy the common European data space for cultural heritage.
Data Providers	National and domain aggregators (organized in the Europeana Aggregators' Forum) work with 3700 + cultural heritage institutions to gather authentic and accurate data and make it accessible through the data space.
User Groups	User groups are the beneficiaries of the data available in the data space (main user groups are European citizens, educators and students at the primary and secondary levels, academia and research, those engaged in vocational education, as well as lifelong education, and API users/developer communities; user groups under investigation are related to tourism and media). Each user group has targeted outreach and engagement activities. See <i>Data Value Creation</i> tables for more details.

Table 11: Participants and roles for Europeana

Regulations and Legal challenges - The most relevant EU regulations and legal challenges are listed in the table below:

EU Legislation (Directives and Regulations) ⁴	Gaps/Legal Challenges
<ul style="list-style-type: none"> • General Data Protection Regulation (GDPR) • E-Commerce Directive • Regulation on the free flow of non-personal data • Data Governance Act (DGA) • Data Markets Act (DMA) • Data Services Act (DSA) • European Digital Identity (EUDI) Regulation 	<ul style="list-style-type: none"> • Patchwork of legal entitlements to data. • Keeping up with new regulations (e.g., DGA, DA, DSA, DMA and AI Act). • Understanding the roles and responsibilities. • Interplay between legal initiatives. • Complying with personal data protection. • Establishing trust between partners. • Antitrust law concerns. • Commercial sensitivity of data. • Uncertainty regarding liability.

Table 12: Regulations and Legal challenges for Europeana

Language Data Space (LDS)

Main Requirement

The value of data sharing of this Data Space is to develop and deploy a European platform and marketplace for the collection, creation, sharing and re-use of multilingual and multimodal language data.

⁴ Note these pieces of legislation have been indicated by the data space initiative itself. We are aware that other legal instruments, e.g., Info Soc Directive or CDSM Directive, would also play a role for this data space initiative. For the sake of authenticity of responses, we did abstain from adding any further items, though.

Applicable business models

The identified business models are cost sharing, joint innovation, combined forces, shared marketplace and greater community good. For the latter, it applies:

1. Collection, sharing and re-use of as many high-quality language datasets as possible, multilingual and multimodal in nature, in line with the previous initiatives (e.g., ELRC) and in respect of EU legislation and values.
2. Furthering of Digital Language Equality

Participants

Any language data holder and/or consumer. They can be, indicatively:

- EC, Member States
- EU industry (from various sectors being language data transversal: e.g., advertising, broadcasting, defense, finance, general ICT, health, legal, LT, NLP, AI, mobility, newspapers & news agencies, pharmaceutical, publishing)
- Research centers
- Public administration
- Libraries and archives
- Cultural associations
- NGOs
- Citizens

Regulations and Legal challenges - Most relevant EU regulations for this Data Space are explained in the table below:

EU Legislation (Directives and Regulations)	Legal Challenges
<ul style="list-style-type: none"> • General Data Protection Regulation (GDPR) • Data Protection Regulation for the European Union institutions, bodies, offices and agencies 1 ('EUGDPR') • Regulation on the free flow of non-personal data • Data Governance Act (DGA) • Digital Market Act (DMA) • Data Act (DA) 	<p>Governance</p> <ul style="list-style-type: none"> • Establish a multi-stakeholder data and services framework • Establish Centre of Excellence of Language Technologies (CELT) and CELT+ to define and implement the governance framework. <p>Legal</p> <ul style="list-style-type: none"> • Ensure GDPR and EUGDPR compliance of personal data processing operations • Ensure that LDS is in line with the principle of compliance “by design and by default” set in the GDPR and EUDPR

Table 13: Regulations and Legal challenges for LDS

2.1.3.2 Operational, Functional and Technical Requirements

The operational, functional and technical requirements are presented for each data space initiative based on the following criteria:

- Data Interoperability Requirements/Building Blocks
- Data Sovereignty Trust Requirements/Building Blocks
- Data Value Creation

Data4Skills

Data Interoperability Requirements/Building Blocks

Data Interoperability	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Data models • Data exchange APIs • Data provenance and traceability 	<ul style="list-style-type: none"> • Need for semantic interoperability, enhanced data quality, and the possibility of producing AI models that could operate with incomplete data. All these aspects together could define a Minimum Interoperability Mechanism (MIM) covering both functional and technical issues. • Recommendation: use of JSON-LD as the appropriate way to describe skills data models and pivot skills ontologies that support broad semantic interoperability among the knowledge domains. • Rather than imposing a single ontology, it is recommended to develop semantic translators to and from pivot ontologies as the right approach. Other recommended future standards to improve interoperability are GraphQL API and Open Badges v3 (OBv3). • Need for implementing a universal plugin that would integrate different solutions, the use of the W3C Verifiable Credentials standard, and a mechanism for digital wallet interoperability. 	<p>Need for a new building block: skills data interoperability</p>

Table 14: Data Interoperability Requirements – Data4Skills

Data Sovereignty and Trust Requirements/Building Blocks

Data Sovereignty Trust	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Identity and access management framework • Access and usage control/policies • Trust exchange/Trust Frameworks 	<ul style="list-style-type: none"> • Use of secure data architectures • Manage consent forms and contracts based on standards, e.g. Kantara Consent Receipt and Open Digital Rights Language (ODRL) to make processes and outcomes easily explainable. Standards such as IDP, SSO, and 	<p>Future implementations and building blocks:</p> <ul style="list-style-type: none"> • Recommended the use of decentralized protocols, self-sovereign identity (SSI) management, smart

	Verifiable Credentials were identified as the main standards in use today.	contracts, and decentralized AI training. <ul style="list-style-type: none"> • Need for building blocks that could guarantee the individual's control over their data through Personal Data Intermediaries (PDIs)
--	--	---

Table 15: Data Sovereignty and Trust – Data4Skills

Data Value Creation

Data Value Creation	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Data usage accounting • Metadata and discovery services • Publication and marketplace services 	Need for a common vocabulary to standardize specifications for describing the datasets and services in the data space. As a reference, the Gaia-X catalogues of services and JSON dashboards are considered the main references and standards that are currently used.	Future implementations and building blocks : <ul style="list-style-type: none"> • Recommended the use of non-fungible tokens (NFTs) for gamification, FAIR principles to optimize the reuse of data, and data vault modelling.

Table 16: Data Value Creation – Data4Skills

Green Deal Data Space - GREAT

Data Interoperability Requirements/Building Blocks

Data Interoperability	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Data models • Data exchange APIs • Data provenance and traceability Note: all the BBs will be needed.	Selection of brokers/connectors to be deployed not identified at this stage. Specific standards to ensure interoperability will be identified, e.g., Open Geospatial Consortium (OGC) Note: Technical components to be defined.	Technical building blocks developments, including semantic standards and interoperability protocols.

Table 17: Data Interoperability Requirements Green Deal Data Space - GREAT

Data Sovereignty and Trust Requirements/Building Blocks

Data Sovereignty Trust	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Identity and access management framework • Access and usage control/policies 	Implementation of the 3 BBs on the 1 st column will be performed via different interoperable solutions and AAI	Looking for best practices to develop a common set of rules and agreements/contracts to

<ul style="list-style-type: none"> Trust exchange/Trust Frameworks <p>Note: all the BBs will be needed.</p>	<p>(Authentication and Authorization Infrastructures).</p> <p>Rules and agreements for this DS to facilitate the exchange of data in a trusted way will be described by developing contracts/common framework for the participants.</p> <p>In terms of a process, to verify the identity of participants when data exchange transactions occur. Based on the AARC blueprint with VOMS/PERUN and to be mapped into the VC/VP in terms of functionalities.</p> <p>Technical approach is based on a federated AAI following AARC guidelines.</p>	<p>ensure trusted data exchanges among the participants.</p> <p>In terms of IAM and personal data management roles (e.g. PDI) this is an open question.</p> <p>Runtime policy enforcement (i.e. control over data after its shared) not contemplated as of today but to be regulated by contractual terms.</p>
---	---	--

Table 18: Data Sovereignty and Trust Requirements – Green Deal Data Space - GREAT

Data Value Creation

Data Value Creation	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> Data usage accounting Metadata and discovery services Publication and marketplace services <p>Note: all the systems/services listed will be implemented.</p>	<p>Most of the services mentioned in the 1st column are already implemented in their key data providers. It is under analysis how to integrate their existing systems and capabilities.</p> <p>A metadata broker is essential and ensure to be implemented by their providers.</p>	<p>Separating data usage accounting scheme from marketplace services is found to be difficult technically.</p> <p>Data market services for certain data assets, data publication, search and data discovery services will be provided.</p>

Table 19: Data Value Creation – Green Deal Data Space - GREAT

Agri Data Space

Data Interoperability Requirements/Building Blocks

Data Interoperability	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> Data models Data exchange APIs Data provenance and traceability 	<ul style="list-style-type: none"> Application of ISO 11783 (ISOBUS) standard which defines the communication between agricultural machinery and facilitates data transfer between mobile far 	<p>No gaps identified</p>

	<p>equipment and farm management information systems (FMISes)</p> <ul style="list-style-type: none"> • Data exchange principles should be standardized (i.e. one interface protocol between machinery and telematics services and from telematics to external APIs through a standardized cloud-cloud interfaces). Need for a common set of APIs. • FMIS might be crucial in the future regarding the B2G connection for reporting and control obligations. It is being explored how to perform automatic reporting by connecting FMIS directly to National Farm Registries or Animal Identification and Traceability Systems. • Application of communication protocols and platform-independent standards for data exchange such as MQTT, HTTP/REST, Kafka, CoAP, DDS, OPC-UA, etc. • Data models and APIs to enable semantic interoperability and use of tools such as OpenAPI or AsyncAPI. • Adoption of common semantic models and technological enablers avoiding vendor lock-in. 	
--	---	--

Table 20: Data Interoperability Requirements - AgriDataSpace

Data Sovereignty and Trust Requirements/Building Blocks

Data Sovereignty Trust	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Identity and access management framework • Access and usage control/policies • Trust exchange/Trust Frameworks 	<ul style="list-style-type: none"> • Implementation of a IAM framework and systems enabling single sign-on systems, two-factor authentication, multifactor authentication and privileged access management. • Common Identity Framework as a building block for authenticating data stakeholders and for checking Verifiable Credentials 	<p>Consent platforms to control the access to data, however explicit control mechanisms to enforce this is still missing.</p> <p>Technical mechanisms to allow increased control over data usage, storage and sharing.</p>

	<ul style="list-style-type: none"> • Need to create trusted environments that incorporate data sovereignty by design, e.g. Gaia-X Trust framework • Mechanisms that guarantee data control or blockchain-based trust models, such as the European Blockchain Services Infrastructure (EBSI) in the case of personal, competitive or otherwise sensitive data, mechanisms like compute-to-data, federated learning or multi-party computing could add instruments to the agro-data ecosystem that ensure privacy while retaining data control, thus enabling data sovereignty by design. 	
--	---	--

Table 21: Data Sovereignty and Trust Requirements - AgriDataSpace

Data Value Creation

Data Value Creation	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Data usage accounting • Metadata and discovery services • Publication and marketplace services 	<p>Implementing a layer of smart contracts that ensures that data is not misused. Marketplace to be perceived as a sort of 'Data Clearing House' building block. There are emerging solutions in the agricultural sector, such as the API-Agro platform, that implements both the building blocks for data exchange and the conditions necessary for data sharing by proposing a solution like a Data Marketplace</p> <p>Requirements for a marketplace offer different options for data sharing such as:</p> <ul style="list-style-type: none"> • Per file (with a wide range of file formats) • Per recurring file (the platform sends files when they are added) • Per pull API • Per push API <p>Trust and transparency attributes: identification of registrants, data exchange traceability, etc. Flexible to be</p>	<p>No gaps identified.</p>

	adapted according to different types of governance and business models.	
--	---	--

Table 22: Data Value Creation - AgriDataSpace

DS4SSCC

Data Interoperability Requirements/Building Blocks

Data Interoperability	Requirements/Technical Solutions	Needs/Gaps
Data models	<p>Data Models</p> <p>Minimum requirement: describe the data sources and make these data available in the data space, e.g. to describe the name of the different attributes present at the data sources (e.g. data type, possible values, list of strings, etc) and explaining the meaning of the values. (Metadata). Related MIM:MIM2</p> <p>Related standards: SAREF, OSLO, Data Privacy Vocabulary (DPV), NTEX, Siri, DataxII, OPIN, PSD2. Related Industry specifications: Smart Data Models (FIWARE/IUDX/OASC/TM Forum), Indian Urban Data Exchange, INSPIRE data specifications, Crop Ontoogy, OPC UA, MaaS Data String.</p> <p>Reference implementations: IoT big data harmonized data model (GSMA), CityGML/CityJSON,</p>	Maturity: Quite mature
Data exchange APIs	<p>Data Exchange APIs</p> <p>Functionality: adopt domain-agnostic common APIs and security schemas for data exchange together with data models represented in data formats compatible with those APIs. Related MIM: MIM1.</p> <p>Related standards: NGSII-LD. Related industry specifications: LDES, MQTT, JSON-LD. Reference implementations: Indian Urban Data Exchange, CEF Context Broker.</p>	Maturity: Quite mature
Data provenance and traceability	<p>Data Provenance and traceability</p> <p>Functionality: provide the means for tracing and tracking in the process of data provision and data consumption/use, e.g. timestamp is a mechanism to provide traceability over the</p>	Maturity: not very mature. Provenance & Traceability still needs definition in detail.

	<p>data, as it does not allow values from the past, using timescale data bases. Related MIM: MIM3</p> <p>Related standards and industry specifications: ETSI-CIM, DCAT-AP.</p> <p>Reference implementation: Canis Major (FIWARE)</p>	
--	--	--

Table 23: Data Interoperability Requirements – DS4DSSCC

Data Sovereignty and Trust Requirements/Building Blocks

Data Sovereignty Trust	Requirements/Technical Solutions	Needs/Gaps
Identity and access management framework	<p>Identity and access management</p> <p>Functionality: to provide authentication and authorization of data space participants. MIM4, MIM6</p> <p>Related standards and industry specifications (Decentralized IM): W3C DID, W3C Verifiable Credentials, SOLID, etc.</p> <p>Related standards and industry specifications (Centralized IM): CEF eID, OAUTH2.0, OpenID Connect, SAML 2.0</p> <p>Reference implementation: Keycloak, Okta, Anubis, Microsoft AD, etc.</p>	<p>Maturity: Quite mature</p> <p>Personal data management (MIM4); variety of technical options available but only implemented at small scale.</p>
Access and usage control/policies	<p>Access and usage control/policies</p> <p>Functionality: to provide guarantee to organizations that are part of the data space to exercise sovereignty on their data, using common building blocks for access and usage control based on security standards, a defined framework and reference architectures. Related MIM:MIM3</p> <p>Related standards: XACML policy definition, W3C ODRL, W3C WAC</p> <p>Industry specifications: Rego, Open Policy Agent.</p> <p>Reference implementation: i4Trust (which relies on exchange of context data via NGSI-LD)</p>	
Trust exchange/Trust Frameworks	<p>Trust exchange/Trust frameworks</p> <p>Functionality: Any data space requires a Trust Anchor Framework and associated decentralized Identity and Access Management Framework to enable the trusted operation of</p>	<p>Maturity: quite mature</p>

	<p>the system without requiring a central entity intermediating in all interactions among participants. It is identified as most appropriate implementation, a decentralized IAM based on Verifiable Credentials, as described in the European Digital Identify Wallet Architecture and Reference Framework. Applicable MIM:MIM6.</p> <p>Related standards and industry specifications: EUDI, ESBI</p> <p>Reference implementation: i4Trust , European Blockchain</p>	
--	---	--

Table 24: Data Sovereignty and Trust Requirements – DS4DSSCC

Data Value Creation

Data Value Creation	Requirements/Technical Solutions	Needs/Gaps
Data usage accounting	<p>Data usage accounting</p> <p>It is only considered specifications which have a standard API defined. Applicable MIM3.</p> <p>Related standards and industry specifications: TM Forum Accounting API, IUDX Metering & Audit API</p> <p>Reference implementation: Business API Ecosystem, IUDX Metering API (NGSI-LD Context broker)</p>	Maturity level: few mature
Metadata and discovery services	<p>Metadata and discovery services</p> <p>Relevant MIMs: MIM1, MIM7</p> <p>Related standards and industry specifications Standards that specifically describe metadata not the data itself, e.g. DCAT (W3C), Dublin Core (ISO), INSPIRE, ISO 19115-1:2014.</p> <p>Reference implementation: Typically, metadata management systems are conceived as “data catalogues”, e.g. Colibra, Apache Atlas and Informatica. For open data, common Data Discovery tools include Tableau, Qlik Sense and Atlan.</p>	Maturity level: few mature

<p>Publication and marketplace services</p>	<p>Publication and marketplace</p> <p>Functionality: to enable the publication and offer data resources linked to data assets and services under defined terms and conditions. Moreover, this building block supports the creation and monitoring of smart contracts. Reference MIM3.</p> <p>Related standards and industry specifications: OpenAPI (TM Forum), ICT Innovation Network reference architecture (Slovenia).</p> <p>Reference implementation: Fiware business API ecosystem (BAE)- joint component from Fiware business framework and a set of APIs provided by TM Forum, CKAN Extensions, IDRA, i4Trust, etc</p>	<p>Maturity level: evolving</p>
---	--	---------------------------------

Table 25: Data Value Creation – DS4DSSCC

DATES

Data Interoperability Requirements/Building Blocks

Data Interoperability	Requirements/Technical Solutions	Needs/Gaps
<p>Data models Data exchange APIs Data provenance and traceability (not defined so far)</p>	<p>Brokers to be deployed: IDS Connector EDC Connector Use of GAIA-X catalogue (not to deploy own catalogue) Standards to ensure interoperability IDS Information model Note: in the process of identifying relevant standards for semantic interoperability. Type of data exchange API IDS protocol REST Data exchange formats CSV Other: data transformation or integration tools for data mapping, cleansing or enrichment are not considered at this stage. Similarly, any advanced automated tool or technique for data interoperability (e.g. Named Entity Recognition, Topic Modelling, Text normalization) is identified at this moment.</p>	

Table 26: Data Interoperability Requirements - DATES

Data Sovereignty and Trust Requirements/Building Blocks

Data Sovereignty Trust	Requirements/Technical Solutions	Needs/Gaps
<p>Identity and access management framework</p> <p>Access and usage control/policies Trust exchange/Trust Frameworks</p> <p>Gaia-X Trust Framework (to be implemented)</p>	<p>All the building blocks listed in first column are applicable.</p> <p>Technical implementation of the building blocks is done via:</p> <ul style="list-style-type: none"> • IDS Identity provider2 (CA+DAPS) • IDS data usage policy rules and enforcement. • Definition of data usage policies to facilitate the exchange of data in a trusted way. • Verification of identify when data exchange transactions occur is done using the IDS Identity provider for IDS connector (not using PARIS3) • In terms of identity wallets and/or EUDI wallets or similar approaches, the selected option based on the IDS approach. Moreover, DATES is aware of specific personal data intermediaries for identity management (e.g. wallet providers) such as MyData initiative and the personal data management from Gaia-X. • IDS Data usage policies are used to enable runtime policy enforcement (control over data after its shared) • In terms of IAM services, the project is using IDS identify and authentication mechanisms. 	<p>Gaia-X Trust Framework - Under analysis specific personal data specific technologies such as Solid pods and data wallets.</p>

Table 27: Data Sovereignty and Trust Requirements - DATES

Data Value Creation

Data Value Creation	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Data usage accounting • Metadata and discovery services • Publication and marketplace services 	<ul style="list-style-type: none"> • None of the building blocks mentioned in the first column are implemented. • In the POC, metadata brokers, data market services or data discovery services are not considered. 	<ul style="list-style-type: none"> • GAIA-X defines Logging and audit data abilities needed to provide an auditable framework for transactions. • GAIA-X is defining the federated catalogue concept to cover the

		<p>metadata and discovery services.</p> <ul style="list-style-type: none"> • TM Forum APIs implementing marketplace services.
--	--	--

Table 28: Data Value Creation - DATES

Int:net

Data Interoperability Requirements/Building Blocks

Data Interoperability	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Data models (High priority) • Data exchange APIs • Data provenance and traceability (not relevant at this stage) 	<ul style="list-style-type: none"> • Selection of brokers/connectors to be deployed will rely on specific implementations, e.g. IDS connectors, EDC connectors. Some will be compliant with Gaia-X. • Specific sw developments/components (algorithms) to optimize power system operations. • Use of data aggregation mechanisms and metadata for traceability of data exchange transactions. • Specific standards to ensure interoperability between the DS participants: CIM, IEC 61850, IEC62325, IEC61970 (Energy domain, commonly used for data models) • Type of data exchange API depends on the implementation, e.g. REST is used for one use case or pilot • Type of data exchange formats/data models used: JSON, JSON-LD, CSV • Use of data transformation and integration tools for tasks such as data mapping, cleansing, or enrichment? -->N/A • Use of advanced automated techniques for data interoperability (e.g., Named Entity Recognition, Topic Modelling, Text normalization)-> N/A 	<p>Use of SAREF Saref4ener, saref4grid standards for semantic interoperability to facilitate the exchange of information.</p>

Table 29: Data Interoperability Requirements – int:net

Data Sovereignty and Trust Requirements/Building Blocks

Data Sovereignty Trust	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> Identity and access management framework (High priority) Access and usage control/policies (High priority) Trust exchange/Trust Frameworks (High priority) 	<ul style="list-style-type: none"> Implementation of the 3 BBs on the 1st column was done using Authentication software component and Authorization software component. Rules and agreements for this DS to facilitate the exchange of data in a trusted way are derived from the energy market as this is regulated by national authorities (only authenticated participants). Different IAM mechanisms depending on the role of participants, i.e. enterprises and individuals. 	No gaps identified.

