DELIVERABLE D3.2 - DATA GOVERNANCE AND METHODOLOGIES - OUTCOMES PHASE I

PROJECT ACRONYM:	DS2
PROJECT TITLE:	DataSpace, DataShare 2.0
GA NUMBER NO.	101135967
WEBSITE:	www.dataspace2.eu
DUE DATE OF DELIVERABLE:	2025-06-30
SUBMISSION DATE:	2025-06-27
LEAD BENEFICIARY:	UoS
LEAD AUTHORS:	Stefano Modafferi (UoS)(Editor). Module and other content: Laura Carmichael, Nicholas Fair (UoS), Jutta Suksi, JP Soininen, Pasi Pussinen, Jarno Halme (VTT), Elias Dakos, Antonios Mpantis (ATC), Carlos Sanchez (INDRA).
REVIEWERS:	Stuart Campbel, Shmuel Bar
TYPE:	OTHER
DISSEMINATION LEVEL:	PUBLIC

DISCLAIMER

STATEMENT OF ORIGINALITY

This Deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

DOCUMENT HISTORY

VERSION	DATE	DESCRIPTION	NAME	ORG
V0.1	2025-05-12	Table of contents	Stefano Modafferi	UoS
V0.2	2025-05-22	Contributions	Jutta Suksi, JP Sonninen, Pasi Pussinen, Jarno Halme, Nicholas Fair, Laura Carmichael, Elias Dakos, Carlos Sanchez	VTT, UoS, ATC, INDRA
V0.3	2025-06-15	Final Draft	Stefano Modafferi	UoS
V0.4		Updated Draft	Jutta Suksi, JP Sonninen, Nicholas Fair, Laura Carmichael, Elias Dakos, Carlos Sanchez	VTT, UoS, ATC, INDRA
V1.0	2025-06-26	Final Document	Stefano Modafferi	UoS

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This document describes the work done in DS2 Work Package until M18 of the project. These solutions or applications are related to the tasks T3.1, T3.2, T3.3 and T3.4 of the WP and the objectives:

- O3: To specify methodologies and models to protect the sovereign rights of data owners and enable compliance with European data regulations over complex lifecycles of data sharing, aggregation & provenance.
    - O3.1: Provide knowledge & tools to support data sovereignty over complex data lifecycles
    - O3.2: Provide support to practitioners for governance and relevant policy regulatory compliance for data
    - O3.3: Enable management of rights to data sovereignty over its complex lifecycle
    - O3.4: Determine motivations and barriers for sharing data & strategies to lower barriers

The formal WP deliverables are all software (Type: OTHER), so the software, possibly source code, documentation, detailed progress tracking are primarily on the DS2 GitHub which will evolve and grow in content over time - https://ds2-eu.github.io/documentation/ . The online documentation includes module overview, architecture, components descriptions, example screen shots, information on installation and use and other matters. Within this accompanying document, an introduction to the WP is provided, followed by a brief description of software and non-software results. A dedicated section presents a summary of the progress for each software module. The deliverable also includes a description of KPI's, Primary Risks, and Primary Issues for the WP, followed by a conclusion section.

In summary, there are few significant risks and issues, all software and other activities is on track or exceeding expectations

1 INTRODUCTION

DS2 WP3 encompasses a comprehensive set of methodologies, survey analyses, and software modules, focussed on the themes of data governance and risk assessment.

This document discusses the results of the survey initiated in D3.1, it provides a manual for the creation of data products, and it discusses the policy management lifecycle and risk assessment.

From a software point of view, the key modules in DS2 which address policy and risk management in this WP are:

- DS2 Sovereignty Decision Support Module SDS
- DS2 Data Rights Management system DRM

In addition, module Policy and Enforcement (PAE) and Policy Creation (PCR) are additional policy centric modules developed in other work packages (resp. WP4 and WP6) and only briefly introduced for presenting the general overview, while SDS and DRM, developed in WP3 are presented according to the DS2 software template in Section 6.

The main partners involved are UoS, ATC, VTT, other technical partners also had resources in the WP and provided a support and harmonization role (e.g. INDRA formally in WP4), All users have minor resources in the WP and participated intermittently/as necessary in meetings and/or email exchanges.

The formal WP deliverables are all software (Type: OTHER), so the software, possibly source code, documentation and detailed progress tracking are primarily on the DS2 GitHub which will evolve and grow in content over time - https://ds2-eu.github.io/documentation/.

For M18, as mentioned the deliverable of WP3, but also similar ones in other RTD WPs (D4.1, D5.1, D6.1), are all formally OTHER/Software and to ensure consistency they follow a common structure and a content template was used. This being said, WP3 has used this deliverable to also report on this continued work from D3.1 delivered at M12 hence its structure is enhanced.

1.1 Document Structure

- 1: Introduction: Introduces the deliverable – this section
- 2: Provides a summary of the findings from the Attitudes and Behaviours Survey and the User Deep Dive Workshop and how they inform WP3 and other WPs activities (in-depth results can be found in Annex B)
- 3: Discusses and proposes a concrete methodology for the creation of data products.
- 4: Discusses the Risk Knowledgebase, which forms the foundation of the risk assessment process and briefly introduces the SDS module.
- 5: Addresses the policy management lifecycle, and it includes a brief description of the software modules relevant in the lifecycle: DRM, PAE, and PCR.
- 6: Dedicated to the presentation of the software modules developed in WP3, namely the SDS and the DRM.
- 7: Lists KPIs, risks and issues
- 8: Conclusion: Conclusion and next steps
- Annex A: References
- Annex B: Detailed User Research results and findings
- Annex C: The Three-Step Approach Manual for Building Data Products

1.2 Glossary and Abbreviations

A definition of common terms related to DS2 as well as a list of abbreviations, is available at https://www.dataspace2.eu/results/glossary

1.3 External Annexes and Supporting Documents

External Documents:

- DS2 D2.2 Requirements, baselines, KPIs, Architecture & Specifications
- DS2 Communication, Dissemination, and Exploitation report (M18)
- DS2 Risk identification spreadsheet

2 FURTHER EXPLORATION OF DATA SHARING ATTITUDES AND BEHAVIOURS

2.1 Data Sharing Attitudes and Behaviours Survey & User Workshop Results

This section describes the activities and results associated with the Inter-organisational Data Sharing Attitudes and Behaviours Survey and the follow-up User Deep Dive Workshop as initially reported in Section 2.1.3 in D3.1. The aim of the survey and workshop was to provide an understanding of the existing attitudes of key stakeholders, with a particular focus on stakeholders within the Use Case domains of SmartCities, Green Deal and (precision) Agriculture, to the sharing of data between collaborative business partners, and to gain some knowledge of current inter-organisational data sharing practices. The results are used to inform further WP3 and wider DS2 work by providing insight into the barriers to the adoption of data sharing practices and how they might be overcome through considered design and development of modules and tools within the DS2 project.

2.1.1 Survey Overview

The survey was developed as a mixed methods, fully anonymised, online survey using the University of Southampton approved and secure instance of the Qualtrics platform. Ethical approval for the survey was awarded under ERGO ID 100306. All relevant participant information, opt-in and consent was provided at the start of the survey. There were no specific gender or other Equality, Diversity and Inclusion considerations as no personal or demographic information was requested.

The survey was disseminated according to the 3-pronged strategy detailed in Section 2.3.1.2 in D3.1 between November 2025 (M11) and February 2026 (M14). All consortium partners were requested to leverage their networks and their social media to promote the survey, and the survey details were emailed directly to 141 EU Digital Innovation Hubs (DIHs) identified via the EU EDIH Catalogue, with a specific focus on DIHs in the following domains: Big Data; Blockchain; Communication networks; Data; and Logistics.

In total 93 responses were received, the upper end of the original target response rate (of 50-100 responses), of which 2 declined to consent and a further 16 lacked the relevant expertise. This resulted in a sample size of 75. The majority of respondents (54%) declared a knowledge of data sharing policies and practice, while a further 29% either had direct responsibility for implementing data sharing or had data sharing decision-making power within their organisation and job role. In addition, the vast majority of the participants (including most of the ‘Other’ category - project/communications/technical/executive/relationship managers) occupied Director or Manager level job roles. The remaining responses were from individuals with researcher, advisor, operator or engineer roles (see below).