Table 30: Data Sovereignty and Trust Requirements – int:net

Data Value Creation

Data Value Creation	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> Data usage accounting Metadata and discovery services Publication and marketplace services 	<p>Implementation of the BBs on the 1st column:</p> <ul style="list-style-type: none"> Data usage accounting is a must. <ul style="list-style-type: none"> Publication and marketplace services can be implemented, even if the participants and main actors, i.e. system operators, aggregators, flexibility providers do not change very often. Metadata and discovery services might be needed. A metadata broker is needed in most of the flexibility use cases. 	Data discovery services can be useful for the system operator that has acquired the prosumers' data.

Table 31: Data Value Creation – int:net

Europeana: The common European data space for Cultural Heritage

Data Interoperability Requirements/Building Blocks

Data Interoperability	Requirements/Technical Solutions	Needs/Gaps
Data models and format	Data models and formats: Europeana Data Model (EDM) and associated profiles, are the interoperability solution for data exchange EDM is based on RDF and re-uses elements from the following	Deployment of brokers/connectors: The Europeana services act as a broker in the sector to bring metadata about

	<p>(linked) data standards: Dublin Core, W3C SKOS, FOAF, IIIF, W3C Web Annotation, W3C DQV, RightsStatements.org, W3C DCAT, W3C ODRL, ccREL, RDA, ADMS, EBUcore, CIDOC CRM etc.</p>	<p>cultural heritage objects available to users across Europe. It creates connectors with data partners, using the Metis aggregation service, to collect metadata on a regular basis and transform it to a single interoperable format that can be used across all organizations (i.e. EDM). On the other hand, the Europeana APIs are the connectors between the data space and the (software) users. It is not foreseen further development in the short term, given that the current approach works well for the sector. Nonetheless, they continue to investigate new methods for data acquisition and exchange.</p>
<p>Data exchange APIs</p>	<p>Europeana APIs - a range of APIs that offer access to a wide range of data available in Europeana, namely:</p> <p>Record API will share metadata that is available in the Europeana database.</p> <p>Entity API will share metadata and receive data from external sources about entities referred to in the items.</p> <p>Annotations API will receive and share additional information about the items which are contributed via external partnerships. User Sets API will manage and share information about groups of items that are created by end-users of the Europeana Website A range of IIIF APIs the Manifest API which implements the Presentation API spec used to share a display Deployment of brokers/connectors: The Europeana</p>	

	<p>services act as a broker to fetch metadata about cultural heritage objects available to users across Europe. It creates connectors with data partners, using the Metis aggregation service, to collect metadata on a regular basis and transform it to a single interoperable format that can be used across all organizations (i.e. EDM). On the other hand, the Europeana APIs are the connectors between the data space and the (software) users. It is not foreseen further development in the short term, given that the current approach works well for the sector. Nonetheless, they continue to investigate new methods for data acquisition and a version of the metadata of the items The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is used for obtaining metadata from data providers. It is also part of the Europeana APIs for data sharing.</p>	
Data provenance and traceability	Data provenance and traceability: Europeana Data Model (EDM) records some level of provenance and traceability of the metadata and data enrichment (data provider, partner project)	

Table 32: Data Interoperability Requirements - Europeana

Data Sovereignty and Trust Requirements/Building Blocks

Data Sovereignty Trust	Requirements/Technical Solutions	Needs/Gaps
Identity and access management framework	An Auth Service is used to manage user's login, accounts and access to the Europeana website and APIs. This service is supported by KeyCloak which can act as a broker to other Identity Providers (IdP) offering the possibility for users to use accounts from other services such as social platforms like Google.	Under investigation, the use of Solid in the context of new and more decentralized methods for data acquisition and exchange.
Access and usage control/policies	Each service has its own access control methods and policies. We have recently	

	started to explore the Auth Service as an SSO and access control solution which we are looking to expand to potentially all services used in the data space.	
Trust exchange/Trust Frameworks	The Europeana Data Exchange Agreement (DEA) is the central element of the Europeana Licensing Framework. The DEA structures the relationship between Europeana and its data providers.	

Table 33: Data Sovereignty and Trust Requirements - Europeana

Data Value Creation

Data Value Creation	Requirements/Technical Solutions	Needs/Gaps
Data usage accounting	<ul style="list-style-type: none"> • Matomo (Europeana website and Europeana Pro) • Internal logging tracking application (Europeana APIs) 	N/A
Metadata and discovery services	<ul style="list-style-type: none"> • Europeana website (service to access digital cultural heritage) • Europeana APIs (services to connect digital cultural heritage) • Europeana Pro (resources for professional audiences) 	
Publication and marketplace services	<ul style="list-style-type: none"> • European citizens (social media Facebook, Twitter, Pinterest, Instagram, Medium, LinkedIn, Giphy, Transcribathon.eu, DailyArt app and DailyArt magazine, Reddit) • Education (educational resources on the Europeana website, the Teaching with Europeana blog; source collections and eLearning activities with Europeana content on Historiana; integrated content in several educational learning platforms from Ministries of Education in Europe, and cultural heritage institutions or edtechs e.g. Portugal, Greece, Lucia) • Blaga Central University, The Moldavian National Digital Library, Unsplash, Canva or EUScreen among others). 	

	<ul style="list-style-type: none"> Academia and research, R&D (e.g. CLARIN Virtual Language Observatory; SSH Open Marketplace with Jupyter Notebooks for Newspapers; N/A EOSC Marketplace; Data.europea.eu, Zenodo) · Developer communities (e.g. Postman) 	
--	---	--

Table 34: Data Value Creation - Europeana

Language Data Space (LDS)

Data Interoperability Requirements/Building Blocks

Data Interoperability	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> Data models Data exchange APIs Data provenance and traceability 	<p>Data models: currently under investigation, with a focus on IDSA, Gaia-X models, and ELG (European Language Grid) model for domain-specific modeling.</p> <p>Data exchange APIs: Investigating APIs of the IDSA and Eclipse Data Space connectors and CEF Context broker.</p> <p>Data provenance and traceability: Under investigation, looking at ELG and ELRC-SHARE platforms mechanisms</p>	Metadata multilinguality

Table 35: Data Interoperability Requirements - LDS

Data Sovereignty and Trust Requirements/Building Blocks

Data Sovereignty and Trust	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> Identity and access management framework Access and usage control, policies 	<p>LDS will cater for data sovereignty (as per definition in IDS RAM)</p> <p>Identity management: examining the eID service and open protocols (such as OpenID connect and OAuth 2.0 protocol) and implementations.</p> <p>Access and usage control: under investigation; based on standard licenses, contracts, agreements, etc., considering ODRL, IDSA and Gaia-X implementations (connector and clearing house)</p> <p>LDS still to define whether they propose: federated, centralized,</p>	Find the right balance between trust and ease of use.

	<p>decentralized or hybrid architecture -> it seems that the project is considering decentralized or hybrid scenarios.</p> <p>The LDS Connector will be a key element that will enable the documentation and negotiation of data, actual exchange and recording of logging, etc.</p>	
--	---	--

Table 36: Data Sovereignty and Trust Requirements - LDS

Data Value Creation

Data Value Creation	Requirements/Technical Solutions	Needs/Gaps
<ul style="list-style-type: none"> • Self-description of data and services • Publication and Discovery • Marketplace and usage accounting • LDS mapping with DSSC: 	<p>Language data is: Text, audio, video – multimodal, language models, language descriptions, lexical and conceptual resources</p> <p>To consider: data in different languages, monolingual/multilingual/parallel, different domains, different communication scenarios, subjective, ambiguous, scope of language data is evolving</p> <p>The point of departure of all metadata schemas for language data has been the META-SHARE schema and the META-SHARE ontology. This is adapted and expanded to cater for the needs of other LR infrastructures, resulting in various application profiles, e.g., ELRC-SHARE and ELG-SHARE.</p> <p>The LDS metadata schema will be a revised and adapted application profile of these existing schemas (in particular ELG-SHARE), further informed by the IDSA and Gaia-X ontologies and by the DCAT vocabulary.</p> <p>Data usage accounting: Under investigation, considering prior work in ELG and ELRC-SHARE and other relevant recommendations.</p> <p>Publication and marketplace services: Under investigation, considering prior</p>	<p>Development of the LDS-SHARE metadata schema for language data, based mainly on ELG-SHARE.</p>

	work in ELG and other relevant recommendations.	
--	---	--

Table 37: Data Value Creation - LDS

2.1.3.3 Ethical/Cybersecurity measures

DS4Skills

- It is identified the need for new **AI building blocks** to evaluate the algorithms against reference ethical criteria, following recognized international guidelines for ethical aspects in AI developments. For example, the Organization for Economic Co-operation and Development (OECD) or the United Nations Educational, Scientific and Cultural Organization (UNESCO).
- The potential future implementations related to these suggested AI building blocks are related to the use of AI modelling platform, ethical AI, edge AI translators and image recognition support.

GREAT

- AI/analytics models and related applications will be provided and intended to adopt a suitable framework to ensure trustworthy AI applications.
- Ethical principles to avoid bias, ensure fairness, prevention of harm and non-discrimination will be considered based on current regulation, e.g., Data Act and FRAND³ basis and usage policy, as a considerable number of data sets are generated by IoT devices (e.g. weather measurements, climate observations). This will be incorporated at the design phase.
- To realize data protection regulations, providing means for easy access to control own personal data, perceived data use, data portability or data deletion through data space for the individuals, will be assessed and translated into technical requirements to identify technical enablers.

AgriDataspace

- Implementation of mechanisms that allow more (technical) protection and control over exchanged data, for example by exchanging and processing encrypted data (e.g. confidential computing)
- In the case of personal, competitive or otherwise sensitive data, mechanisms like compute-to-data, federated learning or multi-party computing could add instruments to the agro-data ecosystem that ensure privacy while retaining data control, thus enabling data sovereignty by design.

DS4SSCC

- AI/analytics models and related applications will be provided and intended to adopt a suitable framework to ensure trustworthy AI applications. Regarding this aspect, Living-in.eu MIM5: Fair Artificial Intelligence (not included in D3.1 mapping with the BBs) is about making sure that cities can have confidence that the AI and the models they use, as well as the goals the AI is programmed to achieve, are fair and transparent and that they are able to use data in a fair and transparent way.

- To realize data protection regulations, providing means for easy access to control own personal data, perceived data use, data portability or data deletion through data space for the individuals, the principles outlined by Living-in.eu MIM4: Personal Data Management will be followed.

DATES

- N/A

Int:net

- The development of AI/analytics models is essential for data analysis, as the outcomes will serve as inputs for system operations. Explainable models and the ability to easily interpret the results is fundamental to obtain accurate and reliable information to operate the power system components.
- Ethical/cybersecurity measures to be implemented to detect and avoid cyber-attacks to the power systems by using bad data (both intentional and unintentional). Other goals are related to revenues generated due to the reduction of CO2 emissions.
 - Data protection measures: data deletion mechanisms are feasible (e.g. data deletion through the data space for the individuals). Also, applying cryptography of personal data by using representative data.

Europeana

- AI applications are developed through several pilots. Enhancement of images and automatic translations are going to be deployed soon. Several data space partners experiment with AI models.
- In terms of trustworthy AI/ explainable models which outputs are easy to interpret, several AI applications that are explored so far show outputs easy to interpret (generation of metadata and improvement of content that is interpretable the same way as the original data is) even though some results may be surprising at first.
 - When considering ethical principles, e.g. fairness, prevention of harm, non-discrimination, avoidance of bias, etc. Generally, it is evaluated the quality of data that results from their AI models, and it is not observed any ethical concern of its own (there may be some in the original data though and we could propagate it). At this stage if there is bias, they are not really equipped to notice and fix it (in a systematic way).
 - The process to ensure compliance with data protection regulations, also in terms of technological enablers, and the data space principles (e.g. specific means for providing easy access to control own personal data, perceived data use, data portability or data deletion through data space for the individuals) is based on the following processes:
 1. When collecting personal data, ensuring for its safe processing, namely through informing data subjects of all the relevant conditions, collecting consent when necessary, and ensuring that all the data stays in the EU.
 2. There are internal processes in place for ensuring the deletion or anonymization of data when necessary and for respecting all other relevant obligations.
 3. When collecting and displaying persona data amongst the data made available through the data space. The responsibility to ensure this type of data is available through the data space with the data provider/s, which accept this responsibility

through the data exchange agreement. They ensure that either no personal data is shared, or that is shared legally.

LDS

- Ensure compliance of all personal data processing operations with GDPR and EUDP.
- Ensure that the LDS is in line with the principle of compliance “by design and by default” set in the GDPR and EUDPR.
- Gather all data protection analysis and assessments in a single document: Data Protection Concept:
 - Transparency, empowerment, trustworthiness, privacy, data protection and security, etc. -> Implement the common functional, legal, operational and technical aspects, such as security, identity, authentication, protocols, metadata, etc.
 - Ensure security, including via encryption mechanisms—to guarantee trustworthiness of the system, trust in it and subsequent mass-scale adoption throughout the EU.

2.1.3.4 Gap Analysis

Note:

- The rows color-coded in Pink are high priority and the ones in Blue are low priority.
- The Building Blocks that are referred in the Gap Analysis table are prior to Blueprint version 2.0

Data Space: DS4Skills

Identified Gap	Assessment
Minimum interoperability Mechanisms (MIMs)	Propose to learn about them and analyse how to leverage on them
Semantic Translators	If it is agreed a common vocabulary, the translation should not be required, but better to check with experts
Digital Wallet Interoperability	According to the EU regulation, it is desirable to propose a common digital wallet, and in case of several ones, ensure their interoperability
Manage consent forms and contracts	It seems that consent is not explicitly included as a feature in any of the two BBs, so we propose to consider the required extensions to the related BB to cope with this concept.
Personal Data Intermediaries (PDI)	Management of personal data
Anonymization & pseudonymization	Not sure if Trust is the only affected BB, but worthy to assess if essential to have
Trustworthy AI	AI related features are frequently mentioned to support the data space development and operation. Not essential at first glance but nice to have in future iterations.

Security	(Cyber) Security is in general not too present in the DS concept and many EU topics about how to integrate security mechanisms into data spaces.
-----------------	--

Table 38: Gap Analysis Data Space DS4Skills

Data Space: Green Deal - GREAT

Identified Gap	Assessment
Trustworthy AI	AI related features are frequently mentioned to support the data space development and operation. Not essential at first glance but nice to have in future iterations.
Introduction of a DS digital ecosystem soft infrastructure for multisector DS	Which is the concrete request? Is he referring to the BBs or to a deployment approach? Not clear how to address the request
Dataset Transformer and Data Catalog (Logical component)	Catalog is already included in the Publication and Discovery BB. The Transformer components rely on the fact that the data space will not use common standards for data models and then need to be converted to a common standard. It can make sense in cross-domain data sharing when different data spaces have agreed different data models. Like row 4, but to be checked
Auditor component, PUI provider and PUI resolver components	Consideration for the Provenance and Traceability BB
Federated AAI, XACML PEP	Consideration for the Identity Management BB
Ethical considerations	Which is the concrete request of extension? Request T4.1 to clarify with the data space

Table 39: Gap Analysis Data Space Green Deal - GREAT

Data Space: AgriData Space

Identified Gap	Assessment
Security	(Cyber) Security is in general not too present in the DS conception and many EU topics about how to integrate security mechanisms into data spaces.
Marketplaces as Data clearing houses	Not clear which is the requirement. Request T4.1 to get more info. Not a request it is a statement.
Code of conduct for data sharing	Propose to create this BB or justify if included in an existing one, as one of the capabilities

Table 40: Gap Analysis Data Space AgriData Space

Data Space: DS4DSSCC

Identified Gap	Assessment
Minimum interoperability Mechanisms (MIMs)	Propose to learn about them and analyze how to leverage on them

Personal Data Intermediaries (PDIs)	Management of personal data
Trustworthy AI	AI related features are frequently mentioned to support the data space development and operation. Not essential at first glance but nice to have in future iterations.
Federation of ODMS	The approach of the marketplace is already federated. No specific distinction about Open Data. We recommend to experts in Marketplaces to analyse the request
Data Cooperation Canvas	Propose to be adopted by the Organizational blueprint as it has been proven in several use cases at the Smart Cities and Communities data space and it has been very much appreciated

Table 41: Gap Analysis Data Space DS4DSSCC

Data Space: DATES

Identified Gap	Assessment
Personal Data Intermediaries (PDIs)	Management of personal data
Connector-as-a-Service	Consideration for the Data Spaces Protocol EG
Use of standards for data quality	Data quality is repetitive request to be included. Several BBs can be affected. We propose the consideration of the experts in the BB
Code of conduct for data sharing	Propose to create this BB or justify if included in an existing one, as one of the capabilities

Table 42: Gap Analysis Data Space DATES

Data Space: Int:net

Potential Request	Assessment
Legal agreements for participants and 3rd parties to be fully compliant with the GDPR	It is not explicitly mentioned in the description of the two related BBs

Table 43: Gap Analysis Data Space Int:net

Data Space: Europeana

Potential Request	Assessment
Trustworthy AI	AI related features are frequently mentioned to support the data space development and operation. Not essential at first glance but nice to have in future iterations.
Security	(Cyber) Security is in general not too present in the DS conception and many EU topics about how to integrate security mechanisms into data spaces.
Wide range of APIs	Which is the concrete request of extension? Request T4.1 to clarify with the data space

Enrich metadata	Consideration for the Data, Services and Offerings Description BB
Licensing framework	IPR terms are not explicit in the contractual framework, so perhaps it can be extended with this aspect

Table 44: Gap Analysis Data Space Europeana

Data Space: LDS

Potential Request	Assessment
Trustworthy AI	AI related features are frequently mentioned to support the data space development and operation. Not essential at first glance but nice to have in future iterations.
Security	(Cyber) Security is in general not too present in the DS conception and many EU topics about how to integrate security mechanisms into data spaces.