Figure 1: Survey Demographic - Expertise and Job role

A wide range of industry sectors were also represented (see below).

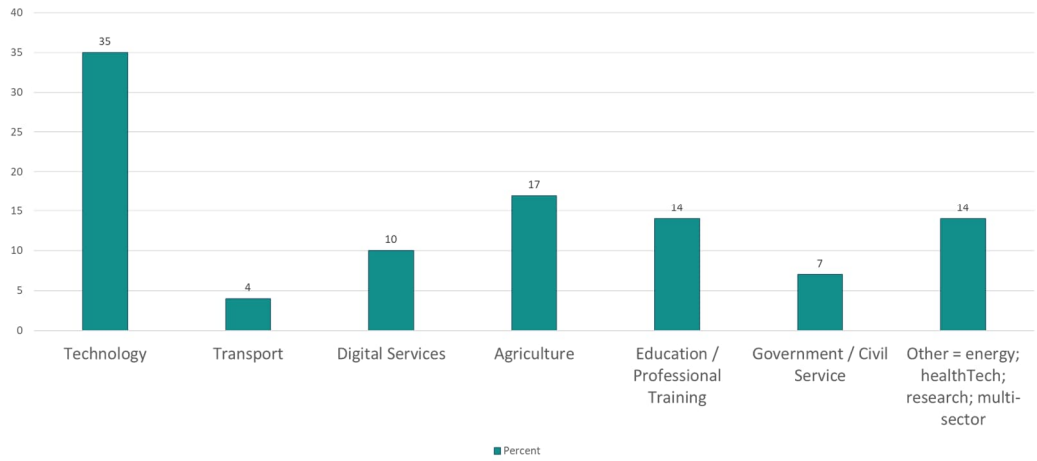


Figure 2: Survey Demographic – Industry Sectors

As a result of the diversity shown, the survey responses can reasonably be considered as representative of the opinions and practices of knowledgeable and expert individuals working in high-level job roles directly related to data sharing at companies from diverse sectors.

A detailed analysis of the gathered data is available in Annex B and below a summary is reported.

2.2 Summary of Research Findings

The survey and Deep Dive Workshop indicate a broadly positive attitude to data sharing and a reasonable amount of data sharing activity, infrastructure and processes already in place. Cross-sector data sharing is considered desirable, although altruistic data sharing with state agencies, umbrella organisations and industry hubs to grow the sector is less favoured. There is also a broad perception that data sharing is complex and has multiple challenges, which results in just half of all stakeholders expressing a likelihood to increase data sharing activities in the future.

The results indicated that a number of important specific gaps were present in the understanding or implementation of:

1. data sovereignty
2. legislation
3. data sharing agreements
4. licensing vs restrictions
5. security / risk assessment
6. training

In addition, a range of important strategic, operational, technical and inter-organisational challenges to data sharing were identified. The most significant of which are:

1. Strategic – establishing trust; potential for misuse; establishing monetary value of data
2. Operational – associated costs; complexities of Data Sharing Agreements; lack of skills
3. Technical – transparency; quality; access controls
4. Inter-organisational – governance and ownership; trustworthy partners; access controls

A number of important topics also remain open for future discussion, including business models / monetisation, and the relationship between Data Space Governance Frameworks (aka Rulebooks) and the Data Policies found in Data Products.

2.3 Recommendations Resulting From Survey & Workshop

It is important that within the DS2 platform, tools, services and use cases as many of the concerns, challenges and open discussion points as possible are adequately addressed so as to effectively support and promote trustworthy, simple and valuable data sharing in the future. As such, the key findings from this research that will be used within DS2 are given in following Table 1.

Topic	Finding	Action
Data Sovereignty	Lack of agreement over definition and understanding of the term	Decide on a comprehensive definition, incorporating both the jurisdictional and control aspects of data sovereignty, and use it to inform Knowledgebase extension (WP3)
Data Sovereignty	Data Space Governance Frameworks (aka Rulebook) should be as similar as possible in order to facilitate sharing across data spaces	All Use Cases to base their Data Space Governance Framework on the SITRA rulebook reference (WP7)
Data Sovereignty	Lack of agreement concerning the priority between the Data Space Governance Framework and the Data Policies found in Data Products	Active design consideration for the PCR, Portal/IDM and PAE modules (WPs 4 & 6)
Legislation	Lack of in-depth understanding of the various pieces of relevant legislation	Explore the possibility of inclusion of legislative compliance within Knowledgebase extension and its potential for inclusion in risk assessment (WP3)
Data Sharing Agreements (DSAs)	DSAs are too complex and many aspects are not well understood	Use the least understood aspects to focus development efforts in the PCR module to ensure adequate simplification and automation in these areas (WP6)

Licensing & Restrictions	Tension between datasets that are required (or chosen) to be Open Access and the data usage and access restrictions that may be present in the Data Space Governance Framework and the Data Policies found in Data Products	Inform the development of the Data Product description process and content (WP3) Ensure PCR module accounts sufficiently for open datasets (WP6) Use the finding to inform on-going discussions concerning Data Space business models and sustainability (WP1 & WP3)
Security & Risk Assessment	Risk assessments only sporadically conducted Main focus of security controls is only on Access Controls Despite this focus, controlling Access remains an important concern when sharing with another organization	Ensure the SDS module(s) supports more frequent risk assessment by simplifying and automating the process (WP3) Ensure the Knowledgebase extension accounts for the complexities of Access Controls as defined in multiple places within and between Data Spaces (see 4.1 below) (WP3) Ensure the DS2 enabled risk assessments provide information about other important controls and control strategies that should be implemented beyond Access Controls (WP3)
Monetary Value of Data	Placing a monetary value on a dataset (that is not open by default) is extremely difficult	Understand how the 'value' of a dataset can be estimated sufficiently well that it can be accounted for in the Knowledgebase extension to contribute to the risk assessment calculations (WP3) Use the finding to inform on-going discussions concerning Data Space business models and sustainability (WP1 & WP3)

Table 1. Key findings from the survey and workshop.

3 DS2 DATA PRODUCTS AND DATA OFFERINGS

The development of the DS2 opportunity-based method aiming to interlink regulatory and business opportunities to the data space data lifecycles and technical implementations through identifying data products (see D3.1), has continued by focusing on the Theme 4 Data Products approach that was described in D3.1 in detail in close co-operation with the DS2 Use Cases.

3.1 The evolving definition of data product

The process of developing DS2 data products with the DS2 Use Cases is made more complex by the fact that the definition of the concept of data products in data spaces is still evolving. The Data Spaces Support Centre (DSSC) introduced the first Data Product Building Block on 11th October 2024 in its Blueprint 1.5. In the DSSC Blueprint 2.0 from 7th March 2025, the concept was further developed and linked to services offered to data space participants through the concept of a Data Space Offering. At the same time, the definition of data products and related concepts have evolved through the standardization initiative CEN CENELEC Trusted Data Transactions Working Group that released its first draft 3rd July 2024 and its Part 2 is currently under draft.

3.2 Co-creation with the DS2 Use Cases

Despite this on-going evolution of the definition of Data Products, methods, manuals and templates for effectively describing DS2 Data Products have successfully been co-created in close co-operation between WP3 and WP7 Use Cases. Working backwards from the Use Cases, it has been possible to identify the DS2 Use Case specific data product specifications and data product flows. In addition to specifying the content of the data product through data product specifications, the DS2 Use Cases have identified the data product flows both within and across data spaces. The results are reported in D7.1. Further development and finalization of the data product specifications and data product flows will continue in M19-M36 of the project.

3.3 Three-Step-Approach for identifying data products and data offerings

As a result of the close co-creation process with WP7 Use Case partners, DS2 has developed a Three-Step-Approach Manual for identifying data products and related service offerings:

- Step 1: Understanding data products in the context of data spaces (Knowledge package)
- Step 2: Specifying value creation services and data products (Method and template)
- Step 3: Identifying data product flows (Method and reference to examples)

The Three-Step-Approach can be found in Annex C.

3.4 Technical implementation of data products

One target of the co-creation with the Use Cases was to identify and describe the data products shared within and between data spaces. A Data Product is a vehicle by which datasets and their associated policies, such as business and usage conditions, can be bundled together into a single 'product' that ensures that the consumer knows exactly what they get and under what conditions. Data Products describe the data itself (metadata definitions), but in this context most such metadata is already available in e.g. the Data Catalog Vocabulary definitions (DCAT) (W3C, 2024). Consequently, in addition to dataset metadata descriptors, Data Products should further include descriptions of the data rights holder, pricing, license terms, service level agreement (SLA), other conditions…etc.

A number of technical implementation options for Data products were reviewed, including DCAT itself, which was found to be mostly limited to dataset metadata descriptions; the TM Forum Product Model, which was primarily intended for full product lifecycles mainly in the telecommunications sector; and the Data Product Descriptor Specification (DPDS), the Open Data Contract Standard, and the Data Contract Specification, all of which, despite having a number of strengths, were either weak concerning the ability to incorporate the business aspects of a Data Product or not sufficiently well-established, or both.

After review, the Open Data Product Specification (ODPS) (https://opendataproducts.org/) was identified as offering the most benefits and advantages for the technical implementation of data products for a number of reasons:

1. It is recommended by the EU's Data Spaces Support Centre (DSSC) Blueprint 2.0 guidelines
2. It is machine-readable in JSON (enabling automating workflows and programmable data management)
3. It can account for the business-related aspects of data sharing, including pricing, SLAs, and terms & conditions
4. It is an open specification being developed under the Linux Foundation
5. It can be linked to existing dataset metadata description systems (e.g. DCAT)
6. It can integrate with data transaction specific policies defined using Open Digital Rights Language (ODRL)

As a result, ODPS was selected as the preferred suggestion for the DS2 modules (inc. in particular the CAT module) to use for Data Product technical implementation solution.

4 DS2 RISK KNOWLEDGEBASE

Part of the DS2 pipeline (prior to reaching a Data Sharing Agreement [DSA]) includes the assessment of the risk of data sharing with a specific focus on data sovereignty. Consequently, DS2 builds the Sovereignty Decision Support module. The module is based on the UoS background of Spyderisk, which, in response to the User research findings (see section 2.2 of this report), is being extended in the ways outlined below in Section 4.1 to support the sovereignty risks. The SDS module itself is presented in detail in Section 6.3.1.

4.1 Risk assessment

Data spaces need robust risk management frameworks in place that can support organisations to identify, prioritise, manage, and mitigate risks including those related to the security of data (e.g., Agencia Española de Protección de Datos [AEPD] & European Agency for Cybersecurity [ENISA], 2024; Pitkänen et al., 2025). Risk assessment is considered one crucial element in supporting the establishment of trust between Data Providers and Data Consumers, and between multiple Data Spaces. For instance, complexities in establishing these trust relationships have been highlighted by Users as both a major strategic and inter-organisational challenge to data sharing (see section 2.2 above). Consequently, the results from the Users' Attitudes and Behaviours research (see 2.2 and Annex B) have been used to inform this extension of the Knowledgebase that underpins the risk modelling and assessment process.

The DS2 D3.1 report on data governance and methodologies highlighted how risk assessment tools can be used to support organisations in making consistent and transparent governance decisions as part of data spaces in terms of e.g., data access, sharing and linkage. DS2 specifically looks at how Spyderisk (https://github.com/Spyderisk) — an existing semi-automated risk assessment tool based on systematic cause-and-effect modelling of threats — can help to automate various aspects of information security risk assessment reporting in the data space context. To do this, extensions to the Spyderisk modelling software and knowledgebase need to be made so that risk analysts can use this tool to assess information security risks specifically for business-to-business (B2B) data transactions in data spaces, including those related to data sovereignty.

Results from the User research findings indicate that there is overall a lack of understanding concerning exactly what data sovereignty means, which is negatively impacting data sharing activities. This has informed an on-going process of developing a comprehensive, clear and consistent definition of the concept, which sufficiently accounts for the jurisdictional/legislative aspects of data sovereignty alongside the 'control' aspects. For more information about meanings and interpretations of data sovereignty in a data space context, see the DS2 D3.1 report on data governance and methodologies.

4.2 Risk modelling: a data sovereignty perspective

From a high-level perspective, DS2 proposes that a key information security goal related to data sovereignty is about: ensuring that a data provider can exercise appropriate and effective control over the access and usage of shared data throughout its lifecycle (e.g., Abbas et al., 2024). Such control over data access and usage should be supported through both organisational and technical security mechanisms (e.g., Abbas et al., 2024). For instance, Lohmöller et al. (2022) state: *"the notion of data sovereignty, i.e., one of the critical concepts of data ecosystems, currently lacks a clear and common definition […]. If used in the context of data ecosystems, researchers generally agree that data sovereignty relates to control and ownership of data items, together with specific claims and obligations made by involved parties […]"*. This also builds on the primary

understanding reported by Users that data sovereignty relates to *"the ability of the data owner to decide how to share and use its data"* (survey question 34), a statement which was developed from the DSSC definition of data sovereignty as *"the ability of individuals, organisations, and governments to have control over their data"* (DSSC Glossary, v3, 2024). From a computer science perspective, the notion of control is commonly used in *"computer security"* with regard to *"access control"* and *"usage control"* (Lazaro & Le Métayer, 2015).

'Control' and its related notions of ownership and governance is highly nuanced in Data Spaces, where various policy documents within and across Data Spaces all have something to say on the matter (e.g. Data Space Governance Frameworks (including for multiple, linked Data Spaces), Data Space Interoperability Agreements, the Data Policies found in Data Products, and Data Sharing Agreements). Further, in some cases, data providers will release data to a data consumer for use in the consumer's own environment over which the data provider has no control (e.g., Gil et al., 2022). In this situation, a data provider will require certain assurances that shared data would be used responsibly in accordance with associated policies and contractual agreements, and that appropriate control strategies are in place on the consumer-side, such as to prevent unauthorised access to and use of shared data which would be in contravention of a data access and usage policy (e.g., Gil et al., 2022; Kelbert & Pretschner, 2018). Hence fully and accurately understanding 'control' as it relates to Data Sovereignty in its various forms has been, and remains, a central aspect of extending the Spyderisk Knowledgebase effectively.

4.2.1 Policy definition, implementation and enforcement

Data access and usage policy definition, implementation and enforcement are viewed as key components of data spaces — and therefore become a means for supporting the 'control' aspect of data sovereignty in practice (e.g., Steinbuss et al., 2021; DSSC, 2024). Hence, attention needs to be paid to how the security of shared data (e.g., its confidentiality, integrity, availability) is protected across its lifecycle to ensure appropriate use and authorised access (Bartsch et al., 2022). Consideration also needs to be given to how the security of data access and usage policies is being upheld — especially where there is a need to protect the confidentiality of a policy containing information that would be viewed as sensitive (Bartsch et al., 2022).

As part of Spyderisk knowledge acquisition, the need for additional controls specifically related to data sovereignty in data spaces, such as those around policy enforcement, are therefore being considered. For instance, data policy enforcement technologies include *"digital rights management"*, *"data leakage prevention"*, *"access control"* and *"usage control"* (Zrenner et al., 2019). Further, *"security protections for data sharing"* in data spaces, include *"watermarking"*, *"certification"*, and *"smart contracts"* (Abbas et al., 2024). Indeed, this also reflects one of the discussion points in the Users' Deep Dive Workshop, where the topic of policy enforcement was a concern raised by some participants.

However, it is crucial to acknowledge that the enforcement of data access and usage control policies on the consumer-side (including security requirements) is somewhat challenging to achieve in practice (Hellmeier et al., 2023) — given that it necessitates *"system-specific approaches and trust relationships with enforcement points, possibly located on remote platforms"* (Brost et al., 2018). Different system architectures enabling usage control may also give rise to various degrees of sovereignty (Zrenner et al., 2019). Moreover, the possibility remains to bypass usage control mechanisms — consider the example outlined by Steinbuss et al. (2021) where technically *"a usage control system may control taking screenshots and printing, but it cannot prevent a person to take a photo from the screen displaying the sensitive data*".

4.2.2 Knowledge acquisition: identifying risks and threats

As part of the knowledge acquisition process for risk modelling, an initial literature survey has been undertaken to identify some key requirements related to data sovereignty — an overview is provided in the following Table 2.