Table 45: Gap Analysis Data Space LDS

2.1.4 Conclusions

The initiatives were exploring the most adequate reference architectures, common building blocks, toolboxes, and technology standards to be implemented in the deployment phase. The organizational and business blocks, use cases and the minimum digital services foreseen each data space play a fundamental role to determine the technical grounds and software components to be developed and/or implemented. Below a list of the major findings from the technical perspective:

Most of the initiatives are analysing a variety of existing reference architectures to build upon, such as IDS-RAM and Gaia-X architecture or sector-specific, e.g. FIWARE reference architecture for smart cities and in combination with the re-use or extensions of other technologies that have been developed in other sector related projects, e.g. the Language Data Space (LDS) is designing an architecture based on relevant components from FIWARE, IDSA and Gaia-X as well as on previous language resource infrastructures, such as ELRC (European Language Resource Coordination), ELG (European Language Grid), CLARIN (Common Language Resources and Technology Infrastructure), etc. Other initiatives which are in a more advanced stage of maturity (Implementation), as for example, the GDI-Genomic Data Infrastructure project, has been primary focused on implementing the architecture coming from previous initiatives such as ELIXIR, standards and frameworks conceived by the GA4GH (Global Alliance for Genomics and Health initiative) or the DS4SSCC data space (Smart cities and Communities) which has adopted the Minimal Interoperability Mechanisms (MIMs Plus) promoted by the Living-in.eu.

In terms of technical building blocks and specific technology solutions and processes that are necessary for ensuring the expected functionalities of each building block, all the projects follow a common standardization approach comprising the gathering of requirements to achieve interoperability and cross-border access to data by engaging with all the relevant stakeholders of a specific domain or domains. These initial set of requirements for data interoperability, data quality, specifications from relevant standards and other relevant aspects (e.g. legal, social) were translated into technical specifications and some high-level guidelines for implementation.

In general, all the initiatives initially evaluated and incorporated the building blocks outlined in OpenDEI, corresponding to the technical pillars of Data Interoperability, Data Sovereignty, Trust, and Data Value Creation in the initial stages. However, as the DSSC released the first blueprint and subsequent updates, the adoption of the DSSC approach became evident in the latest deliveries of the CSAs within the CoP.

Main technological challenges are very diverse due to the inherent complexity of setting up data spaces and the variety of data, use cases, stakeholders and local, national regulation among other factors. Some examples, just to highlight some of them are related to: i) differences in terms of technology maturity and size of different key players involved in the data space (e.g. farmers and large industry in the agrifood sector) together with the different needs regarding the use of data, are of the challenges for scaling up; ii) how to achieve the federation of national nodes with multiple infrastructures (centralized, distributed, federated or hybrid) with multiple governance frameworks (e.g. GDI), iii) many of the initiatives are investigating the most appropriate technological means towards a more decentralized approach as for example the use of Solid specifications (Home · Solid (solidproject.org))

Other aspects that have been considered by all the initiatives are related to the use of Trustworthy AI applications, ethical considerations, (cyber) security and personal data protection. These aspects are not explicitly considered in the existing frameworks and will be subject to further analysis within the DSSC project to derive extensions of building blocks, identifying new ones or add additional specifications in the Blueprint. For instance, the need of a personal data intermediary to adequately handle personal data beyond the GDPR requirements is a common aspect to be further analysed and raised by several initiatives such as DS4SKILLS, DS4SSCC and DATES.

2.1.5 Results

Task	Document	External reference	Document internal reference
T4.1	Requirements		Chapter 2.1
T4.1	Identified Building Blocks Overview	Blueprint 0.5	Chapters 2.1 and 2.2

Table 46: Summary of results of requirements collection

2.2 Building Block identification and data space development

2.2.1 Scope and purpose

Data space designers are the active developers of data spaces facing challenges at various levels in their development effort. An ecosystem, which the data space builds upon, is a complicated context, starting with the motivations for data space participants, concerns on trust and benefit, questions on governance and

contracts, and decisions on technological choices. Typical issues in this multifaceted endeavour are how one can build a sustainable, functionable, and lucrative (social) system to support trusted collaboration, meaningful data sharing, for the common benefit of stakeholders and how one can create a scalable solution supporting the bigger purpose. It is evident that a common methodology would be able to provide economies of scale in such a complex environment where everyone adheres to the same approach, speaks the same language, and (re-)uses solutions which provide interoperability, not only to the data space, but between many data spaces.

This chapter describes such a common approach on a more detailed level: the **co-creation method** is a practical methodology for a data space designer that gives instructions on how **building blocks** must be considered when setting up a data space. While the co-creation method helps and guides in using the building blocks for data spaces, building blocks are the ingredients for the data space and therefore at the heart of data space development. More specifically, building blocks are the means to understand what elements one should take into consideration when planning to set up a data space. In addition, to that, they are the resources for data space developers who want to go deeper on what exactly is available or even recommended as a solid foundation for a data space. (cf. Figure 4)

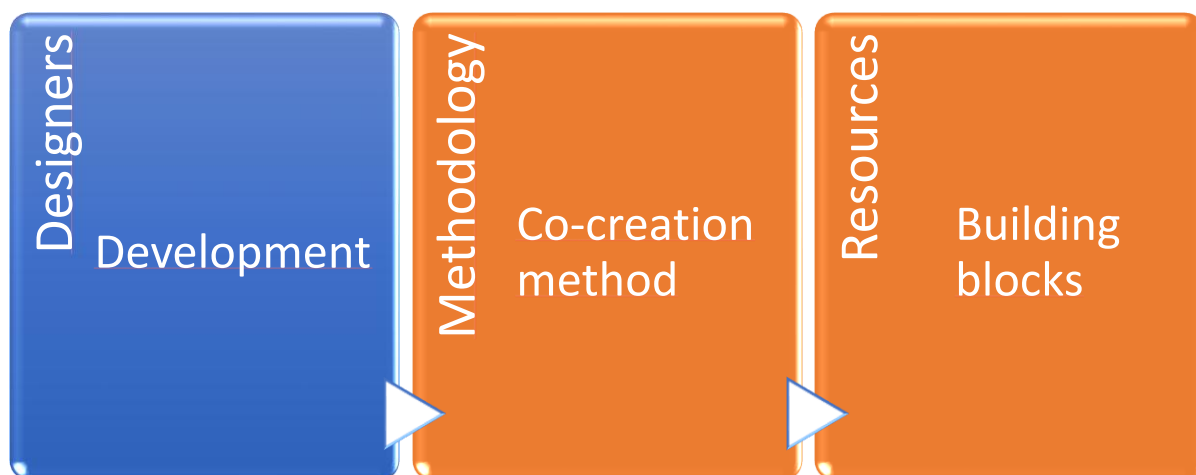


Figure 4: Designers, development activity and the relationship between co-creation method and building blocks

This chapter explains the relationship of design methodology in form how co-creation method has been composed. We delve into how the process of data space building blocks identification occurred, exposing the rationale of choices made, and what are the end results from that work.

2.2.2 Methodology

The approach of DSSC project is community-driven and the outcomes are the crystallization of many interventions: surveys, thematic group meetings, and interviews. The intention of this section is not to list all these occasions but to give an overview of the timeline and examples of methods exploited in the course of work. Furthermore, identification of building blocks has been a creative effort, where something new emerges from the need of the community. Thus, catching, analysing, and modelling the often-complex reality into a digestible conceptualization is a process of workshops, ideas, summaries, and further analysis.

Although releasing some of the ideas in a Starter Kit earlier in 2022, the first attempts to collect a comprehensive understanding of underlying building blocks were done in March 2023 in two workshops. One of the cornerstones in this effort was the work done in the previous OpenDEI project, which provided a framework and the idea of building blocks and design principles.

To concretize and share the ideas a document called ‘building block taxonomy’ was used on the one hand to find a common understanding, but on the other hand to communicate with the community of practice. Eventually, the ‘building block taxonomy’ was renamed as ‘**building blocks overview**’, the name which we use from now on. The paper was circulated within DSSC and with selected experts starting May 2023; the first release to the DSSC Community of Practice was made in early July 2023. Feedback was collected from both CoP and DSSC Architectural Board on several occasions in August 2023. Security Advisory Board (SAB) passed the material in its two review rounds.

The main new original finding was that there is a pronounced need to provide a conceptual “home” for all **organizational aspects**. Although originally started as a separate paper, it soon was clear that this work is part of the DSSC building blocks structure resulting in the adoption of a distinct business and organisational building blocks part (green elements in Figure 5). In August and September 2023, the building blocks overview was used as a basis for Blueprint v0.5 in DSSC platform release. Later, a smooth hand over from identification and distillation (cf. figure 2) towards synthesis and deeper analysis of the building blocks, first validity, and second their content.

2.2.3 Contributions to data spaces Blueprint

2.2.3.1 Building Blocks

The aim of this section is to present the ‘Building Blocks Overview’ and describe the contributions of the identification work package to the common asset, i.e. the Blueprint. To start with, figure 4 depicts the current ‘Building Blocks Overview’ as it is released in Blueprint 2.0. Furthermore, it clearly shows the differentiation between business and organisational building blocks (green elements) and technical building blocks (blue elements). Additionally, the internal sub-structure of these two domains is also clearly shown. The chapter, however, is not going to reproduce actual content which is available at <https://dssc.eu/> website as part of the Blueprint, but we delve more into the rationale and development process of the building blocks, and how identification effort has contributed to the work.

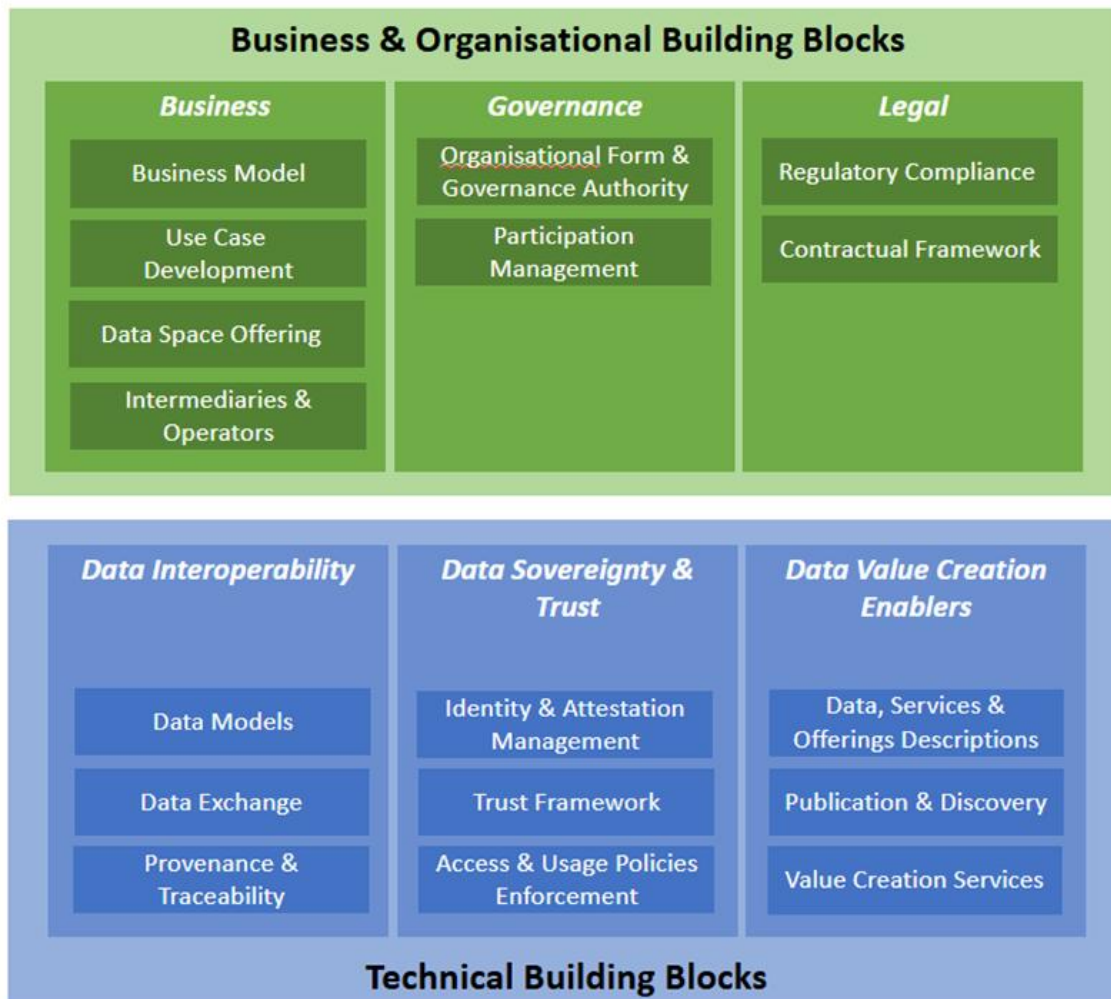


Figure 5: Building Block overview in Blueprint 2.0

Business building blocks

For the business building blocks we started in the Starter kit asset (cf. Figure 1) with describing the Business perspective that included value and models. In this description various business case patterns are identified as well as various examples of these business case patterns.⁵ Besides that, Business-model radars have been presented for data-driven applications and data space participants. After that we described the building block overview and incorporated it in the blueprint with 4 business building blocks: Business Model development, Use case development, Data product development, Data space intermediaries. After that we updated the building block content and names, meaning that the four business building blocks are called Business model, Data product, Use case development, Data product and Intermediaries and operators. In the next version of the blueprint v2.0 the Data Product BB will be renamed into Data Space offering. This will be done to focus more on the business perspective of the data product and services. The technical

⁵ See [Starter Kit for Data Space Designers | Version 1.0 | March 2023 - Starter Kit - Data Spaces Support Centre](#)

perspective of the data product will be covered on the technical side of the blueprint. The aim for the next two versions of the blueprint is to make the business building blocks more user oriented.

These four building blocks have been developed in cooperation with the business expert group, with input from the business thematic group. To make the business blocks more user oriented we are currently working on a paper about data space examples and their business models.

Governance building blocks

The main starting point for the work on the building blocks under the DSSC blueprint, the [OpenDEI design principles for data spaces](#), did not yet put a strong emphasis on governance. Nevertheless, the DSSC received many questions from the community on governance issues, ranging from how to identify the relevant stakeholders for a data space initiative to the legal form for a data space governance authority. The DSSC has since put much effort into expanding more on the governance side of data spaces. The BB taxonomy paper, which was preceded by a number of workshops, initially put forward “data governance framework”, “roles and responsibilities” and “decision-making, rules and policies” as candidate building blocks to cover governance issues within a data space. Governance of data spaces remains, however, a complex topic with many uncertainties and open questions, in particular compared with the topics covered by the technical building blocks on which discussions are generally considered more mature. This in part explains why the governance building blocks have evolved significantly throughout the different iterations of the DSSC blueprint.

An often-recurring request from the community was to provide more information on different options for setting up a governance authority, such as what the role is of European Digital Infrastructure Consortia, potential legal forms, etc. In the Blueprint v1.0 the “organisational governance” building block was revised to provide with more information on legal forms (e.g., unincorporated versus incorporated) and the building block was renamed to “organisational form and governance authority” to clarify that its scope envelops both the organisational form of the data space and the legal form of the governance authority.

The “data sharing governance” was also revised. Initially, the emphasis in version 0.5 was on setting “rules and standards for participants” and “establishing and organising enabling services”. Feedback from the community was to focus more on the onboarding and offboarding of participants. At the same time, the business building blocks cover more extensively the enabling services. Consequently, the building block was revised with the release of blueprint v1.0 and renamed “participation management”. With the release of Blueprint v2.0, this building block will also cover what internal data governance practices can be required of individual participants to successfully partake in a data space.

Legal building blocks

Initial work done on legal issues related to data spaces was done in preparation of the DSSC Starter Kit v1.0. In line with the principles of the BLOFT framework (comprising business, legal, operational, functional and technical) the DSSC Starter Kit contained separate sections on business, legal and operational issues. The initial chapter on legal was inspired by the prior work done by the [EUHubs4Data project](#), expanding on the pillars identified by EUH4D to make them relevant to different data spaces described in the European Data Strategy. These areas identified in the Starter Kit were:

- Cross-cutting legal frameworks (relating to, e.g., contract law, data protection, intellectual property, competition law and cybersecurity)

- Organisational aspects
- Contractual (or transactional) dimension

In the building block taxonomy developed by the DSSC, the “legal pillar” eventually includes two building blocks:

- The “Regulatory Compliance” building blocks focuses on cross-cutting legal frameworks, with its current iteration (v1.5) focusing both on compliance at the level of the data space and facilitation of compliance within a data space by its participants.
- The “Contractual Framework” building block covers both the high-level contractual framework required for a data space (constitutive agreement, terms and conditions, etc.) and agreements at the transactional level.

The elements identified under “organisational aspects” are integrated either into the governance building blocks (e.g., legal form is part of the “Organisational Form and Governance Authority” building block) or into the “Contractual Framework” building block (e.g., information on agreements at the data space level).

Other candidate building blocks that were identified but were not included in the taxonomy (i.e., related to change management and dispute resolution) are (intended to be) covered within the existing legal building blocks.

Regulatory compliance

In preparation of the draft of the “Regulatory compliance” building block (DSSC Blueprint v0.5), a legal mapping exercise was conducted. The purpose was to scope the initial version of this building block. The initial version of the legal mapping covered (i) data-related legislation (comprising of new data legislation such as the DGA and DA, AI-legislation, non-personal data and personal data), (ii) intellectual property rights, covering trade secrets, copyright and sui-generis database rights, (iii) contract law, (iv) platform regulation, (v) competition law, (vi) trust & security and (vii) sectoral legislation. For each of these elements (except for sectorial legislation), the building block contained a short paragraph on the relevance to data spaces. The preparatory work for this building block built upon pre-existing work done by the Support Centre for Data Sharing and EUHubs4Data.

In version 1.0 of the DSSC Blueprint, the content was significantly reworked. Rather than a concise overview of different “clusters” of relevant legislation, the attempt was to provide more detailed information on the different relevant legal frameworks. In order to present the contents of the building block in a way that provides more accessibility (i.e., customers are easily able to filter out the information they are interested in), the contents were structured through the perspective of “triggers” (essential elements or criteria whose existence or specific function in a given context result in the application of the legal framework). Work on identifying the relevant triggers was inspired on research into automated compliance. In Blueprint v1.0, three types of triggers were identified: data (comprising personal data, intellectual property/trade secret protected data), actors/participants (comprising of data intermediation service providers, data altruism organisations, public sector bodies and participants with specific market power) and use cases (particularly referring to highly regulated sectors).

Version 1.5 of the DSSC Blueprint built further upon version 1.0, introducing more detailed information related to the different triggers, including providing some examples of specific use cases. It also introduced specific provisions related to data spaces, stemming from the Data Act and European Health Data Space proposal.

Feedback from the community (gathered through the Governance Thematic Group and Community of Practice, organised under WP2) learned that the regulatory compliance building block was considered a good source of information. The concept of triggers was found useful and should be expanded upon, introducing more granularity. However, the feedback also learned that the major challenge is bringing it all into practice. The community clearly requested more focus on resources and tools to make the content more actionable. Examples given include templates, decision trees, check lists, etc. In response to this feedback and to reduce the overall length of the building block, blueprint v2.0 will include a more layered approach. The building block will provide with more visual content (including checklists and decision trees), whereas more in-depth information will be included in a separate document, to be further developed within WP4 (T4.6 in particular), that will provide a supporting resource for the DSSC Blueprint.