Source	Brief description
Akbari Gurabi et al. (2024)	Explores *"Privacy-Preserving Machine Learning [PPML] in Sovereign Data Spaces"*, providing a *"Conceptual Approach to Include PPML into a Data Space"*, which includes a workflow that incorporates risk assessment.
Altendeitering et al. (2022)	Focuses on data sovereign *"collaborative AI pipelines"* and categories *"lessons learnt"* using three types of *"data challenges"* identified by proposed by Gröger (2021) related to *"data management"*, *"data democratisation"* and *"data governance".*
Antony & Sarshar (2025)	Highlights both *"technical"* and *"non-technical"* requirements for supporting data sovereignty in smart cities.
Brost et al. (2018)	Outlines some *"security requirements*" for industrial data space systems.
Da Silva Carvalho et al. (2023)	Focuses on *"personal data sovereignty"* and *"personal data governance framework"* in the context of *"cross-border digital public services"*.
Gil et al. (2022)	Centres on data sovereignty and distributed systems, offering an overview of a *"distributed usage control framework"* as a means for *"providing data sovereignty in distributed systems".*
Hellmeier et al. (2023)	Outlines different types of challenges with implementing data sovereignty in practice — i.e., *"organizational", "technical",* and *"personal & emotional"* challenges — and approaches for addressing these challenges.
Lohmöller et al. (2022)	Centres on data sovereignty in *"data ecosystems"* ("data-driven business models"), highlighting the need for *"trusted remote policy enforcement, verifiable data tracking, and integration of resource-constrained participants".*
Marino et al. (2023)	Sets out some *"properties of secure data spaces*".
Opriel et al. (2024)	Proposes some *"design principles"* for *"data sovereignty in inter-organizational information systems"* with focus on the *"automotive industry"*.
Otto & Jarke (2019)	Considers *"design requirements"* for *"multi-sided data platforms",* includes an *"international data spaces"* case study.
Steinbuss et al. (2021)	International Data Spaces Association Position Paper on Usage Control.
Zrenner et al. (2019)	Focuses on data sovereignty in business data ecosystems.

Table 2.: Overview of initial literature survey to identify key data sovereignty requirements.

Further, some initial work has been undertaken to identify sovereignty threats and prepare them for the encoding in the knowledge base can be found at the following links:

A more detailed data sovereignty requirements table above: https://github.com/ds2-eu/Sovereignty-Decision-Support-System/blob/main/Sovereignty%20Requirements

A list of some initial threats related to loss of control over data access and use: https://github.com/ds2-eu/Sovereignty-Decision-Support-System/commit/4598081b3188a59e91c0137310cd43d5d55da1b3

A validation process is ongoing with the Project Use Cases to understand how relevant and valid the threats are for them, and they are now being encoded in the extended Spyderisk Knowledgebase.

4.3 Supporting Data Sovereignty-Based Risk Assessment

To address issues identified during the User research, and as presented in D2.2 and D3.1 , the DS2 Sovereignty Decision Support system (SDS) will provide a mechanism to enable the automated identification of data sovereignty-focused risks when undertaking inter-organisational data sharing. This is achieved by requiring that both the data provider and data consumer perform a local risk assessment using one (sub)module of the SDS tool and then share the result with their counterpart via a second SDS (sub)module that enables local risk assessment comparison while ensuring that individual organisational system privacy is maintained. In this way risk assessments and comparisons not only highlight system-level risks (and control strategies), but also provide a mechanism for an inter-organisational "trust assessment". In this way the SDS module(s) directly addresses both the strategic (establishing trust) and inter-organisational (trustworthy partners) challenges that were highlighted during the User research (see 2.2).

Furthermore, the SDS module is intended to be used BEFORE a Data Sharing Agreement (DSA) is reached, hence the outcomes of both the local risk assessment and the inter-organisational 'trust assessment' can, where relevant, be used to support decision making and also to inform the DSA content itself. Provision for this has been communicated to the PCR module developers. The SDS module(s) does not have an active role once actual data sharing is operational.

5 DS2 POLICY AND POLICY MANAGEMENT

In Data Spaces, the policies are part of the data sharing contract and are used to implement the contract-based transaction process. There are several components needed to manage the data sharing contracts. The components control data sharing based on the conditions defined in the signed data sharing contract. The different modules attaining the policy management are developed across WP3, WP4 and WP6. A brief description is provided in this section to provide the overarching view, while a detailed description for the modules developed in WP3 is presented in Section 6.

This Section addresses the technical implementation of Policies (Data Access and Data Usage) as part of the data-sharing contracts. The basic structure of the data-sharing contract can be described as below in Figure 3:

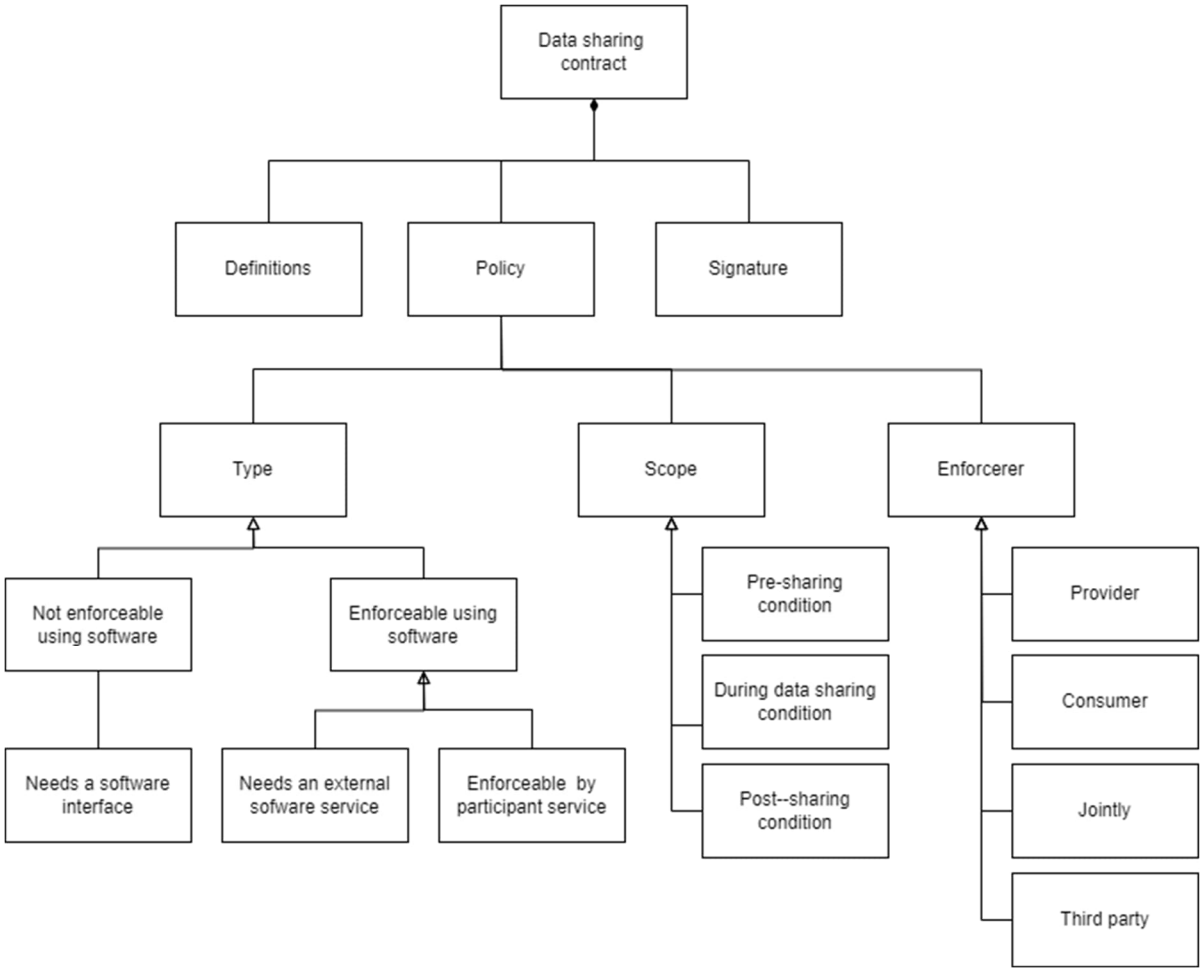


Figure 3: High level data model of the data sharing contract

Policies applicable to data transactions that are not enforceable using software (i.e. rules), e.g. rulebook level terms and non-automated data license terms, are left outside the scope of this Section, which will be covered in the final phase of DS2 in connection with the development of business cases (Theme 3).