Contractual framework:

The work on the contractual framework building block was to a large extent informed by the work on the legal mapping, as well as the analysis of contracts and contractual needs. Specifically relevant in this regard, were the [SITRA rulebook](#), which is used by many as the basis for a governance structure and contractual framework, and the work of the [Support Centre for Data Sharing](#) on model contract terms.

The initial version of this building block provided a concise overview of various agreements that would be required to establish the respective rights and responsibilities of data space participants, providers of enabling services and the data governance authority. Additionally, it listed a few resources that a data space could provide to allow transaction participants to regulate and execute specific data transactions (e.g., vocabularies and contract templates).

In the version 1.0 of the blueprint, the content of this building block was expanded, mainly with regards to the agreements at the level of the data space. It contained a section on the different agreements (constitutive agreement, general terms and conditions, accession agreements and agreements related to enabling services) and provided of example agreement-specific clauses of the constitutive agreement and general terms and conditions.

With version 1.5 of the blueprint, this building block underwent further refinement with the inclusion of data products/data space offerings to more closely align with the business and technical building blocks. At the same time, interdisciplinary discussions were started on the contractual framework and its alignment with other building blocks, including the “Access & usage policies enforcement” building block. However, these discussions did not result in any major changes already in version 1.5 but continued into the writing phase for DSSC blueprint v2.0.

With Blueprint v2.0, this building block brings an increased emphasis on services and smart contracts. Next to this, realignment was needed with the newly renamed “Data space offerings” and the revised glossary entry “data product”. Finally, an effort was made to further align this building block with technical building blocks, such as the “Provenance and traceability” building block.

2.2.3.2 Co-creation method

The co-creation method is meant to guide data spaces and data space initiatives (and their participants) through setting up, developing and operating a data space. With the co-creation method, the DSSC aims to provide a practical and hands-on method that gives instruction on how building blocks must be considered whilst constructing a data space. The co-creation method will guide the data space initiatives through the

different stages of the development cycle. Finally, the co-creation method should support data spaces and data space initiatives on different maturity levels.

The co-creation method explains how to apply the building blocks described in the blueprint. Additionally, it defines how the building blocks (containing specific functionalities or capabilities) and components described in the blueprint interact with one another to establish a data space. Furthermore, it explains how to decide which implementation of the building block is required or advised for a data space initiative.

The co-creation method provides a flowchart of the different data space processes required to set up a data space. It was for the first time incorporated in Blueprint v1.0 (see Figure 6), in which each data space process included a brief description, the goals associated with each process, and high-level steps, actions and considerations required per process. Finally, the relevant building blocks for a data space process and the connections between the data space processes are described. In the versions v1.5 and v2.0 as well as in future versions the co-creation method is further expanded. The DSSC developed this method based on earlier work (e.g. European Zero waste project) based on which it is published in the blueprint. The co-creation method has an own expert group.

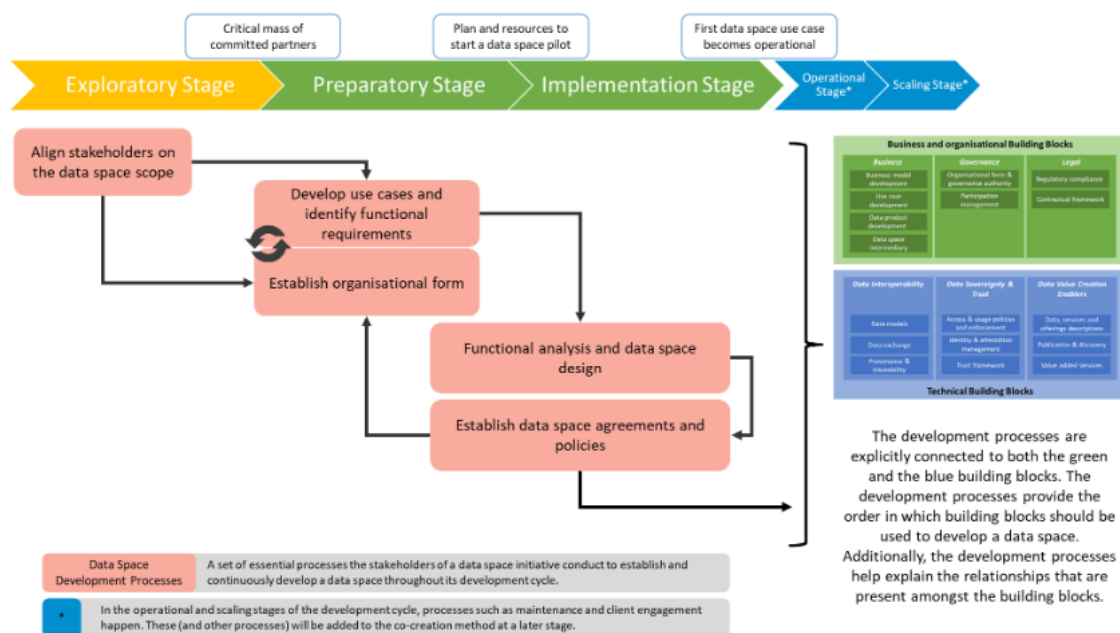


Figure 6: Development processes and the relation to the building blocks presented in v1.0 of the blueprint

2.2.4 Results

Task	Document	External reference	Document internal reference
T4.1	Identified Building Blocks Overview	Blueprint 0.5	Chapters 2.1 and 2.2

T4.5	Identification of and contribution to organisational, business, governance building blocks	Blueprint v0.5 , v1.5 , v2.0	Chapter 2.2.3
T4.6	Identification of legal building blocks	Legal building blocks in blueprint	Chapter 2.2.3.1
T4.6	DSSC role in addressing legal challenges	Blogpost	-

Table 47: Summary of results of building block identification

2.3 Standards landscape

2.3.1 Scope and purpose

Standards are an essential element to ensure the interoperability in a data space. Thus, their identification, monitoring and engagement are at the core of the DSSC. Originally in the contract, the plan was to collect the standards (domain and cross-cutting) from the various data spaces. However, in discussion with EC, the activity was extended to cover also liaison with standardization bodies and contribution to standardization reports.

The figure below summarizes the diverse activities included in the project in relation to standards:

- **Collection of the standards landscape.** The first need was to discover and collect the most used and adopted standards by Common European Data Spaces. Thus, through various interviews and surveys, we made an overview of the most used standards by them and provided an analysis to give some insights and identify the cross-domain standards.
- **Cross-domain standards recommendation.** Leveraging on the conducted analysis of standards across all data spaces, and relying on the experts working in the blueprint, the DSSC proposes some cross-domain standards which are foundational at the basis and foster the cross-data spaces interoperability. These standards are clearly stated in the blueprint description, here just mentioned: W3C DCAT, W3C Verifiable Credentials and ODRL.
- **Alignment and contribution to the European standardization activities.** Every standardization activity under DSSC must be aligned with other existing initiatives at European level and in full synchronization with the EC directives. Thus, the foreseen plan and activities were discussed and aligned with the EC before putting them into place. The main contribution of DSSC to the European initiatives is:
 - Provide inputs to the standardization reports and workshops (i.e. High Level Forum standardization report or ICT standardization rolling plan)
 - Identify gaps for the new standardization request proposed by the EC
 - Report standardization about data spaces to the European Data Innovation Board (EDIB)

- **Engagement with Standard Developing Organizations (SDOs).** Collecting, recommending and aligning are not enough to provide a complete value, so DSSC decides also to engage with those SDOs in relation to data spaces standardization. The goal is twofold: follow and be aware of the ongoing committees and workshops defining standards; and influence them by informing about project results. The DSSC role in these committees/workshops will be mainly as observer and advisor.

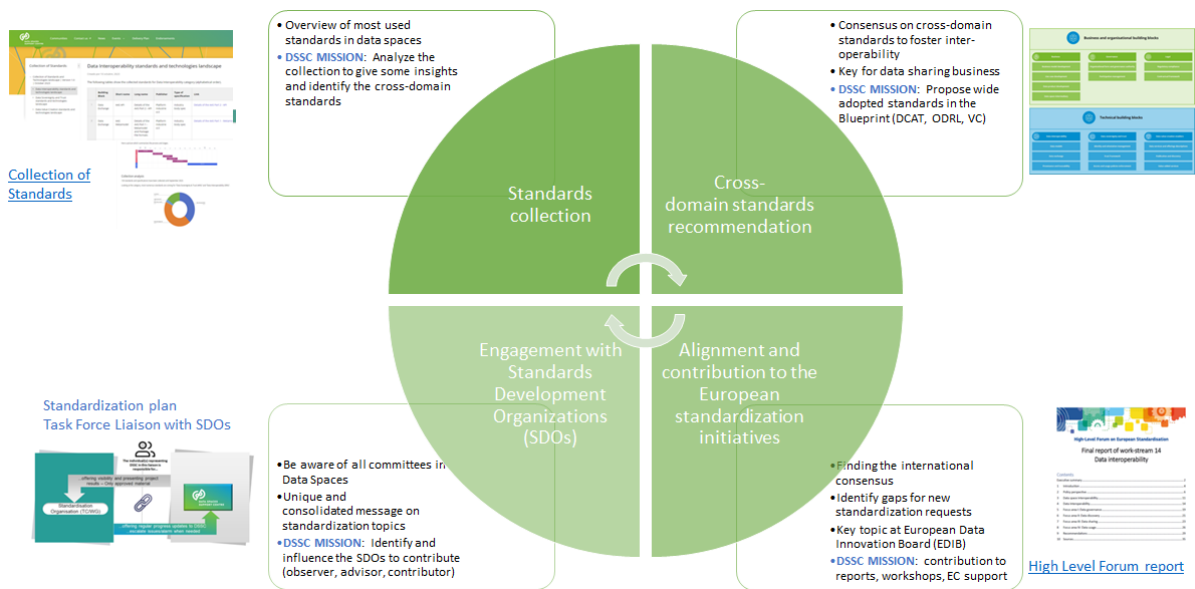


Figure 7: Standardization activities in DSSC

2.3.2 Standardization relevance and activities in DSSC

The standards are technical norms developed by consensus, accessible for use, published by a recognized entity and of voluntary compliance. They favor common understanding and interoperability amongst participants in certain domains. Complementary the regulations and legislations are legal norms of mandatory compliance that may use standards for the implementation of the norm (e.g. the Data Act is a mandatory regulation in Europe that may require the use of certain standards for being deployed in practice).

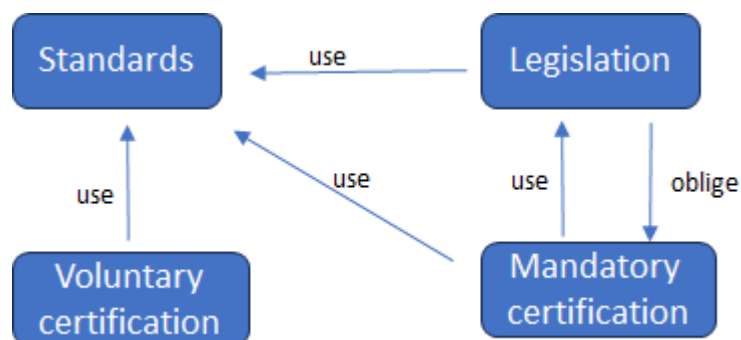


Figure 8. Standards vs Legislation

The standards provide different types of documents: technical specifications, technical reports, workshop agreements and norms; and the consensus instrument used to be the technical committees. Typically, the elaboration of a standard follows the following process:

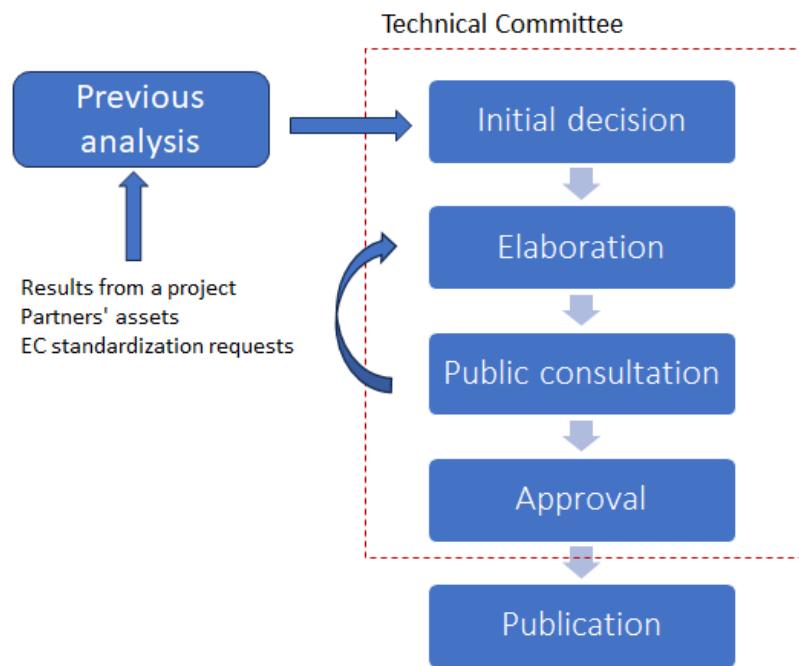


Figure 9. Usual standardization process

Before a Technical Committee for a concrete standard creation or modification is launched, a previous analysis is required to assess the convenience of starting the process. This analysis can be started from the results of a research project, from the own assets of partners in the project or by proposal of the EC through the standardization requests by DG CNECT G1 unit. All the standards require launching a public consultation before any kind of approval step to ensure the wider consensus as possible.

Due to the nature of the DSSC, as a funded project, its purpose is not to contribute formally to standardization initiatives, but to observe and advise various ongoing committees and workshops. Some of the DSSC partners are active in some related SDOs and directly contribute to the standards definitions.

Once the standardization activity was decided, the main action was defining the goal and scope of the plan. Thus, in alignment with the DSSC partners, task T4.3 elaborated a standardization roadmap for the DSSC in order to cope with the following objectives:

- Collect all standardization activities carried out by DSSC partners in relation to data spaces.
- Map the planned standardization activities with the collection of standards (following DSSC taxonomy of BBs).
- Identify gaps and propose prioritized standardization actions.
- Engage the relevant standardization bodies in coordination with T2.4 in charge of the engagement activities.
- Align with the various standardization initiatives as suggested by the EC (High Level Forum, Multi-stakeholders platform for ICT standardization)

- Provide joint recommendations to EDIB with these initiatives on possible standardization requests.
- Create a common message about standardization on behalf of the DSSC project.

Bearing in mind these objectives and considering the available resources to put in this focus, the DSSC has decided to cope initially only with formal standards proposals. Thus, the collection of standards delivered by DSSC as the corresponding asset covers all types of standards except the ones promoted by a sole company, while the recommendations about standardization requests will tackle only the formal standardization



Figure 10. DSSC focus on standards

Another aspect of the scope is the geographical application of the standardization bodies. The DSSC priority will be on European standardization bodies like ETSI, CEN or CENELEC, but without preventing any international standard if considered relevant for the data spaces.

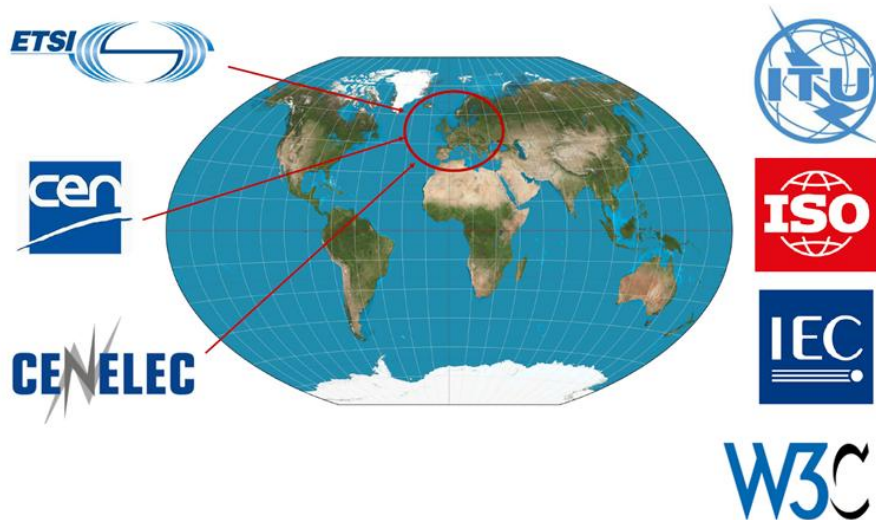


Figure 11. European and international standardization bodies in scope of DSSC

But **why are standards so relevant in data spaces?** First of all, because data spaces require to be compliant with the current EU regulation on data sharing (Data Governance Act (DGA), Data Act (DA)) and adhere to the FAIR principles, which recommend that data must be accessible in a Findable, Accessible, Interoperable and Reusable.

Standards are also essential for achieving interoperability, which means having mechanisms for data interoperability (like vocabularies, APIs, metadata...); addressing the data spaces interoperability by the use of governance schemas, catalogues and agreements; and fostering the cross-data spaces interoperability and federation. Without the adoption of common standards in multiple domains, the ambition of sharing data across all of them will be a utopia.

Finally, European standards must be aligned with international standards since most of the business are global. There is international aim at reaching the regulatory convergence. Promoting the European standards as international standards would favour the access of European companies to global market.

As said, the DSSC will not propose any specific standardization process, but it will collect, analyse and identify possible standards that could be used, extended or proposed by any organization in Europe willing to do it. Therefore, the **DSSC strategy towards the standardization** leverages on these pillars:

- DSSC will foster their partners and data spaces community to participate in standardization technical committees to propose extensions of existing standards or creation of new ones in line with the recommendations provided by the DSSC and liaised initiatives.
- DSSC will find the liaison with related standardization bodies by:
 - Attending related technical committees to provide feedback (no vote) and present DSSC assets and view.
 - Commenting on draft standardization proposals during public consultations.
 - Facilitating the data spaces community the engagement with SDOs
- DSSC will collaborate with other initiatives (i.e. [High Level Forum on European standardization](#), [European Multi-stakeholders platform for ICT standardization](#)) in joint recommendations for future standardization actions.

The following picture illustrates the role of the DSSC in the overall landscape of players. The approved laws by the European parliament land in standardization requests proposed by the European Commission. In relation to data related laws, the European Data Innovation Board (EDIB) has been created to implement these laws in all members states and provide directions for the required standards. The corresponding SDOs take the mission to develop the requested standards by EC, and they can be participated by organizations as observers. The DSSC plays this role by establishing a liaison with SDOs and reporting this activity to the EDIB, while contributing to other initiatives like the ones referred to above.