The technical implementation of the policies in data spaces is connected to the contract-based data transaction process as detailed in the logical architecture model of data spaces[1]

The contract-based data transaction process is depicted in Figure 4:

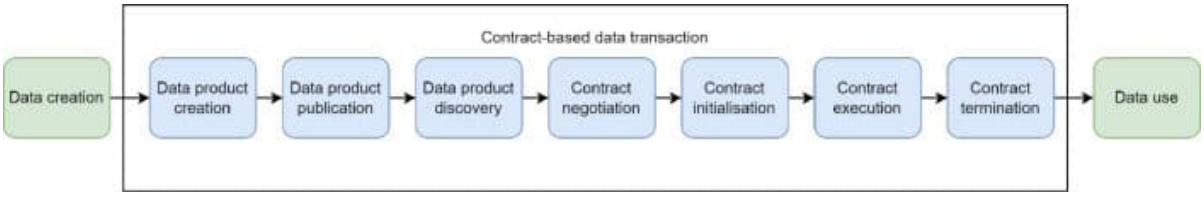
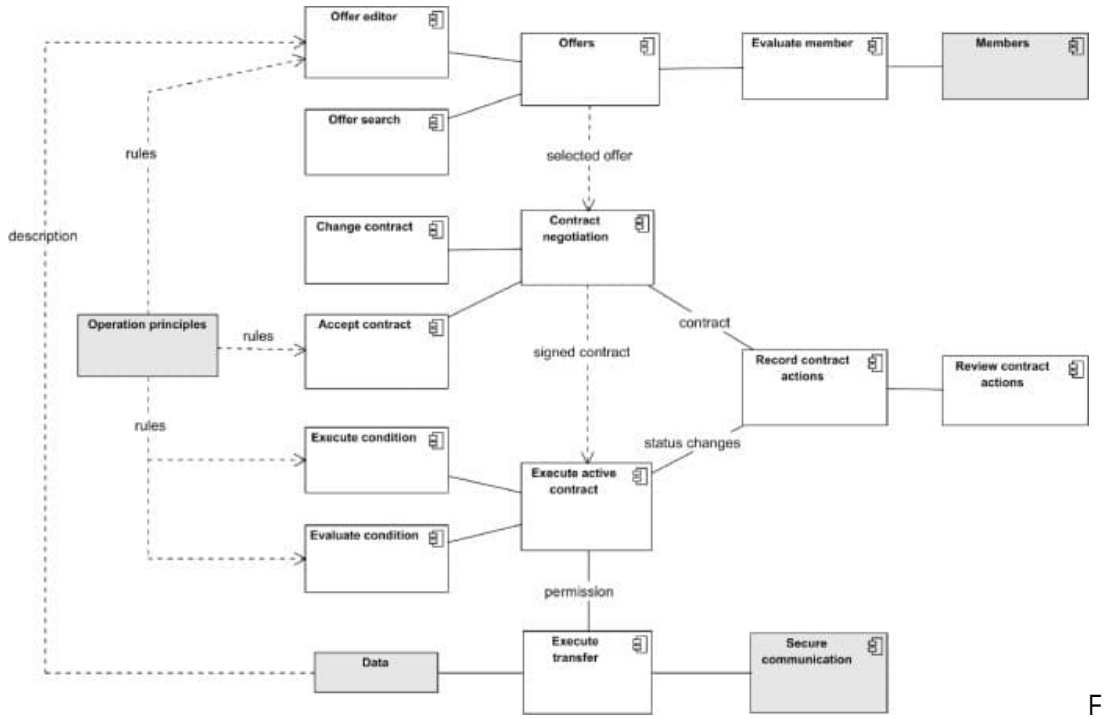


Figure 4:Contract based transaction process

Data spaces and DS2 will implement the contract based transaction process. Functional Components for management of data sharing contracts[1] are presented in Figure 5:

[1] (J.-P. Soininen and G. Laatikainen, "What is a data space—Logical architecture model," *Data in Brief*, vol. 60, p. 111575, 2025, doi: https://doi.org/10.1016/j.dib.2025.111575 .)



F

Figure 5: Management Components

The technical implementation of policy management can be divided into three phases:

- Policy creation;
- Policy enforcement; and
- Policy monitoring.

In the policy creation phase the data provider creates the terms under which the data products can be shared. Policy enforcement verifies that these rules are being obeyed – e.g. if the provider has granted access only on a certain time period, or the policy enforcement validates the time period. Policy monitoring refers to methods to manage what policies, or elements of policies, are valid

The policy management phases encompass the above data sharing contract components (see Figure 5) as follows:

1. Policy creation: Expressed in the Offer andsubject to Contract negotiation
2. Policy enforcement: Takes place at the moment the contract is signed
3. Policy monitoring: Executed and evaluated through conditions.

Below in Figure 6 is an example of the contract conditions processed from initialising the contract to terminating the contract:

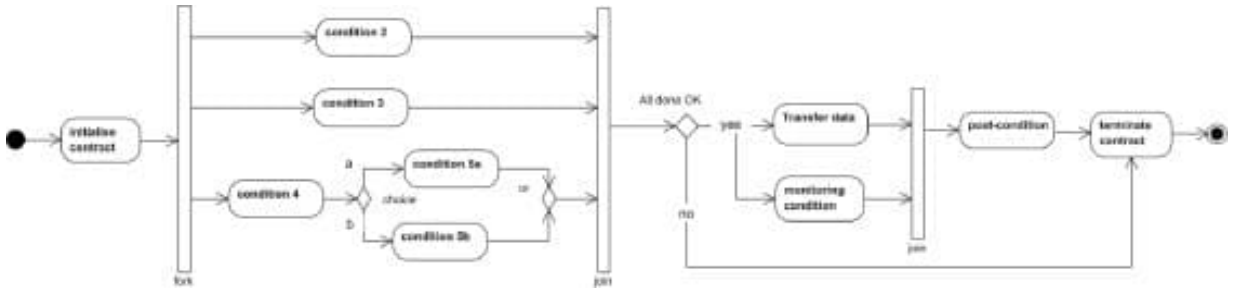


Figure 6: Policy Management Phases

5.1 Policy Creation

The DS2 Policy Creation Module (PCR) serves as a tool for Data Space participants (e.g. authorities, providers), facilitating the creation and comparison of dataspace policies through an intuitive user interface (UI), which can then be matched against other dataspace policies. This module leverages the Open Digital Rights Language (ODRL) to offer a sophisticated yet simplified approach to policy generation. Each policy created through the UI is automatically exported as a JSON file, containing all necessary metadata and ready for storage or sharing. The PCR module also enables dataspace authorities to manually compare policies either within a single dataspace or across multiple dataspaces through a text-based comparison feature. It is developed in WP6 and presented in D6.1.

5.2 Policy enforcement

The primary function of the Policy Agreement and Enforcement Module (DS2 PAE) is to ensure compliance with the established policies and regulations governing data exchange among participants of different data spaces. Policies, regulations, and agreements are considered synonymous under the term policy. The module is built as a set of extensions of the DS2 EDC Connector which is part of IDT.

Policies serve two main purposes: Access Control and Usage Control. Access Control determines whether access to data is granted or denied. Usage Control dictates how the data can be used once access is granted.

In the context of dataspaces, policies define who can access the data and under what conditions it may be used. The PAE module enforces the policies associated with a data-sharing contract—an agreement between a provider and a consumer.

Policies are evaluated before any data transmission begins and during it. Rules that cannot be automatically enforced are simply recorded alongside the contract for accountability purposes. These rules may later require monitoring or human oversight.

The Policy Agreement and Enforcement Module could optionally notify relevant components when such actions are required, ensuring that data sharing complies with the agreed-upon policies throughout its lifecycle.

PAE is described in detail in D4.1.