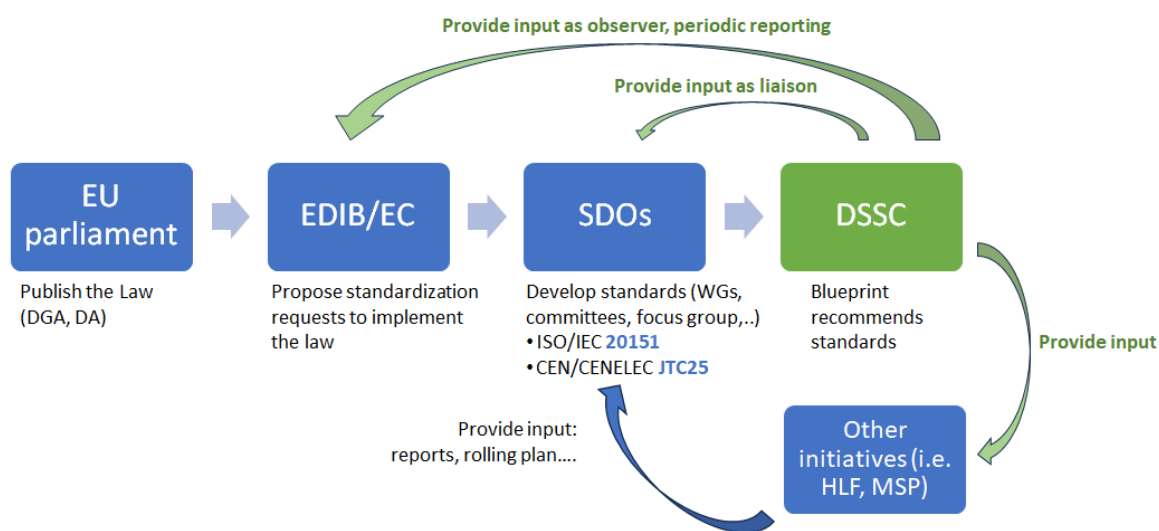


Figure 12: DSSC role in standardization process

The main challenge of all this standardization process is **being concrete within the regulation**. In a bottom-up direction, we have data spaces already using both cross-domain and specific standards, as well as with technologies and implementations in use. While, in a top-down direction, we have EU institutions regulating the context and formal standardization bodies developing industry-driven standards. Both directions cannot ignore each other and must work in the convergence which minimizes the effort of investment and maximizes the consensus.

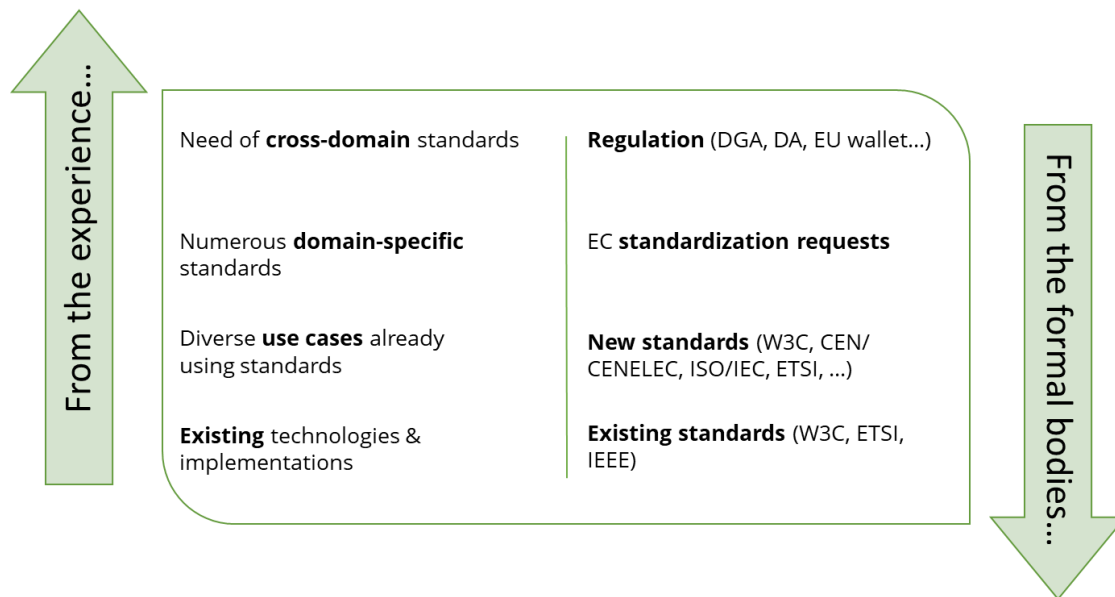


Figure 13: Experience vs regulation

Standardization activities timeline

Considering the stated objectives and the alignment with the EC proposal on the process, the following timeline graphic shows the work plan for the DSSC in this regard. Once the collection of standards has been fully consolidated by integrating all the inputs from relevant stakeholders, we elaborated the present plan which includes the ongoing standardization activities of the DSSC partners as well as a scanning of relevant technical committees and workshops working on standards about or in relation to data spaces (data trust, data sharing, data exchange, data privacy, data publication, etc).

Then, both the collected activities and the scanned committees were mapped into our structured collection of standards to find correspondence with the BBs taxonomy on one hand, and to identify gaps on the other hand. This work was presented in the online workshop held by HLF and EC on 22nd September 2023, and together with the other initiatives, a set of recommendations was proposed to the EDIB when launched. Another workshop in the context of EBDVF in Valencia was used to consolidate these recommendations for the DG CNECT G1 unit to prepare the corresponding standardization requests.

In parallel, the collection of candidate standards was included in Blueprint v0.5 and circulated to the Expert Groups for the consideration in the Technical Specifications of the Building Blocks, where they propose the most recommendable standards to use for the implementation of the BBs. Since then, a continuous monitoring and update has been carried out, following the requests collected in our ticketing system (Jira issues from the [support form](#)).

In May 2024, we launch the task force between T4.3 and T2.4 about the liaison with SDOs. Since then, we have established already a liaison with four SDOs. Despite the original proposal, the T4.3 is planned to finish in March 2025, we proposed the extension till the end of the project in order to continue monitoring the updates in the collection of standards and the liaison with ongoing and upcoming SDOs.

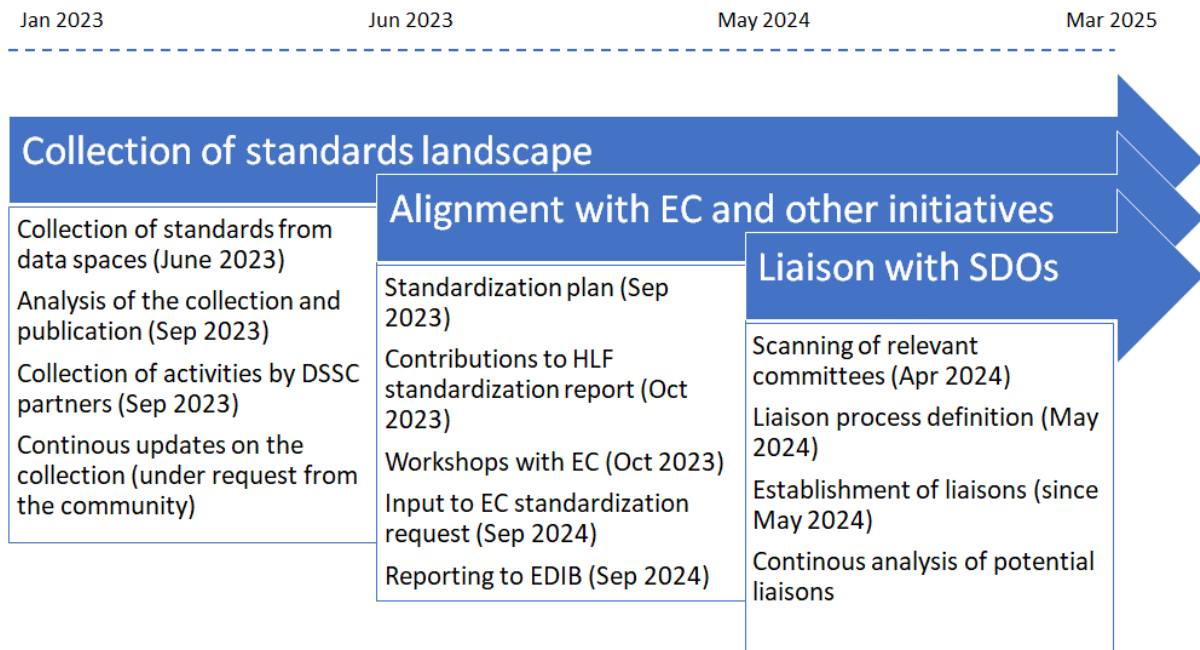


Figure 14. Timeline for standardization activities in DSSC

DSSC Partners activities

Following the previously mentioned strategy of the DSSC towards standardization, the DSSC will foster that their partners and community participate in standardization processes.

In order to know the degree of involvement of DSSC partners in standardization activities, one of the actions in the plan is the collection of the ongoing and foreseen activities by partners.

In order to help them in identifying these activities several questions have been put to them. They should communicate any of these actions, so they can be included in this standardization plan:

- Have you initiated any standardization process?
- Are you participating in some technical committee?
- Do you plan to propose a standardization action?
- Would you like to do it in collaboration with other partners?
- Do you have engagement with some standardization body?

Thus, a template has been created in the DSSC Support Platform at Confluence, in the Data Space Standards and Technologies landscape space, and the partners have been requested to include their activities.

The information that has been requested is:

- Involved DSSC partner or partners in the standardization activity.
- Related Building Blocks in the DSSC taxonomy that the standard request is affecting.
- Related standardization body where the activity is taking place.

- Technical committee or sub-committee within the SDO that has been created or where the standard is being discussed.
- Stage in the standardization process.
- Type of document is under development. The documents can be a workshop agreement, a technical report or technical specifications.
 - Description of the proposed standard or contribution
- Relevance for the data spaces
- Link to the public information of the committee, the proposal or the document if it is already under public consultation process.

The information provided by the partners involved is summarized in the following table:

DSSC Partner	Related DSSC asset	SDO (CEN, CENELEC, ETSI, ISO, ITU, IEC, W3C)	Technical Committee (and/or subcommittee)
IDSA	Organisational form and governance authority. Data sovereignty and trust. Data interoperability. Trust framework. Control and data plane. Data Exchange	CNE/CENELEC	JTC25 Data management, data spaces, cloud and edge
TNO, Fraunhofer, BDVA, GAIA-X, IDSA, FIWARE	Data Exchange & Trust Building Blocks	CEN	Trusted Data Transaction Workshop
IDSA	Organisational form and governance authority. Foundational concepts on Dataspace interoperability, Data sovereignty and trust BB	ISO/IEC	ISO/IEC SC38 – Information technology — Cloud computing and distributed platforms CD20151 dataspace concepts and characteristics
Fraunhofer, IDSA	Control&Data plane - Dataspace protocol Foundational Standard	ISO	ISO Fast Pas under Eclipse Dataspace Working Group – Dataspace protocol
Fraunhofer	Decentralize Claims protocol	ISO	Eclipse Dataspace Working Group
IDSA	Data Space reference architecture	IEEE	IEEE on P3800 Data Trading System
FIWARE	Data Exchange Building Block	ETSI	CIM Specialist Task Force

FIWARE	Data Exchange Building Block	ETSI	CIM Testing Task Force
--------	------------------------------	------	------------------------

Table 48: DSSC partner activities in standardisation summary

2.3.3 The collection of standards landscape in data spaces

The Data Spaces Standards and Technologies landscape is a map of standards and reference implementations as candidates for implementing a building block. As described in the DSSC glossary, a building block is a basic unit of functionality required for the creation and deployment of a data space. There exist business & organizational and technical building blocks. Standards are mostly related to the technical building blocks. A technical building block is designed by providing both a functional description and a technical specification. The technical specification includes recommended standards which should be used to implement a technical building block. They may provide also some reference implementations to illustrate how to implement a building block and use them in case of need.

The users of this collection can be several:

- **Expert Groups** working on technical building blocks for choosing the recommended standards and reference implementations in the Technical Specifications of the BBs. These Expert Groups have been organized in the scope of WP5.
- **Data Spaces Initiatives** working in designing the Common European Data Spaces funded by the European Commission. They can use the collection to pick up the standards applicable to their domain, indicating which ones are their radar and see which ones have been endorsed by the other data spaces. Use of common standards across data spaces may foster interoperability amongst them. They can also propose additional standards or reference implementations that can be used cross-cutting to all domains.
- **European Commission**, especially the DG-CNECT G1 unit in charge of proposing standardization requests, the High Level Forum of Standardization and the DG-DIGIT in charge of digital systems at the EC. They can consult what the DSSC has identified as potential standards to take them into consideration for their standardization processes; and on the other hand, they can also propose some standards promoted by the EC.
- **Research and Industrial communities** around data spaces in general. The collection allows them to be aware of the existing standards and technologies that can be used as a baseline for data spaces deployment. Thus, they can propose new research activities and standardization proposals to fill the existing gaps, as not all the challenges are solved, and some new contributions are required.

The collection has been persisted in an online spreadsheet initially, available for everyone with the link to the repository, but only editable by the DSSC team. It is also available through the DSSC Support Platform in a Confluence space, accessible by registered users on the platform. In the final stage in autumn 2024 the [collection was made fully public through the DSSC web site](#). The spreadsheet will be then kept only for internal management of the collection, as it is very practical for analysis and tracking, and including also internal metadata for easier management of the entries.

identified the following types: EU Specifications, Formal standard spec, Industry standard body spec, Industry body spec, Industry consortia spec, Product spec.

There are other fields of internal use to the DSSC which are:

- *Identification status*: shows the status of the item with regards to the assessment by T4.3. This task only assesses the correctness of the provided information about the standard and the adequation of the standard to the scope of data spaces. Possible status are: 1. Initial / not checked; 2. Analysis in progress by T4.3; 3. Validated by T4.3; 4. Dismissed by T4.3; Assessment by T4.3.
- *Blueprint status*: Once the standard is included in the collection and validated by T4.3, the different Expert Groups decide which out of them should be included in the DSSC Blueprint. Thus, the possible values here are: 1. Initial / not checked; 2. Assessment in progress by WP5; 3. Dismissed by WP5; 4. Included in DSSC blueprint
- *Endorsement by data spaces*: lists all the data spaces which have considered this standard in their scope and then endorse it. The more endorsed a standard is, the more transversal to all the data spaces, and then more in the scope of DSSC as generic standard.
- *Feedback from data spaces*: the data spaces initiatives have had the chance to provide comments on the proposed standards. This field includes such comments, if exist.

Similarly, the "Reference Implementations" section includes following information per item:

- *Category* (same as above)
- *Building Block* (same as above)
- *Brand*: public or commercial name of the implementation
- *Provider*: implementer organization
- *Implemented standards*: which standards from the "Cross-domain standards" is compliant with the proposed implementation
- *Website*: public site with information about the reference implementation
- *Link to repository*: if open source, URL to access the code repository where the implementation is hosted
- *Open/closed source*: indicates the nature of the implementation, open source or closed source

And other fields internal for the DSSC are:

- Identification status (same as above)
- Assessment by T4.3 (same as above)
- Blueprint status (same as above)
- Endorsement by CSAs (same as above)
- Maturity level: level of maturity of the implementation according to the Solution Maturity Model proposed by OpenDEI and summarized below. So far, none of the collected implementations have been yet assessed with this regard.

	Level	General characteristics of solution	Keywords
1	Ad-hoc	Only reactively offered upon demand.	<i>Initial: chaotic, very poor, initial, basic</i>
2	Low	Slightly structured, guidelines are available, an attributed task for someone in the project, partners have some experience in providing the solution.	<i>Repeatable: Organised, defined, managed, poor, repeatable, accepted</i>
3	Intermediate	Structured and fully documented, there is a "trusted" organisation providing the solution as a business process	<i>Defined: standardised, supported, defined, average</i>
4	High	Adopts and applies best practice, actively support from operational users.	<i>Managed: predictable, measured, mature, developed, systematical</i>
5	Excellent	Accepted as best practice and active contribution to other initiatives such as IDSA and Gaia-X.	<i>Optimized: innovation black belt, synergised, optimising, best practice, sustained</i>

Table 49. Solution Maturity model for BBs

The list includes also Domain-specific standards, which collects all the standards gathered from the data spaces initiatives specific to their domain, that is, only applicable to their domain, and then out of the scope of the cross-domain collection.

The asset description was proposed to the Project Steering Committee of the project in February 2023. It was approved and since then, it has been continuously evolving with the joint effort of partners and external contributors till today. It is a live collection which can be extended and modified according to the needs and evolution of technology in the area of data spaces.

Since the beginning of this activity, and in alignment with the project spirit of transparency and openness, the collection of standards and technologies has been created in collaboration with the overall DSSC consortium first, with the Community of Practice in a second stage (data spaces initiatives) and with the overall Network of Stakeholders in later phase. Thus, we have been incrementally incorporating new players in the collection to come up with the most complete and consensus version of the asset.

Bearing this in mind from the start, the collection methodology defines these steps:

1. An initial version was provided by DSSC partners working in T4.3 with their background knowledge. It is worth mentioning that the partners involved are quite experienced in the data spaces ecosystem. Collaboratively, the T4.3 partners decided on the collection structure and initial content.
2. This version was then opened to the Common European Data Spaces for endorsement and contributions. A copy of the online spreadsheet was shared with the data spaces initiatives to allow them to contribute to the collection with their own inputs, commenting on proposed ones and endorsing those ones that are relevant for their data space.
3. In a second level of openness, the collection was also open to overall DSSC Consortium and Technology Thematic Group for feedback. In this TG, there are representatives from data spaces initiatives, from associated partners and from the Stakeholders Forum. In this level, the contributions have been sent by email, downloading the spreadsheet locally, including the inputs and sending it to the leader of T4.3. As we were working on the online spreadsheet to consolidate the inputs from the previous step, we did not want to overwrite by chance any provided content.
4. Once all external and internal inputs have been consolidated and validated by T4.3, it is time to transfer the collection to the Experts Groups. The EGs are working on describing the different BBs, and as part of the Technical Specifications, they need to refer to this collection for recommending

the most suitable standards and technologies that can be used to implement the BBs. The EGs could also provide new inputs to the collection.

5. After the use of the collection by EGs, and considering the inputs they could provide, new consolidation is required before producing the final report to be included in the D4.1 and making the collection public.
6. Then the asset will follow the same steps as other assets to be published and delivered through the DSSC public web site, as indicated by the T6.1 guidelines.
7. Once delivered, the collection will be continuously contributed and evolved. Thus, the DSSC team will monitor the contributions received through the Support Platform and will address them properly.

The following picture shows the work plan (in 2023) that has been followed including the continuous monitoring and management required until the end of the project.

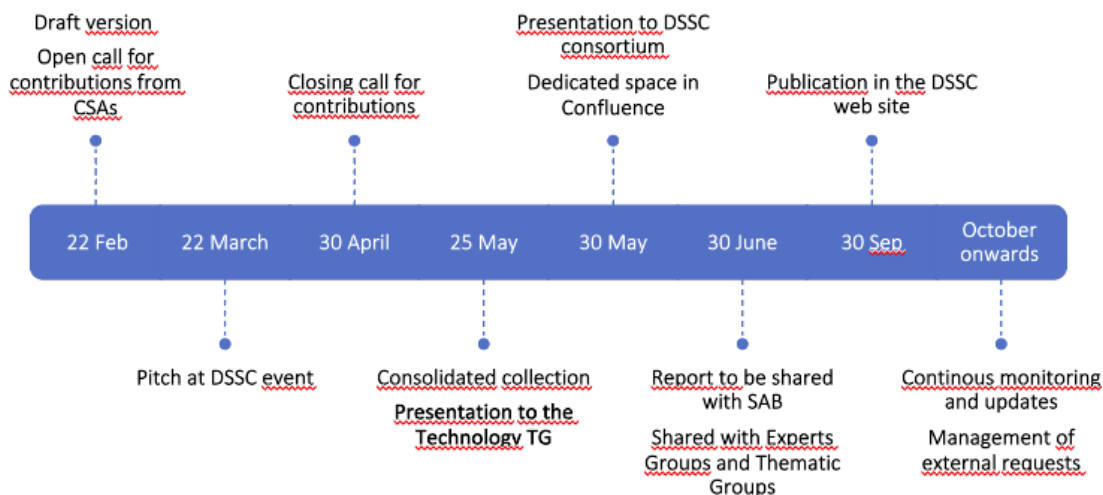


Figure 16. Asset work plan in 2023

This section helps to understand the amount and nature of collected standards. Up to now, 130 candidate standards are included in the collection. All of them have been already validated by T4.3, which, as said, means that they are well documented and fit into data spaces scope.

Looking at the category, most numerous standards are coming for “Data Sovereignty & Trust” (44%) and “Data Interoperability” (38%).

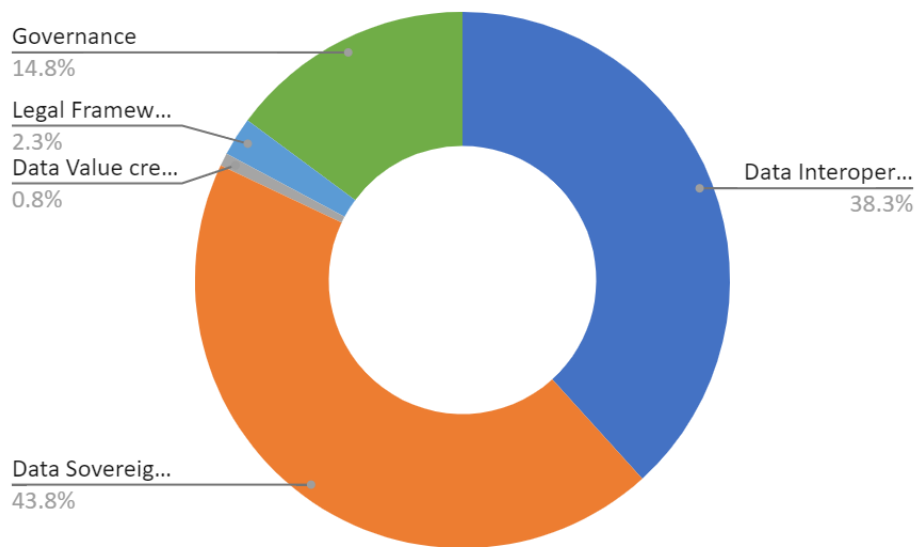


Figure 17: Collected standards per category

Regarding the relation to the Building Blocks, the majority are about “Data Models & Formats” (24%), “Identity Management” (18%), “Trust” (18%) and “Data Exchange” (17%).

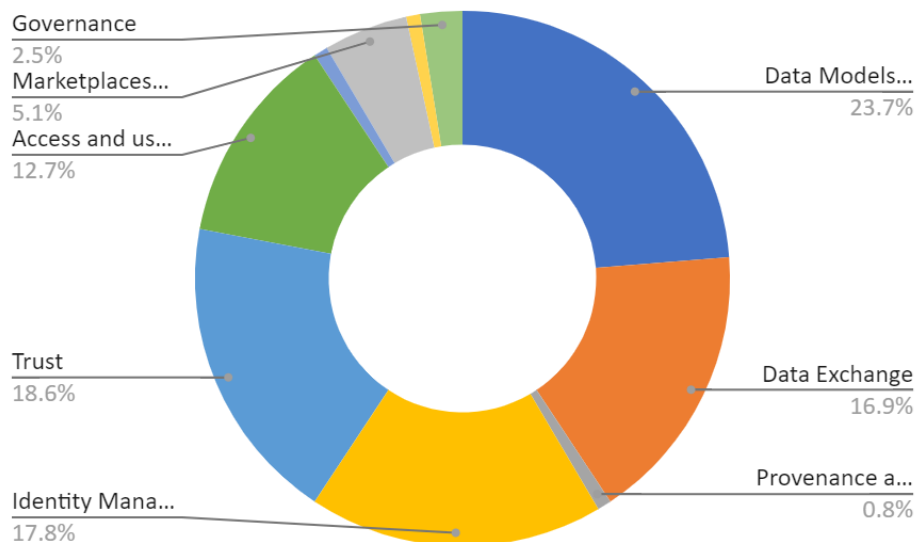


Figure 18: Collected standards per building block

Finally, by the type of bodies and specifications, they most frequent are: “Industry standard body spec” (31%) like W3C, ETSI or TMForum; “Industry consortia spec” (26%) like Catena-X or Smart Data Models; “EU specifications” (20%); and “Industry body spec” (17%) like GAIA-X or IDSA. Only 4% of the collected standards are coming from formal standardization bodies like ISO, ETSI or CEN/CENELEC.

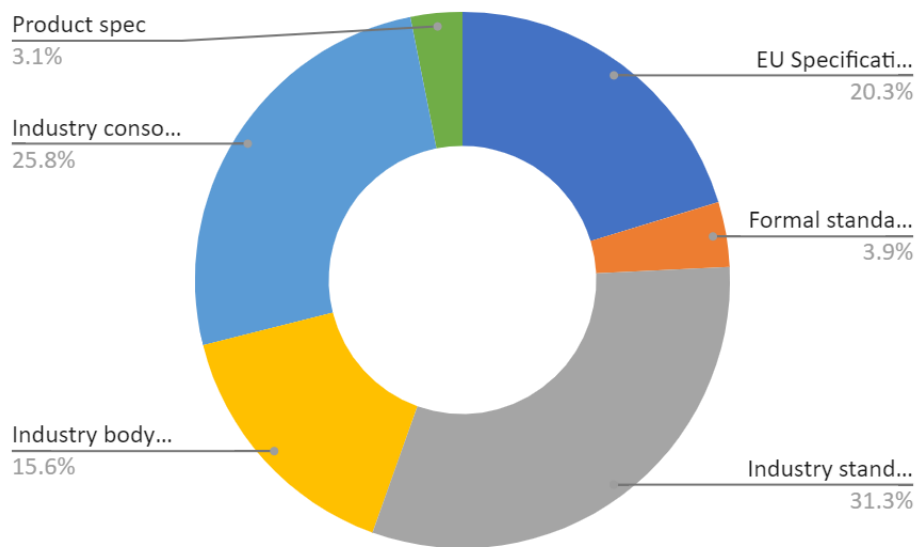


Figure 19: Collected standards by the type of bodies and specifications

During the open phase to data spaces initiatives, nine (9) out of them have provided their feedback, endorsement and inputs to the collection: DS4SSCC (cities), DATES (tourism), INT:NET (energy), DS4.0 (manufacturing), PrepDS4Mobility (mobility), GDI (genomics), LDS (language), DS4SKILLS (skills), AgriDataSpace (agriculture). Additionally, DG DIGIT from EC has also provided a significant contribution.

In total, these data spaces have provided 140 specifications in addition to the ones proposed initially. Approximately 30% of them have been included in the collection, as they fit well in the scope of data spaces and are applicable to all data spaces, independently of the domain. The remaining 70% has been included in an additional section in the collection named "Domain specific". If we observe that in the future there are several data spaces proposing the same standards, we will move them to the Cross-domain section. Otherwise, the recommendation would be to keep them at data space level and federate the DSSC collection with the domain-specific collections of data spaces.

Here below, we can observe the positioning of various data spaces in relation to the number of cross-domain and specific-domain standards they have provided. Some of them, like Tourism or Languages have informed about the use of many cross-domain standards and few specific-domain. While others like Mobility or Genomics have reported more specific-domain and less cross-domain. Skills and Smart

Communities have reported both a number of cross and specific domain.

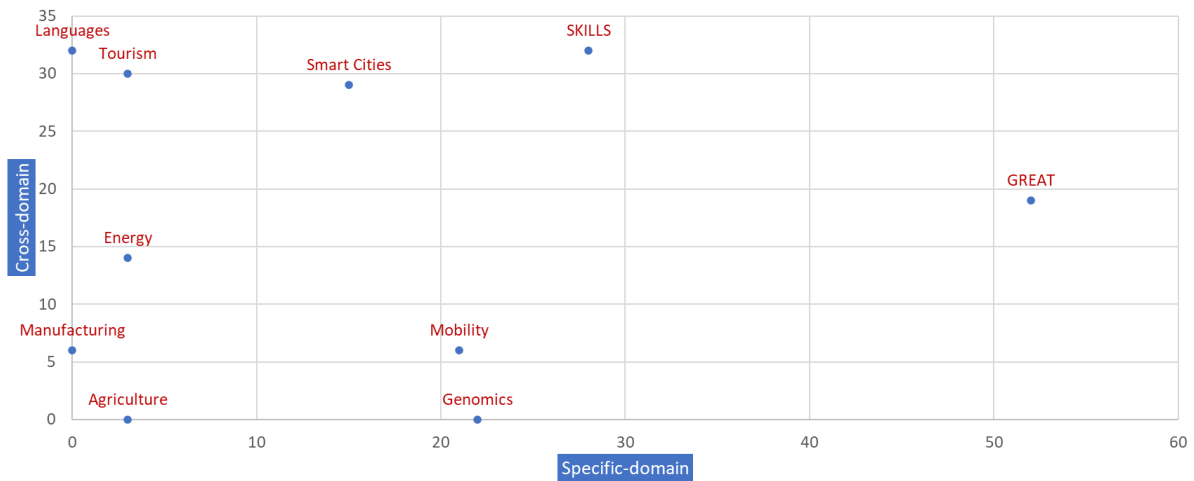


Figure 20: Cross-domain vs specific-domain in data spaces initiatives

The following table shows the most endorsed standards by the data spaces that have been contributed to the collection till now. The mentioned standards have been endorsed by 3 or more data spaces initiative.

Candidate Specification	Related Building Block	Publisher	Data Spaces endorsing
NGSI-LD	Data Exchange	ETSI	DS4SSCC, int:net, DATES, DS4.0
OpenAPI Spec	Data Exchange	OpenAPI initiative	DS4SSCC, LDS, DATES
JSON-LD	Data Models	W3C	DS4SSCC, int:net, LDS, DS4SKILLS
Smart Data Models	Data Models	SDM initiative	DS4SSCC, DATES, DS4.0
XACML	Access and usage policies & control	OASIS	DS4SSCC, LDS, DATES
eIDAS/eIDAS2	Identity Management	EU	DS4SSCC, int:net, LDS, DS4SKILLS
ODRL	Access and usage policies & control	W3C	DS4SSCC, LDS, DATES
EUDI RAF	Identity Management	EU	DS4SSCC, LDS, DS4.0
Oauth2	Identity Management	IETF	DS4SSCC, int:net, LDS, DATES, DS4SKILLS
OpenID	Identity Management	OpenID Foundation	DS4SSCC, LDS, DATES
ICAM	Identity Management	GAIA-X	int:net, LDS, DATES
SAML 2.0	Identity Management Access and usage policies & control	IETF	DS4SSCC, LDS, DATES, DS4SKILLS

DCAT-AP	Publication & Discovery	W3C	DS4SSCC, int:net, LDS, DATES
DID	Identity Management	W3C	LDS, DATES, DS4SKILLS
SOLID	Identity Management	MIT	DS4SSCC, LDS, DS4SKILLS
Verifiable Credentials Data Model	Identity Management	W3C	LDS, DATES, DS4SKILLS
	Trust		

Table 50. Most endorsed standards by data spaces initiatives contributing to the collection

We can observe that the Foundational Standards recommended in the DSSC blueprint (DCAT, ODRL and Verifiable Credentials) are included in the list of most endorsed ones.

2.3.4 Liaison activities with Standardization Development Organization (SDOs)

Once concluded the activity of collecting all the standards from data spaces and depicted the standardization plan in alignment with the EC, we acquired a good knowledge of the role that DSSC may play in the overall standardization ecosystem. Thus, we decided to analyse the most related SDOs with data space and start an engagement process to contribute closely as liaison project.

We found several standardization committees and initiatives related to data spaces. These organizations work on developing standards and frameworks for data sharing, interoperability, and management. Here are a few notable ones:

Committee	Related topic	Involved DSSC partners	Priority for DSSC purposes
ISO/IEC JTC 1/SC 27	Data protection		Medium
ISO/IEC JTC 1/SC 38 (WG 5)	Data in cloud computing, distributed platforms, and related technologies.	Fraunhofer, IDSA	High
ISO/IEC JTC 1/SC 38 (WG5) ISO/IEC TS 10866	Organizational autonomy and digital sovereignty	IDSA	Medium
ISO/IEC JTC 1/SC 38 (WG 5) PWI 20151	Data spaces	IDSA	High
ISO/IEC JTC 1/SC 32	Data Management related to data spaces - metadata		Medium
ISO/IEC JTC 1/SC 42	AI standardization and Big Data	BDVA	Low

W3C - Dataset Exchange WG	Update DCAT and application profiles	Insight	High
W3C - Decentralized Identifier WG	DID schema, syntax, specifications	Insight	Medium
W3C – RDF WG	Unique and deterministic hash of RDF Datasets.	Insight	Low
W3C - Verifiable Credentials WG	Verifiable Credentials Data Model specification	Insight	Medium
W3C- ODRL CG	Permissions, Prohibitions and Duties expressed by the Policy	MyData	Medium
IEEE - P3800 Data Trading System workshop	Standards for data Trading, specification of connectors and data usage rights (IDS-RAM)	IDSA	Medium
CEN-CENELEC Trusted Data Transaction workshop	Trust in data transactions	Fraunhofer, TNO, BDVA, IDSA, GAIA-X, FIWARE	High
ETSI CIM ISG	Context Information Management – NGSI-LD data Exchange API	FIWARE	Medium
ISO TC/211	Geographic information/Geomatics		Low
CEN-CENELEC JTC 21	Artificial Intelligence	BDVA	Medium
CEN-CENELEC Focus Group on Data Spaces (embedded in JTC25)	Data Spaces, cloud and edge	IDSA, BDVA, FIWARE, Fraunhofer	High
CEN-CENELEC JTC25	Data Spaces, cloud and edge	IDSA	High

Table 51: Relevant standardisation committees

A more detailed description of each of these committees can be found in Appendix 2.

A thorough analysis of all these committees and working groups was carried out by the task force team, and we selected those ones that were considered of the most relevance for data spaces and where the DSSC contribution could be more valuable. The table below shows the reasons why they were selected, and which are the expected contributions from DSSC.

Selected SDO	Relevance for data spaces	DSSC contribution
W3C - Dataset Exchange WG	<ul style="list-style-type: none"> • Maintain and revise the DCAT, taking into account feature requests from the DCAT user community. • Define and publish guidance on the specification and use of application profiles when requesting and serving data on the Web 	<ul style="list-style-type: none"> • Extensions of DCAT when needed and specifically for services descriptions, mainly from Data Value Creation Enablers Expert Group
CEN-CENELEC Trusted Data Transaction workshop	<ul style="list-style-type: none"> • Reinforce trust between all parties involved in the data transactions and by that contribute to the acceleration of the data ecosystems. • Provide a better understand of the concept and implications of Trusted Data Transaction and to anticipate future related standardisation requirements 	<ul style="list-style-type: none"> • Ensure Alignment in a framework for trusted data transactions • Common perspective on: <ul style="list-style-type: none"> ○ Glossary ○ Data Product ○ Use cases (stages)
CEN-CENELEC Focus Group on Data, Dataspaces, Cloud and Edge (moving to JTC25 in Jan 2025)	<ul style="list-style-type: none"> • Identify the standardization gaps in the data spaces, cloud and edge standards • Connect various stakeholders working on the topic 	<ul style="list-style-type: none"> • Feedback on proposed standardization requests and plans • Engage with wider range on stakeholders on the topic

Table 52: DSSC contribution in certain standardisation bodies

The DSSC project management board (PMB) voted and approved either the selected SDOs and the proposed liaison persons to engage on behalf of the DSSC with each of the SDO. A **liaison person** is one participant of the DSSC project who establish the link with the SDO and coordinates the relationship in both directions: by providing inputs on behalf of the project to the SDO; and reporting to the DSSC partners about what is ongoing in the SDO committee. To avoid any conflict of interest, this person cannot represent their own organization in the same committee. The figure below represents graphically the interaction with the SDO through the liaison person.

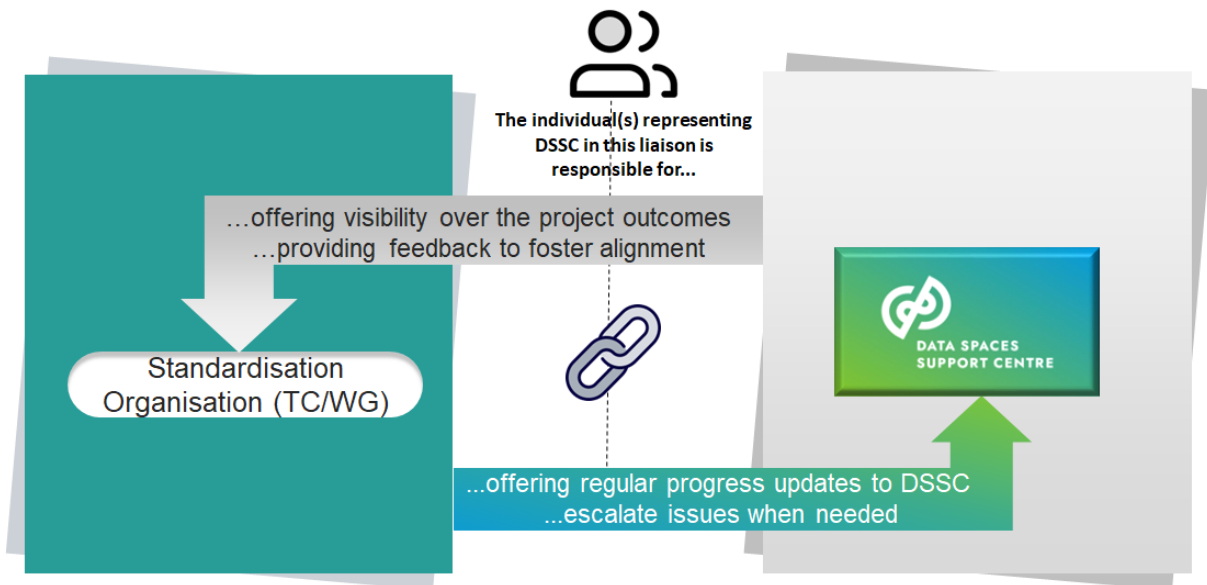


Figure 21: Bidirectional information flow supported by the project liaison

The work around liaisons with SDOs is framed in a joint task force between T4.3 and T2.4, which is in charge of the liaison of DSSC with external organizations. This task force is monitoring all the liaison actions and keeping the potential SDOs updated through an internal project dashboard. For every new SDO selected, we launch an internal call for candidates to act as liaison person. If there are no objections from the partners about the nominated person, the liaison is established. In some cases, the liaison implies subscription to a fee which is paid by the partner acting as liaison (as general rule). The assigned liaison persons must report periodically to the DSSC partners about the SDO activities and raise any issue that may require an intervention or decision by the DSSC partners.

The current outcomes out of the liaison with above mentioned SDOs is summarized below:

- W3C Data Exchange. The liaison process is being quite complicated as the liaison person proposed as individual expert is not accepted since his affiliation is already a W3C member (University of Galway), and the DSSC cannot become a member of W3C since it does not have legal entity. In the meantime, we overcome this issue, the contributions to DCAT standard can be channelled through the corresponding working group in JTC25.
- CEN/CENELEC Trusted Data Transactions. The main purpose of this liaison is to align the contributions of the various DSSC partners participating individually in the workshop, to provide at least a non-contradictory input with the DSSC vision. Thus, the main work here has been to consolidate a common view across the involved partners beforehand every meeting.
- CEN/CENELEC Focus group on data spaces. For one year, DSSC has participated in the group meetings represented by the nominated liaison person. Through this group, DSSC has contributed to the standardization request proposed by the EC and to the Rolling Plan for ICT standardization. The work of this group has been taken over by the JTC25.
- CEN/CENELEC JTC25. The liaison request has been accepted and liaison person nominated. She will attend the next working groups meetings.

2.3.5 Takeaways for data spaces blueprint

The “Data Spaces Standards and Technologies landscape” has been a perfect occasion for collecting in one single place all the standards and reference implementations in the scope of data spaces so far. Thanks to the open and collaborative process put in place, the collection may provide a common source to find the consensus about the most relevant standards for the interoperability of data spaces. The collection can be an inspiration and reference for the experts at the Experts Groups in defining the specifications of the Building Blocks at the blueprint.