5.3 Monitoring

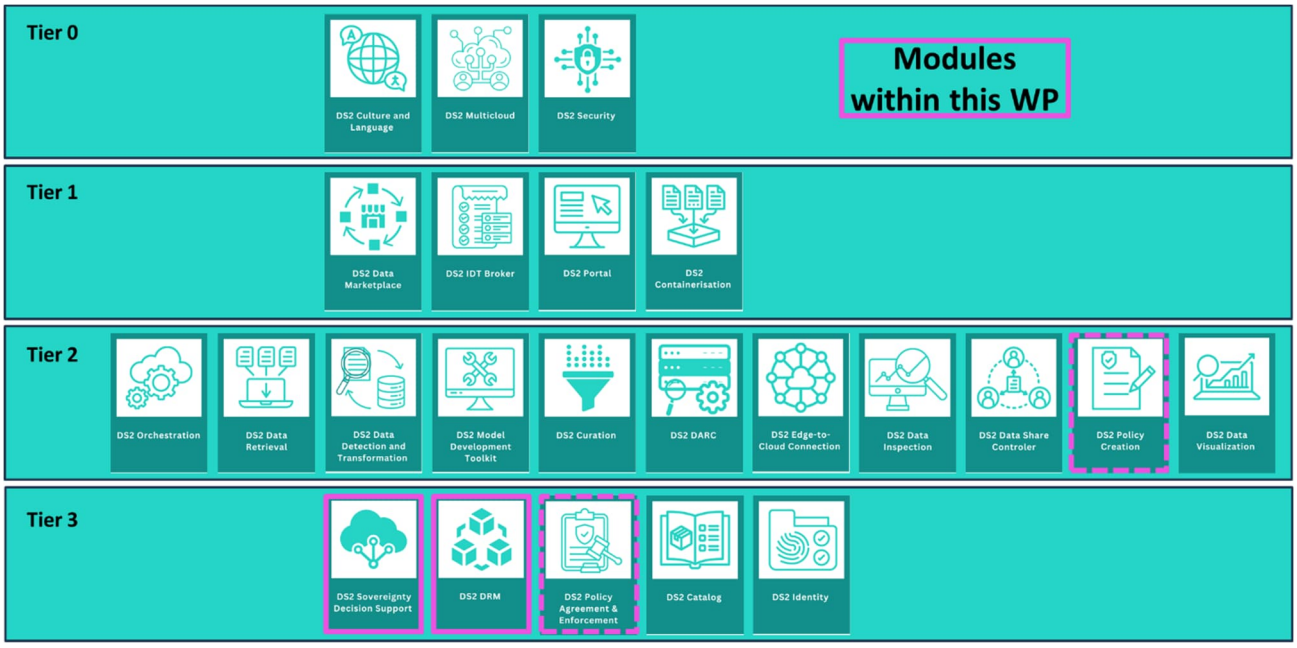
The DS2 Digital Rights Management (DRM) module provides a blockchain-based mechanism for managing, tracking, and validating data rights transactions within and across dataspaces. It records key actions—such as data access, usage, and policy enforcement events in a secure, tamper-resistant way, creating immutable and verifiable logs. These records help ensure transparency, traceability, and accountability among participants,

while supporting compliance with data usage agreements. The DRM module can be used by individual participants or by multiple participants in the same or different dataspaces. Additionally, other DS2 modules that interact with data can also submit logs to the DRM system, contributing to a consistent and trustworthy audit trail across the DS2 framework. A user-friendly interface is provided to allow participants to view their transaction and policy logs, enabling easy access to detailed records and improving the visibility of data-related activities.

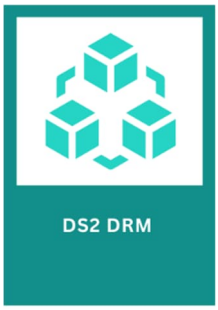
6 WP3 MODULE DESCRIPTION AND OVERALL STATUS

This section focusses on the technical presentation of the software modules developed in WP3 (SDS and DRM). In the figure also the other two main relevant modules are presented.

6.1 Modules



This WP has the following modules:

Module	Purpose
Tier 3	
SDS 	SDS, based on the UoS background of Spyderisk, supports the risk assessment of potential data sharing int eh context of Data spaces. It works requiring that both the data provider and data consumer perform a local assessment and then share the result with the counterpart. Separate sub modules of the SDS allow the execution of the local assessment and the comparative analysis of the results supporting "risk and trust assessment".
DRM 	To enhance the management and security of digital asset transactions through a robust blockchain-based Data Rights Management (DRM) system. It is designed to perform critical functions, including the notarization, tracking, and validation of all data rights transactions both within individual Dataspaces and across multiple participating Dataspaces

The modules fit is as follows:

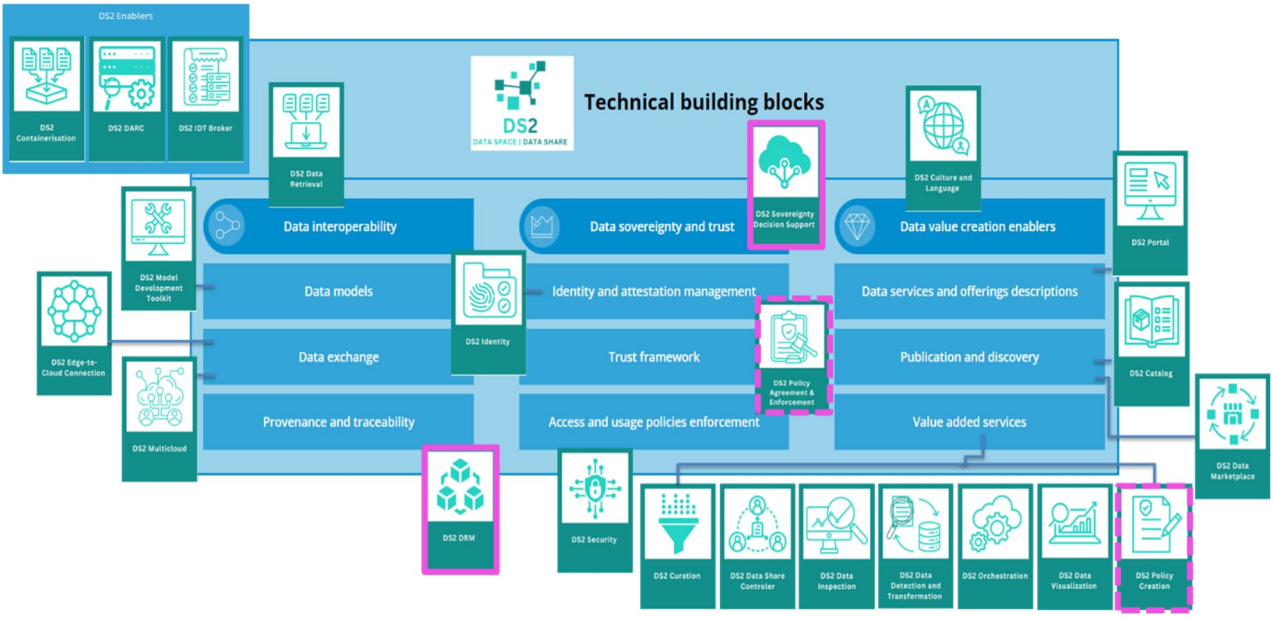


Figure 7: Modules fit

6.2 Status

Module	Task	Partners	License Type	Software Status @ M18	Estimated Completion Month	How to Install "vs 0.1"	How to Configure "vs 0.1"	Marketing Video	PORTAL Helm Chart	Marketplace Entry
SDS	T3.2	UoS	Apache 2.0	50%	M30	On Github	On M18 by education session	For M30	M30	For M30
DRM	T3.3	ATC	Apache 2.0	50%	M30	On Github	On M18 by education session	For M30	For M24	For M30

6.3 Software progress

The software, documentation, and progress for the modules developed is located at the DS2 GitHub repository accessible by the links below. These links give the current module documentation and at the top further links to the software and module progress. At this interim state of the project the modules are still under development by-plan and until more stable documentation is limited to more of an overview. The framework for documentation and progress monitoring is identified in Annex A of this WP6 deliverable "DS2 D6.1 - FEDERATED IDT PLATFORM - PHASE I". Progress is monitored through two-weekly sprints with detailed highlights. In terms of the documentation, it will be improved with How-Tos, API definitions, etc. overtime.