Through the proposed standardization plan, the DSSC may identify the gaps in the current standards and propose new standardization requests to be initiated. The liaison persons help in understanding what is under standardization and which standardization requests are in development. This information is again useful for the experts of the blueprint, to envisage standards that can be proposed in upcoming versions of the blueprint.

All this impacts on the data spaces initiatives which get access to a blueprint which is aligned with the relevant SDOs for data spaces.

2.3.6 Upcoming work

If extension of T4.3 is approved under the current amendment, the collection of standards will be continuously updated by attending to all requests coming through the DSSC Support Platform. Also, other SDOs could be identified and liaised by the DSSC, and the already liaised will be continuously attended and reported, both to the DSSC partners and to the EDIB, when required.

2.3.7 Results

Task	Document	External reference	Document internal reference
T4.3	Data Space Standards and Technologies Landscape	Report	Chapter 2.3
T4.3	Dashboard for all potential SDOs		Summary included in chapter 2.3.4 and full description in Appendix 3.
T4.3	List of analysed SDOs		Appendix 2
T4.3	Standardization plan	A separate document delivered to EC by	-

		request, not in the original project plan	
T4.3	Liaison person roles		Appendix 1, monitoring dashboard

Table 53: Summary of results of standardisation

2.4 Architectures landscape

2.4.1 Scope and purpose

Once the analysis of the standards landscape was completed, and following the contractual description of T4.3, we started the activity to analyse the architectures proposed by the various data spaces initiatives. The objective of this work is, on one side, create a landscape of most adopted architectural approaches, but also to identify the commonalities across the various proposals. The final expected outcome is the creation of archetypes or patterns that could be applicable under certain domains and/or requirements.

In order to carry out such work, the process was as follows:

- Collect all the available architectural documents from the data spaces initiatives, that is from the Common European Data Spaces, especially those ones in the deployment phase, since they require to define an architecture for actual deployment. Many of these documents are not public, so we cannot disclose the source, and all is consolidated in the analysis.
- Desk research analysis of all the collected documentation. Of course, the focus of the analysis is on technical information, since an IT architecture is the design of the arrangements and interactions of a system components to meet certain requirements.
- Report the findings under a common dashboard with the following information:
 - Overall description of the architecture approach
 - Deployment status of the proposed architecture. It can be just a concept, or implemented but not yet in production, or already in an operational environment.
 - Which data space services are proposed in the architecture (federation, participant agent, value creation)
 - The components proposed by DSSC which are used in the architecture (in some cases they use a different name but the functionality that represent is the same than in DSSC proposal).
 - The main used technologies, that is, details about the chosen implementation technology.
 - The main adopted standards and followed technical specifications.
 - The deployment strategy which specifies where, from whom and how to get the required components and deploy them in place.
 - And, if available, the graphical image which illustrates the architecture.
- With all this information in place, next step was to analyse the cross-cutting approaches, technologies and components.

- For the identification of archetypes and patterns to document all for the sake of the current data spaces themselves and for guidance of future data spaces.

2.4.2 Architectures analysis

At the time of writing this report, we have collected architecture document from 9 data spaces, from which, 5 out of them are in the deployment phase. Thus, we decided to start the analysis by the data spaces in this phase since they have probably a more mature architecture and leave for a second stage of analysis those ones at earlier stages. The 5 data spaces analysed so far have been: Smart Communities, Tourism, Media, Languages and Manufacturing.

The main insights obtained in the analysis can be summarized below:

- All the data spaces architectures are in a conceptual deployment status
- The approaches we have observed all have in common are:
 - They follow a bottom-up approach by leveraging on existing legacy systems in their ecosystem.
 - They all rely on data space connectors facility to manage the onboarding and data transactions within the data space.
 - All are implementing decentralized identities systems and federated models for data sharing.
- Three of them already selected components that they plan to implement or reuse. The ones in common are:
 - Central/federated catalogue (metadata descriptions)
 - Credentials registry
 - Common vocabulary (called vocabulary hub or semantic hub)
 - Clearing house (only in two)
- The DSSC Building Blocks are quite followed in general, with wider adoption of the technical ones than of the business and organizational ones.
- Regarding the technologies in common, the unanimously mentioned was Verifiable Credentials, in some cases concretely talking about GAIA-X Trust Framework. In some cases, they were also referring to Data Spaces Connectors (either Eclipse Data Connector or FIWARE Data Space Connector).
- In relation to standards (provided by three of them), the commonly referred were: W3C DID, W3C VC, W3C DCAT, ODRL.
- In some cases, the deployment strategies are not fully clear, but some of them referred to the use of testing facilities for assessment and benchmarking; and use of SIMPL component when relevant and mature.

The following figure shows the common pattern and workflow that we have observed from the analysed architectures. First, the data space establishes a governance to set up the rules and roles; then, a central registry of credentials, vocabulary hub and catalogue are deployed in the data space; once all deployed, data spaces connectors (participant agent as it is called in DSSC) help the participants to onboard; and finally, the peer-to-peer data transactions may happen to actually share/exchange data and services.

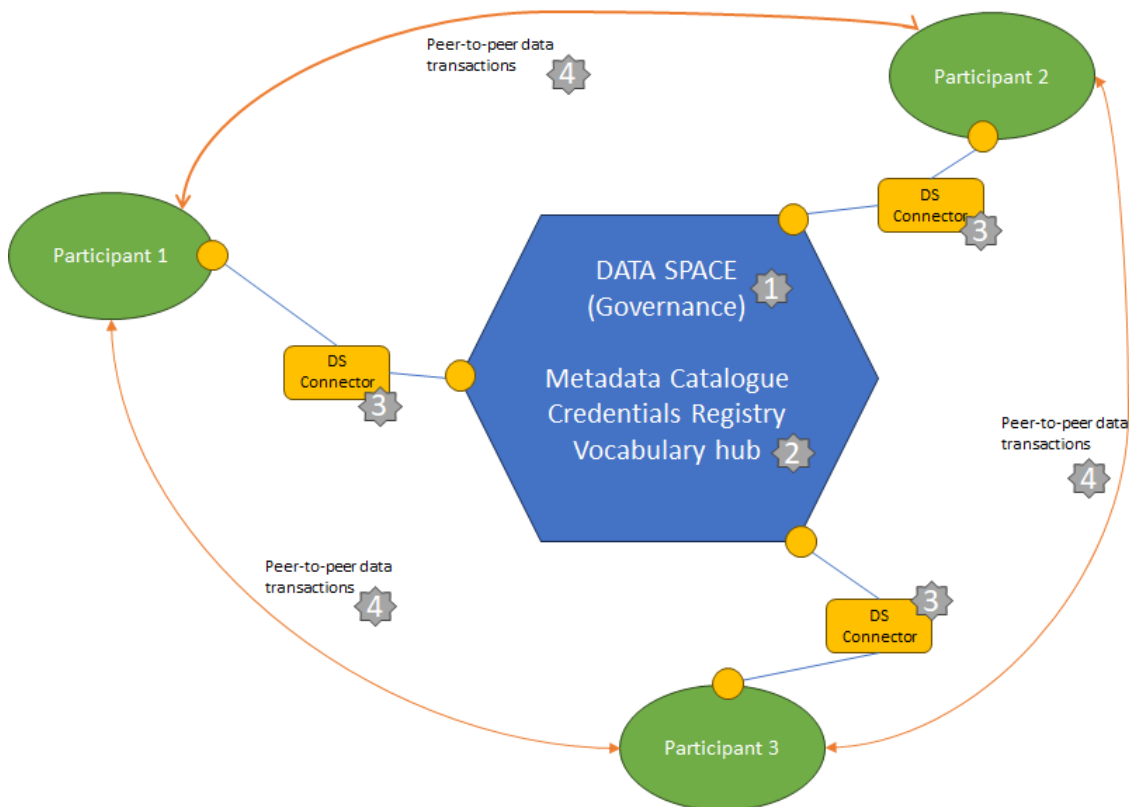


Figure 22: Depicted common pattern from analysed data spaces architectures

2.4.3 Takeaways for data spaces blueprint

The current work was presented to the community in the scope of the Technology Thematic Group in February 2024. After presentation, we offered an interactive board for discussion and feedback, and here below the inquired questions and the provided input by the community:

- Do you see realistic, useful and feasible the proposed objective and taken approach?

In general, the community consider useful this work, although it is perceived as difficult to get valuable outcomes if we do not go into right level of detail, if there is not agreement in the scope and if there are multiple bodies discussing rules, agreements and standards. In order to overcome all these barriers, they recommend setting up a strong governance framework.

- Are the proposed criteria sufficient for the analysis or would you suggest any additional criteria to analyse the collected architectures?

The community would add additional criteria such as: test for availability of minimal standard data exchange, consideration of user friendliness aspects, degree of adherence to standards, followed governance schema.

- Do you see promising the obtained results so far in order to end up with a set of archetypes? Do you observe interesting commonalities that can favour such endeavour?

There were no comments about the presented outcomes.

- Please propose alternatives, other data spaces architecture to look at, additional steps

The community suggested to look also at the economical impact of the technology choices, that is evaluating “money making” aspects in implementations. We are not sure if the business plan must be part of this pure technical analysis of the architectures, but for sure it is impacted. Full building block is dedicated to business models, and the business model must be aligned with the technical choices represented in the architecture. Other suggestions were the use of GenAI for discovering the architectural patterns and the evaluation of time and difficulty of components integration.

We offered the community the possibility to be contacted to provide their data space information or provide further comments on the analysis or additional sources we may consider.

All this material has been shared with the experts and members of the Architecture Board for exploring the implications in the future versions of the Blueprint. In some cases, it will also generate new requirements for the building blocks (to be recorded under T4.1).

2.4.4 Upcoming work

In the following months till the end of the project, we expect to continue exploring other architectures that will arise from the existing or upcoming data spaces. A new presentation of the updated material is planned for the future, both internally in the DSSC consortium and externally to the community.

The evolution of the data spaces towards the deployment and operations phases will bring us the validation of the proposed concepts and components. In the light of these results, this work can be revisited and extended with additional inputs to produce more accurate and detailed architectural patterns. If we are able to discover what the data spaces are commonly using and deploying, we may support them for addressing the optimal reusability and interoperability across them.

2.4.5 Results

Task	Document	External reference	Document internal reference
T4.3	Architecture landscape		Chapter 2.4

Table 54: Summary of results of architecture landscape

2.5 Requirements of demonstrators and test infrastructures

2.5.1 Introduction

As part of work package 4 (WP4), task 4.1 (T4.1) had, *inter alia*, the objective to identify **test infrastructures** that have been implemented as well as **core software components** developed by data space initiatives, with a particular focus on the deployment actions funded by the Digital Europe Programme (DEP).

This section 2.5 provides a concise overview of the methodology employed and the results achieved at the time of publication of this deliverable D4.1.

2.5.2 Methodology

The methodology followed the following structured process.

Step	Name	Explanation
01	Definition Exploration	<ul style="list-style-type: none"> exploration and research of existing generic definitions and definition attempts for “demonstrators” and “test infrastructure” provide an overview of existing definitions
02	Initial Definition	<ul style="list-style-type: none"> scan and evaluate existing definitions from step 01 distil common parts relevant for the application of the identified definitions for the purposes of data spaces generate an initial definition for both terms
03	Definition Review 1	<ul style="list-style-type: none"> publish and disseminate the initial definition collect feedback (critique, improvement suggestions, other remarks and issues) from within the WP4 team
04	Definition Refinement 1	<ul style="list-style-type: none"> evaluate and prioritize feedback received in step 03 refine the initial definitions of the two terms based on the prioritized list of refinements create a v2 of the two definitions
05	Definition Review 2	<ul style="list-style-type: none"> publish and disseminate the definition v2 to the Technical Thematic Group (TTG) collect feedback (critique, improvement suggestions, other remarks and issues) from the TTG team during an interactive workshop session
06	Definition Refinement 2	<ul style="list-style-type: none"> evaluate and prioritize feedback received in step 05 refine the initial definitions of the two terms based on the prioritized list of refinements create a final version (v3) of the two definitions publish the final version (v3) of the two definitions for the whole DSSC project
07	Survey	<ul style="list-style-type: none"> identity the most effective way to survey existing data space initiatives regarding their usage of data space demonstrators and data space test infrastructure conduct the survey (e.g., interactive, questionnaires) record survey results
08	Assessment	<ul style="list-style-type: none"> assess results found with regards to inconsistencies or missing pieces of information refine and consolidate the findings as far as required (e.g., by asking survey respondents for clarifications)
09	Review & Finalization	<ul style="list-style-type: none"> compile final results document final results

Table 55. Methodology for identifying test infrastructures and demonstrators

2.5.2.1 Methodology for step 05 (Definition Review 2) by the TTG (Technical Thematic Group)

Besides introducing and discussing version v2 of the definitions of the two concepts, the project employed the following interactive opinion and response elicitation approach with the TTG experts to establish that the definitions focus on the right target area for the purposes of the project. The experts were confronted with the following two questions.

Q1: When looking at the definition of [demonstrator | test infrastructure]:
What things you believe to be a valid demonstrator **will we miss** when using this definition?

Q2: When looking at the definition of [demonstrator | test infrastructure]:
What things will we get as response which you would **not qualify** as valid [demonstrator | test infrastructure] for the blueprint?

The respective responses were duly collected and directly worked into the final definitions given below.

2.5.3 Data Space Demonstrator

2.5.3.1 Definition

The following definition has been adopted after step 06, Definition Refinement 2.

Definition. A **demonstrator of a data space** is a set of software components able of showcasing capabilities of a data space relevant for a specific set of concerns of certain stakeholders in a non-productive form and usage environment.

Notes

1. The key ingredient of this definition is that demonstrators, by design, are **limited** to show things relevant only for **some concerns** of **some stakeholders**. These concerns typically address certain capabilities (in our DSSC sense) and may regard functions (e.g., identification and authentication of participants) or processes (e.g., how two participant agents negotiate a mutually accepted contract) and may also include non-functional requirements such as performance. Limits of a demonstrator may be tighter, resulting in a smaller demonstrator, or looser, resulting in a larger demonstrator.
2. **Stakeholders** are the key aspect for any type of demonstrator. Identifying stakeholders and their concerns is the guide for answering the typical questions
 - Who would be the target audience of the demonstrator?
 - What would be the purpose of the demonstrator (and hence its features)?
 - Who is expected to provide a demonstrator
3. We are mostly interested in demonstrators that are set up and run in **non-productive** environments - maybe even set up just for the purposes of the demonstration. "Non-productive" means that the demonstrator does not use actual, real, or "live" business or operational data or other pieces of information of participants. We understand that one can always use the operational (productive) services of a data space (maybe just in need of setting up test accounts for

participants) but this is not in the fore of the concept as it requires the full operationalization of a data space as opposed to the more limited understanding of a demonstrator here.

4. A **click dummy** may be regarded as a *demonstrator for business people* to the extent that features and processes relevant to business users (= stakeholders) are demonstrated.
5. A docker container consisting of two data space connectors with a single access token server is a **demonstrator for technical people** even though other elements of a full data space, such as the data space registry or the inclusion of proper trust anchors are missing in this particular demonstrator.
6. Even a **single data space connector** may be regarded as a data space demonstrator for some stakeholders and their particular concerns, for instance, for IT infrastructure people (= stakeholders) trying to understand the subtleties of how to install and run (= concern) this instance of a participant agent.
7. The demonstration of certain **use cases**, including end-to-end scenarios, is also a valid concern and potential subject of a demonstrator.
8. Depending on the breadth and capabilities, a demonstrator may even become or be used as a test infrastructure. Additionally, one may legitimately use a certain test infrastructure to “demonstrate” some aspects of a data space. This indicates that the technical differences between demonstrators and test infrastructures are not as clear cut as it might seem and that one may be used in the context of the other.

For the purposes of the survey, we suggest taking a stakeholder and concern-driven approach: If some stakeholder in their own view use a particular software for demonstration purposes, we qualify it as demonstrator. This, *mutatis mutandis*, shall apply for test infrastructure as well.

2.5.3.2 Miscellaneous other definitions

During step 01, research, the following other definitions of (generic) demonstrators have been identified which readers may find useful on other occasions:

A demonstrator is ...

- a piece of merchandise which can be tested by potential buyers
- a product (such as an automobile) used to demonstrate performance or merits to prospective buyers
- an item for sale that customers can use and test so they can decide whether they want to buy it.
- a technology demonstration (or tech demo), also known as demonstrator model, that is is a prototype, rough example or otherwise incomplete version of a conceivable product or future system, put together as proof of concept with the primary purpose of showcasing the possible applications, feasibility, performance and method of an idea for a new technology. They can be used as demonstrations to the investors, partners, journalists or even to potential customers in order to convince them of the viability of the chosen approach, or to test them on ordinary users.

2.5.4 Data Space Test Infrastructure

2.5.4.1 Definition

The following definition has been adopted after step 06, Definition Refinement 2).

Definition. A **data space test infrastructure** consists of the

- hardware
- software (including tools)
- other resources (e.g., network, test equipment, test data)
- processes covering the whole test lifecycle, and
- methodologies

used to systematically assess and validate a certain set of functionalities and quality attributes of a data space implementation.

Notes

1. Definitions of “test infrastructure” frequently refer to the following terms and concepts which are included in the above definition as explained below:
 - testing frameworks are subsumed under processes as well as frameworks
 - test environments are certain collections of hardware, software, and tools
 - test management is one (important) process of the test lifecycle processes
 - test automation is a combination of test hardware, software, and processes
2. A single test infrastructure does not need to test all functions or all non-functional requirements of a data space but may concentrate on suitable subsets thereof. Typically, functional testing and performance (e.g., load, stress) testing is often performed by different testing infrastructures.
3. Because a data space typically needs several different and distributed software components (e.g., participant agent services, catalogue service, logging service), the testing infrastructure will become more involved and complex the more functionalities or quality attributes one wants to test. Depending on a test infrastructure’s capabilities, one may also test larger use cases or whole processes.
4. The reference to a “data space implementation” distinguishes this IT-related notion of testing from other methods to assess and evaluate a data space such as maturity models.
5. While software engineering practices and processes (e.g., Continuous Integration / Continuous Delivery, CI/CD) discriminate between a testing phase and any demonstration of the capabilities of the software under test, it is conceivable that the actual testing software may also be used to demonstrate certain features and functions of the software under test. This means that, in some situations, a data space demonstrator may also be used as a test infrastructure. For our purposes, we would like to regard all software components which are used in such a testing phase as part of the test infrastructure - including, if appropriate for a particular data space, any demonstrators.
6. The above definition is very much focused on the technical infrastructure of a data space. However, data space implementors will also need to test non-technical aspects of their endeavour such as the governance model, how the various processes (e.g., on-boarding, off-boarding, registering and

de-registering data assets) work in practice, how interoperability is achieved between the different data space participants (e.g., testing data types and semantics), and various additional areas.