Module	Link
SDS	https://ds2-eu.github.io/documentation/modules/SDS/
DRM	https://ds2-eu.github.io/documentation/modules/DRM/

Note the progress tables of Section Error! Reference source not found. and 6.3.1.3 represent the list of Functionalities and status of completion at M18 which were defined for each module in the D2.2 Architecture annexes. Progress was measured via bi-weekly highlights/sprints/meetings. Highlight 1 represented 2025-02-04 and highlight 9 is the last two weeks of June 2025. Post reviews and the features lists will be updated/enhanced. Also indicated in that table on the right is an expectation of when each feature will be completed (100%) at milestones M18, 24, 30, and 36 noting that to ensure a suitable validation plan all software should be completed by M30. For M18, since some features will still be in-situ then the percentages may be less than 100%.

6.3.1 SDS: Sovereignty Decision Support System

6.3.1.1 Architecture Diagram

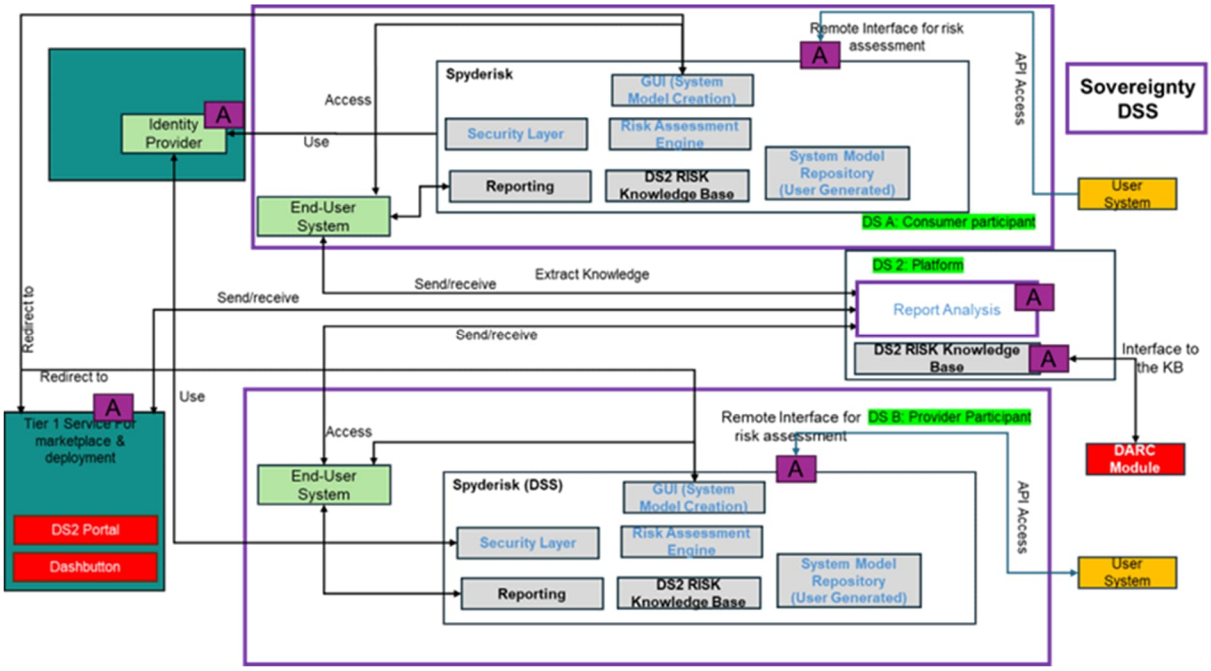


Figure 8: SDS architecture diagram

6.3.1.2 Sample Interface

There are two different user interfaces for this module. One is the interface for risk assessment derived and customized directly from Spyderisk (see below the NON customised version) and one is defined for the report analysis which is still under design.

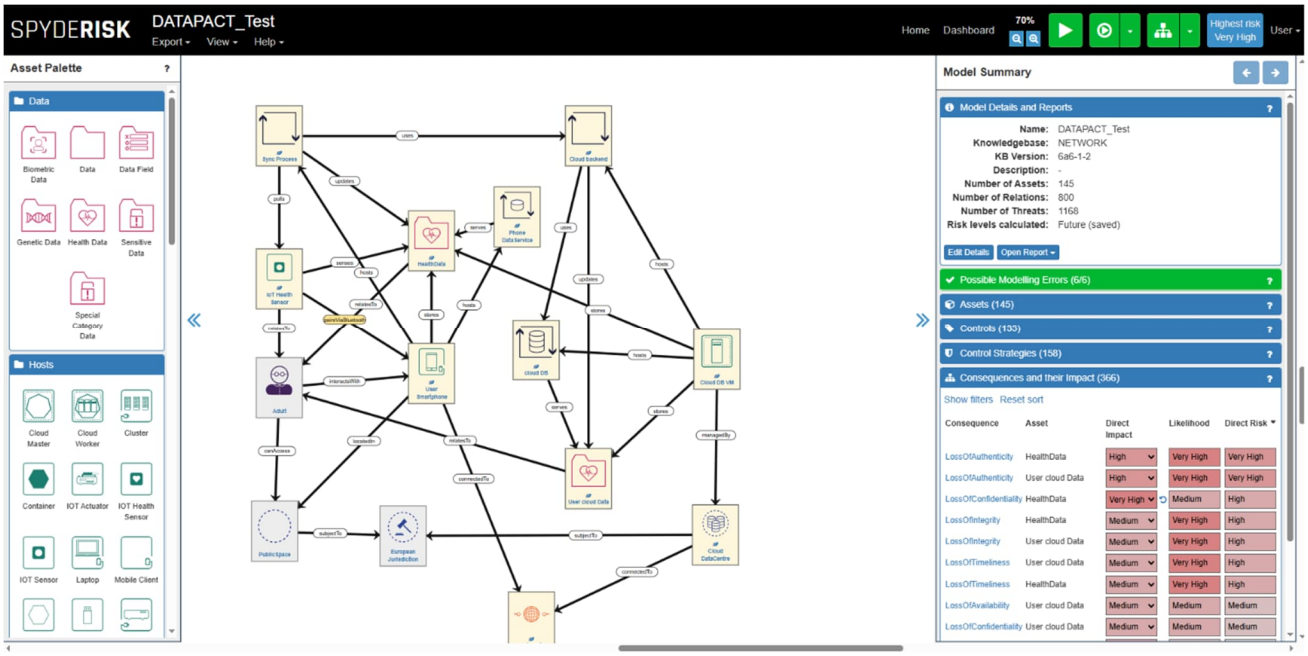


Figure 9: SDS interface

6.3.1.3 Primary Feature Progress

												Planned			
Functionality	0	1	2	3	4	5	6	7	8	9	10	M18	M24	M30	M36
KB Interface with DARC and back end	20	25	40	50	60	70	70	75	75	75	--	90	100	100	100
Improve Information Retrieval for Reporting	30	40	60	80	80	90	100	100	100	100	100	100	100	100	100
GUI (DS2 portal Integration)	10	20	30	50	50	60	70	80	85	90	--	100	100	100	100
Security Layer	10	20	30	50	50	60	70	80	80	80	--	100	100	100	100
Report Analysis interface	10	10	10	10	15	15	15	20	25	30	--	30	50	100	100
Deployment Documentation	20	20	20	20	30	30	30	30	30	35	--	60	100	100	100
UI for risk reports	10	10	10	20	20	30	40	40	45	50	--	40	20	100	100
DS2 Risk Knowledge base v1	40	50	50	60	70	70	70	80	80	90	--	90	100	100	100
report Analysis software	10	10	10	10	10	10	10	20	25	25	--	20	50	100	100
DS2 Risk Knowledge base v2	--	--	--	--	--	--	--	--	--	--	--	--	50	100	100

6.3.1.4 Activity

The focus on the first part of the project was to setup the interoperability with the other DS2 modules, (e.g. the Dash button and the DARC module).

In parallel an activity on sovereignty threats identification has been performed to support the improvement of the DS2 Knowledge base. The identification of the threats (cf. Section 4 above) is not a software feature, but it is a required activity to extend the Knowledge base. The Encoding of the threats is an ongoing software activity allows the inclusion of the new threats in the risk assessment.