2.5.4.2 Miscellaneous other definitions

During step 01, research, the following other definitions of (generic) test infrastructure have been identified which readers may find useful on other occasions:

Testing Infrastructure

- A testing infrastructure is the essential framework and tools used in software development to systematically assess and validate a software system's quality, functionality, and performance. It encompasses testing environments, automation frameworks, and methodologies, ensuring thorough testing and defect detection.
- Test infrastructure refers to the hardware, software, and other resources necessary to support software application testing. This can include things like test environments, test tools and equipment, test data, and other resources that are needed to plan, design, execute and evaluate tests.
- It is an integral part of the testing process by providing the resources and support required for testing activities. It is typically designed and set up before testing, and it is commonly managed by a dedicated team or individual.
- It covers testing exercises and processes to ensure the fastest test execution, providing a shorter release cycle and faster time to market.
- Test infrastructure refers to the environment, tools, and resources used to test software. It includes all the components and systems required to execute tests, such as test management tools, test automation frameworks, testing environments, and other supporting tools.
- Test infrastructure is about the relevant environments, tools, resources, and even any procedure that can assist in enabling software testing. This includes hardware, software, test automation frameworks, network configurations, and any process that can make software testing more efficient and effective.
- Test infrastructure consists of the hardware, software, and resources required for software application testing. This includes test environments, tools, equipment, data, and other necessary resources to plan, design, execute, and evaluate tests.
- Test infrastructure refers to the collection of tools, resources, and frameworks that facilitate the planning, creation, execution, and analysis of software tests. It encompasses both hardware and software components, forming a cohesive environment that supports the testing process throughout the software development lifecycle. Test infrastructure empowers developers to identify and rectify issues early, resulting in more stable, secure, and performant software products.

Test infrastructure also includes performance testing tools, security testing frameworks, and monitoring systems that provide developers with comprehensive insights into the software's behavior under different conditions. These components help identify bottlenecks, vulnerabilities, and potential areas for improvement, enabling developers to optimize software performance and enhance its security posture.

In summary, test infrastructure serves as the backbone of software testing in modern development practices. It encompasses the test environment, automation frameworks, CI/CD pipelines, and various testing tools. By investing in a robust test infrastructure, organizations can ensure the quality, reliability, and user satisfaction of their software products.

Test environment

A fundamental element of test infrastructure is the test environment, which provides a controlled setting to evaluate software behavior under various conditions. This environment mimics the real-world scenarios in which the software will operate, allowing developers to validate its functionality, performance, and compatibility across different platforms and configurations. It encompasses hardware, operating systems, databases, network configurations, and other dependencies necessary to reproduce the software's operational environment accurately.

Test automation

Test automation frameworks and tools are integral parts of test infrastructure. These tools enable developers to automate the execution of tests, minimizing manual effort and ensuring consistent and repeatable test results. Popular test automation frameworks like Selenium, JUnit, and PyTest provide developers with a wide range of capabilities for creating and managing test cases, executing them across different environments, and generating insightful test reports.

2.5.5 Upcoming work

At the time of production of this deliverable D4.1, the project had completed steps 01-06 and was in the process of scheduling and initiating step 07. That implies that the following methodology steps (cf. section 2.5.2) need to be carried out during the remainder of the project:

- Step 07 - Survey
- Step 08 - Assessment
- Step 09 - Review & Finalization
- The requirements for demonstrators and test infrastructures are described in this document. Next step in this field is to create a collection of the relevant demonstrators and test infrastructures.

2.5.6 Results

Task	Document	External reference	Document internal reference
T4.1	Requirements for Demonstrators and Test infrastructures		Chapter 2.5

Table 56: Summary of results of demonstrators and test infrastructures

3 Conclusions

This deliverable D4.1 presented the results of candidate building blocks identification and analysis, more distinctively 1) collection of requirements, 2) emerging building block overview and contributions to Blueprint in its various phases 3) standards landscape, 4) architecture landscape, 5) demonstrators and test infrastructures definitions. Figure 2, presented earlier in subchapter 1.2, depicts the DSSC internal collaboration process, and at which phase the DSSC offering these results contribute and how. Within DSSC, the results presented in this deliverable have been synchronized with the Blueprint 2.0 release (Figure 1) to ensure contribution to Blueprint. In some of the chapters also further work is reported, thus serving the continuum of especially the tasks T4.1 and T4.3 to fulfil the finalization of the work as planned in the proposal.

Data space initiatives and the whole genre of data spaces have evolved remarkably since the planning of DSSC project. WP4 tasks reflect on what are the needs in the community in different stages of progress. Thus, support for Common European Data Spaces initiatives is rather a continuous process than a series of exactly planned steps.

As reported in chapter 2.3.6, in February 2025 the timeline of the task 4.3 (standardization) was extended from end of March 2025 until the end of the project. In this way the task continues the work with the standardization liaisons and standards landscape monitoring. In addition, the architecture analysis is planned to continue since the deployment data space initiatives are running and defining their architectural choices. Material updates are planned when feasible.

Task 4.1 continues still after the submission of this deliverable as planned in project timeline and any further results will be reported in D4.3

There were also topics identified and closely linked to the work presented in this deliverable. Here is an exemplary list of backlog for future work and reporting.

- Need of a personal data intermediary to adequately handle personal data beyond the GDPR requirements
- Legal mapping and analysis, This will involve further work on a number of specific topics (e.g., personal data, essential requirements, data intermediation service providers, competition law and cybersecurity). (D4.3)
- Legal barriers (D4.3)
- Collection of demonstrators and test infrastructures (Task 4.1, D4.3)

Contributions will continue in relevant parts of Blueprint development.

Appendices

1. Appendix 1: Liaison person roles

DSSC Liaison in Standardisation Bodies

Rationale & Motivation

Collaborative projects have the opportunity to directly engage with Standardisation Organisations (SDOs) through establishing liaisons with Technical Committees (TCs) or Working Groups (WGs). The project, acting as a liaison initiative, can assign individuals to represent it and act typically as observers, without voting rights but with the possibility to present and be and be informed about the ongoing standardisation activities taking place. This kind of participation can come without any cost for participation from the project.

The motivation for establishing this connection is to enhance the project's visibility from one hand, but also its capacity to influence bringing to the discussion a strong common voice of the consortium members.

Thereby, the DSSC project liaison is an individual that undertakes a multifaceted role which is crucial in ensuring effective coordination between the Data Spaces Support Centre and the relevant SDO. This individual acts as a bridge or a hub between DSSC and the SDO, bringing the "hat" of the DSSC and aiming to support the strategic alignment of the project goals with industry standards (analogous to the Relationship Manager for CoP)

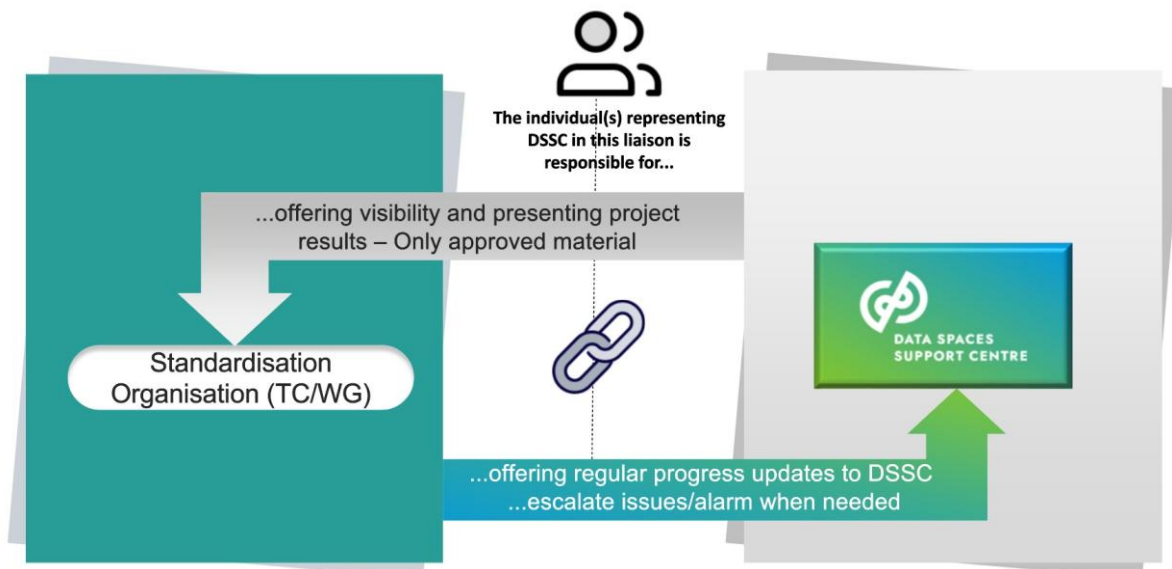
Role & Responsibilities of Liaison

Responsibilities of the project liaison include:

- Following up the activity of the SDO TCs/WGs (participating in meetings, documents, consultations, etc.) bringing the DSSC hat;
- Offering visibility to the project and its activities, mission and objectives;
- Presenting the project results when asked/needed, ensuring internal DSSC alignment at all times. Only approved/published materials can be presented by the liaison (not working materials that haven't reached consensus among DSSC consortium partners);
- Staying updated on the relevant activities of the SDO
- Report back and offer updates to DSSC consortium with regards to the SDO WG/TC activities (general meetings/townhall meetings could be used) and make materials available for DSSC (teams folders)
- Notify/alarm when necessary for any issues encountered (help in escalating)

The effort needed from an individual to undertake and sufficiently fulfil the associated responsibilities of this role is estimated: from 1 working day/month (5% of monthly working time) to 4 working days/month (20% of monthly working time).

This bidirectional information flow supported by the project liaison is graphically depicted below:



Governance of the Liaison

For better supporting the connection of DSSC with SDOs (including defining the process and plans, landscaping, identification of new opportunities to influence, etc.) and an internal task force has been created comprised of project representatives from 2 project tasks and namely:

- T2.4 (as per DoW establishes liaison with Standardisation bodies & organisations)
- T4.3 for common standards and open-source technologies that has been working on fostering a standardisation plan for DSSC

The following points are relevant for the governance of the liaison:

- Liaisons to be decided at plenary level (which bodies need to be prioritised and why). Internal task force is responsible for running this prioritisation exercises and internally exploring and nominating individuals for undertaking this role (selection will be based on expressed interest and relevance to the SDO TC/WG activities).
- Partners with effort on the tasks 4.3 and 2.4 could be a natural starting point for assigning individuals for undertaking the role of project liaison in relevant SDOs
- To avoid any conflicts of interest, the same person should not represent DSSC and its own organisation interests (should be a different individual if the organisation is already represented in the SDO or come from an organisation not represented).

Tools

Liaison individuals need to report back to the internal DSSC task force offering engagement status update and raising issues/alarms when necessary. Thereby,

- TF meetings can be used as an instrument for reporting and open discussion between its members and liaison individuals.
- A dedicated spreadsheet will be used for the liaison individual to report outcomes and raise issues out of each touchpoint with the SDO WG/TC.

2. Appendix 2: Description of relevant SDOs for DSSC

ISO/IEC JTC 1 is a consensus-based, standard group under the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). Joint Technical Committee 1 (JTC 1) focus on standardization in information technology; developing standards related to data management, including data interoperability, metadata, and data quality. The ISO/IEC JTC 1 subcommittees (SC) provide the governance, administration and strategy to enable the development of international standards in relation to its particular field of expertise. The Working Groups (WG) are established under a technical or subcommittee for the specific task of creating standards in line with their defined and agreed scope. Following we identify the subcommittees and working groups related to data spaces that DSSC members are part of:

ISO/IEC JTC 1/SC 27 - Information Security, Cybersecurity and Privacy: ISO/IEC JTC 1/SC 27 is responsible for developing standards related to information security, cybersecurity, and privacy. The committee addresses various aspects of securing information and data, including the development of guidelines, methodologies, and best practices to protect data and information systems from unauthorized access, breaches, and other security threats. This includes standards related to cryptography, security management, privacy protection, risk management, and more.

ISO/IEC JTC 1/SC 38 - Cloud Computing and Distributed Platforms: ISO/IEC JTC 1/SC 38 focuses on standards related to cloud computing and distributed platforms. The committee is responsible for developing guidelines and specifications that address various aspects of cloud computing, including interoperability, portability, security, and management. Standards developed by this committee help ensure that cloud-based services and platforms are consistent, reliable, and secure, enabling organizations to effectively utilize cloud resources. The SC38 committee has initiated important works to support dataspace standardizations under the Working Group 5 – Data in Cloud Computing and Related Technologies where the terms of reference are:

- Standardization in the area of data in cloud computing, distributed platforms, and related technologies, including but not limited to:
- Edge computing, federated or multi-cloud data platforms.
- Transparency in the control, use and sharing of data.
- Interoperability and portability in these environments.
- Mechanisms to address data and platform sovereignty concerns
- Establish liaisons and collaborate with other entities within and external to JTC 1 as appropriate. SC 38 has also established Liaison Coordination Groups (LCGs) for between Plenary interaction with SC 7, SC 27, SC 41, SC 42, and JTC 1/WG 13.

Under the Working Group 5 – Data in Cloud Computing and Related Technologies two projects has initiated that has an impact on data spaces standardization (1) ISO/IEC TS 10866 Organizational Autonomy and Digital Sovereignty and (2) PWI 20151 Dataspaces.

- **ISO/IEC AWI TS 10866** Information technology – Cloud computing and distributed platforms – Framework and concepts for organizational autonomy and digital sovereignty

This evolving document aims to elucidate the intricate nexus of digital sovereignty, organizational autonomy, and digital platforms. It is currently in the development phase and seeks to furnish organizations with a robust framework, enabling them to:

- Grasp the pertinent issues surrounding digital sovereignty and autonomy.
- Discern the essential attributes required for their potential digital platform and the requisite decision-making processes to address these challenges.
- Evaluate and harmonize the array of available options to realize their desired level of organizational autonomy.

This document serves as a valuable resource for all types of organizations and policymakers involved in matters related to organizational autonomy and digital sovereignty in the realm of cloud services and distributed platforms. Presently, it is in the preparatory phase and will be presented to the committee in due course.

- **ISO/IEC PWI 20151** Information technology – Cloud computing and distributed platforms – Data spaces. Proposal for new project received in February 2023.

ISO/IEC JTC 1/SC 41 - Internet of Things (IoT) and Related Technologies: ISO/IEC JTC 1/SC 41 is dedicated to standards related to the Internet of Things (IoT) and its related technologies. The committee focuses on developing standards that facilitate the interoperability, connectivity, and security of IoT devices and systems. This includes standards for data exchange, communication protocols, IoT architectures, and the integration of IoT technologies into various domains such as healthcare, smart cities, industrial automation, and more.

These subcommittees play a crucial role in shaping the standardization efforts within their respective domains, contributing to the development of consistent and interoperable solutions in information security, cloud computing, and the Internet of Things. Please note that the information provided is based on my knowledge as of September 2021, and there may have been developments or changes since that time.

ISO/IEC JTC 1/SC 32: This is Joint Technical Committee 1, Subcommittee 32. It focuses on data management and interchange standards, including concepts related to data spaces.

ISO/IEC JTC 1/SC 42 stands as the preeminent international standards committee entrusted with the task of standardization within the realm of Artificial Intelligence (AI). It operates as a collaborative endeavor, jointly established by the esteemed ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission). Functioning as the epicenter of AI standardization within both ISO and IEC, SC 42 is devoted to comprehensively addressing the expansive AI landscape. Its program of work extends to providing guidance to ISO and IEC committees engaged in the development of AI applications.

The current portfolio of work undertaken by SC 42 encompasses various critical domains, including: (1) Foundational AI standards, (2) Data standards associated with AI, (3) Big Data and Analytics, (4) AI trustworthiness, (5) Exploration of use cases and applications, (6) Examination of governance implications arising from AI, (7) Computational approaches integral to AI, (8) Testing protocols for AI systems and (9) Scrutiny of ethical and societal concerns linked to AI. Moreover, SC 42 maintains active collaboration with numerous ISO, IEC, and JTC 1 committees. This collaborative effort spans across horizontal technologies intertwined with AI and the multifaceted vertical applications of AI.

W3C, also known as the **World Wide Web Consortium**, is dedicated to developing web standards and guidelines that uphold the fundamental principles of accessibility, internationalization, privacy, and security. This commitment is further amplified through its various working groups, each focusing on data-related standards such as RDF (Resource Description Framework), OWL (Web Ontology Language), and other technologies contributing to the concept of data spaces.

Here's an overview of some of the key working groups within W3C:

- Dataset Exchange Working Group (Dataset Exchange WG): This group is tasked with the following objectives:
- Maintaining and updating the Data Catalog Vocabulary (DCAT), taking into consideration feedback and requests from the DCAT user community.
- Defining and publishing guidance regarding the specification and utilization of application profiles when requesting and sharing data on the World Wide Web.
- Decentralized Identifier Working Group (Decentralized Identifier WG): The mission of this group is to standardize several aspects related to Decentralized Identifiers (DIDs), including:
 - Establishing the DID URI scheme.
 - Defining the data model and syntax of DID Documents, which contain pertinent information about DIDs for specific use cases.
 - Outlining the requirements for DID Method specifications.
 - RDF Dataset Canonicalization and Hash Working Group (RDF WG) and RDF-star Working Group: The RDF Dataset Canonicalization and Hash Working Group's mission is to establish a standard for calculating a unique and deterministic hash of RDF Datasets. This hash is useful for various purposes, including the detection of changes in datasets. The group will also define RDF Dataset Canonicalization algorithms. In parallel, the RDF-star Working Group aims to extend a set of recommendations related to RDF and SPARQL. This extension will enable the concise representation and querying of statements about statements.
- Verifiable Credentials Working Group (Verifiable Credentials WG): This working group is responsible for maintaining the Verifiable Credentials Data Model specification and related Working Group Notes.

IEEE is a nonprofit formed in 1963 after the merger of two other professional associations, the Institute of Radio Engineers and the American Institute of Electrical Engineers. Today, it has one of the largest membership bases in the world. It provides nonbinding standards, conferences and other professional development activities for engineering, computing and technology, all with the aim of “advancing technology for the benefit of humanity.” There is one committee active since 2020 on Workshop on P3800 Data Trading System that will be finalized until the end of 2023. The IDS-RAM is part of the foundation for IEEE P3800.

CEN CENELEC workshop on Trusted Data Transaction. The objective and scope of this Workshop is Establish terminology, describe concepts and mechanisms in the field of data exchange to form a foundational understanding on which trusted data transactions can be based. And identify **attribute-based criteria** for the decision-making grid that baselines **how to create trust** in data transactions.

CEN-CENELEC Joint Technical Committee on Artificial Intelligence (JTC 21) was created on 1st of June to shape standardisation that ensures the development of trustworthy AI systems that respect fundamental

values and human rights recognized in Europe. JTC 21 produces standardisation deliverables in the field of AI and related use of data to address European market and societal needs. This work aims to advance EU legislation, policies, principles and values. CEN-CLC/JTC 21 is currently developing European standards which, in the future, would be able to provide manufacturers the presumption of conformity with the upcoming Artificial Intelligence Act. It includes aspects such as data quality, data governance and trustworthiness framework.

CEN-CENELEC Focus Group on Data Spaces. The concepts "Data" & "Dataspaces" are spread across several TCs. In order to support Data Economy as a foundational driver for Europe's Digital Transition and without the intention to remove any work from those TCs, there is a need to cover gaps and ensure a horizontal approach. Similarly, Cloud and Edge have industry-related use cases that could benefit from a systemic/horizontal approach to be combined with vertical inputs from connected TCs. For these reasons, and to connect to various stakeholder groups currently not yet connected to CEN and CENELEC, a Focus Group was created in March 2024. The Secretariat of the CEN & CENELEC Focus Group on "Data, Dataspaces, Cloud and Edge" was allocated to UNI. The group lasted till the creation of the JTC25 which absorbed the group purpose.

CEN-CENELEC JTC 25. It is the new Technical Committee established in September 2024 on "Data Management, data spaces, cloud and edge" to develop standards to support the widespread adoption of digitalization and the establishment of a fully functioning Single Digital Market for the EU. The goal is to ensure the European industry is more efficient, productive, competitive, and fully integrated in the global digital market. The JTC's work will help reduce compliance costs and market entry barriers, enhance innovation and competitiveness, facilitate seamless data exchange and create business opportunities. Special focus will be placed on ensuring interoperability. Furthermore, data-driven innovation can benefit a range of areas, such as healthcare, transport, and environmental sustainability.