Finally, while the UoS background Spyderisk already supports a form of reporting, it is not modular, and it presents challenges when trying to prove the correctness of the risk assessment. For this reason, effort has been dedicated to improving the information retrieval from the Knowledge base which is now modular and to some extent explainable to the end-user.

6.3.1.5 Use Case Validation Plan

	KEY:	Default install, UC = Validation	Specific To Case UC = Validation
Module ID	Precision Agriculture	GreenDeal	CityScape
SDS	All UC	All UC	All UC

6.3.2 DRM

6.3.2.1 Architecture Diagram

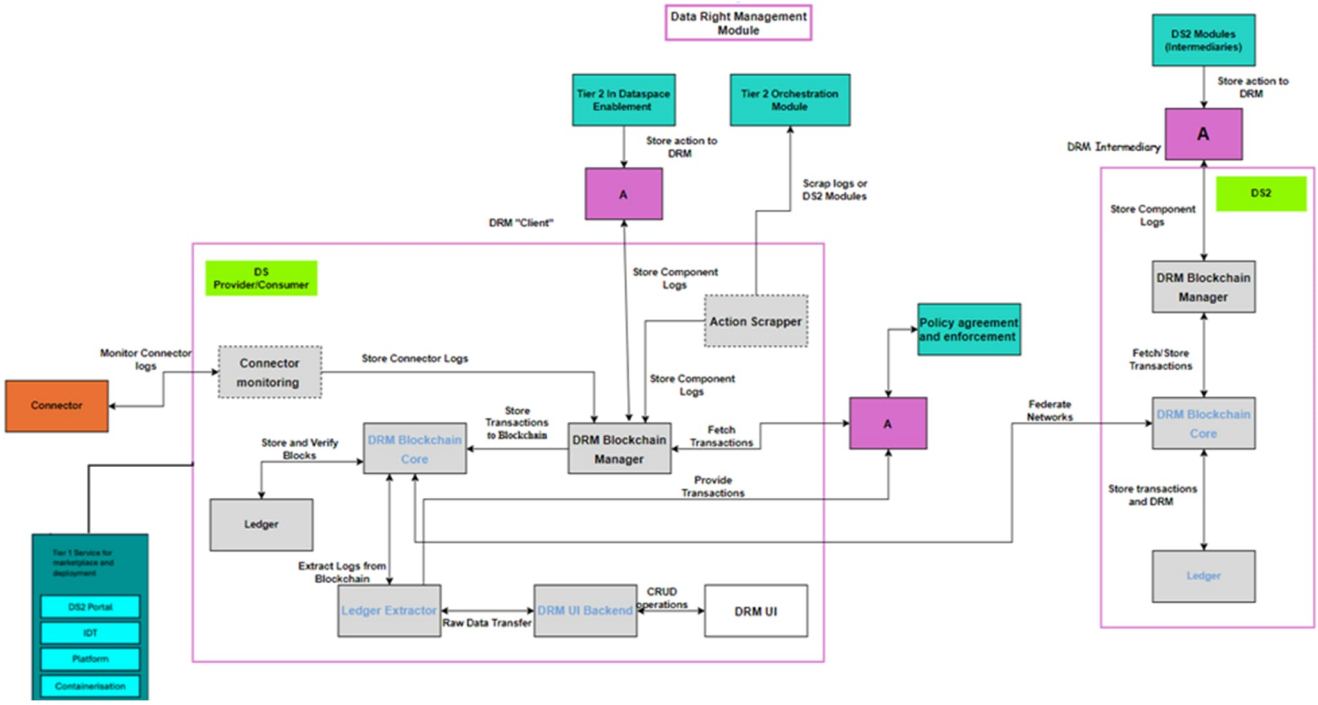


Figure 10: DRM architecture diagram

6.3.2.2 Sample Interface

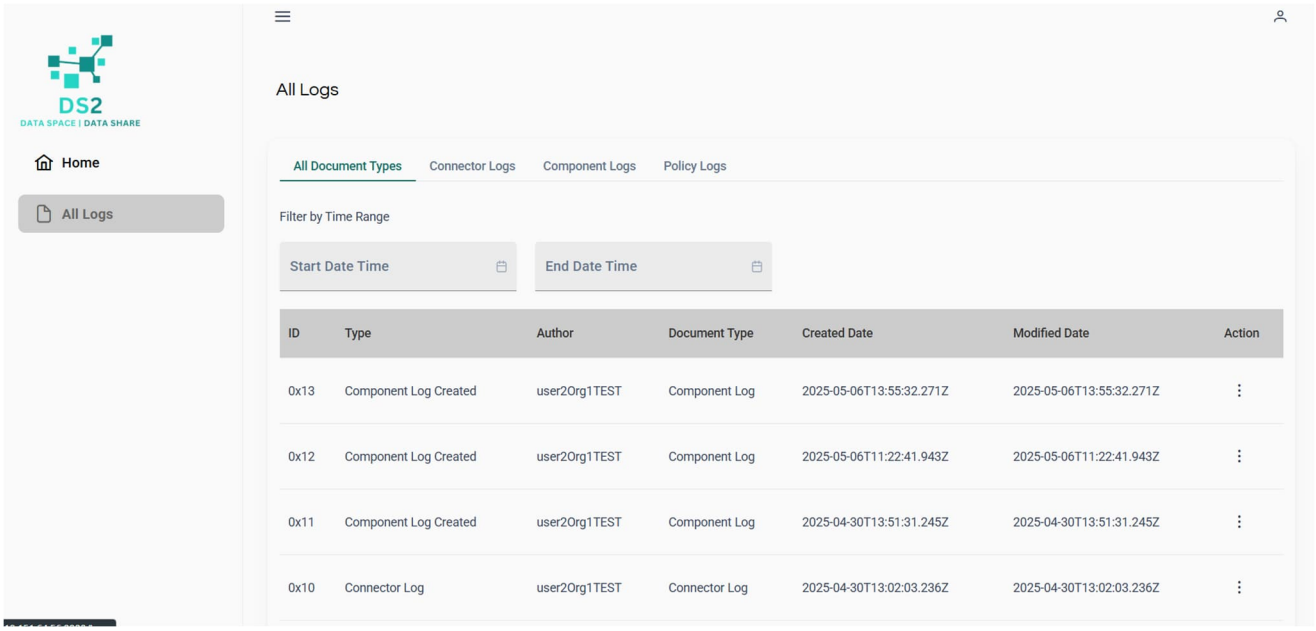


Figure 11: DSM user interface

6.3.2.3 Primary Feature Progress

Functionality	0	1	2	3	4	5	6	7	8	9	--	M18	M24	M30	M36
DRM Blockchain Core	20	30	30	30	40	40	40	40	60	60	-	60	100	100	100
DRM Blockchain Manager	20	35	35	40	50	60	60	65	70	70	-	70	100	100	100
Connector Monitoring	5	10	10	30	30	50	50	50	50	50	-	50	75	100	100
Ledger Extractor	25	35	35	35	35	50	55	60	60	60	-	60	80	100	100
User Interface	0	5	5	5	10	15	30	35	35	40	-	40	60	100	100
DRM User Interface Backend	10	20	20	20	20	20	35	40	40	40	-	40	70	100	100

6.3.2.4 Description

The software development of DRM module began with a focus on the DRM Blockchain Core and DRM Blockchain Manager. For this reason, custom chaincodes were created to record events from Dataspaces, while core features such as user and admin enrolment, as well as Create Read Update Delete (CRUD) operations, and secure data management were implemented early on in the implementation phase. Furthermore, a Docker-based setup was used for streamlining the deployment of DRM module, and Swagger documentation was added across the APIs to improve usability and integration of DRM module.

Once the stated sub-components became fully operational, additional sub-components were introduced/ implemented, namely the Connector Monitoring service, the DRM UI backend and the Ledger Extractor. These sub-components enabled the transformation and display of Ledger data, making it easier to trace and understand data-related activity from Dataspaces. This action was followed by localized testing in a Docker environment that helped to uncover and fix various bugs before moving forward.

Next, ATC's efforts shifted towards integration and preparation of module deployment. The DRM Manager was linked to the Connector Monitoring sub-component, and a basic extension for the Eclipse Dataspace Connector (EDC) was developed. At the same time, deployment of the DRM stack onto the DS2 Virtual Machine was prepared for upcoming demonstrations of the module.

As regards the work carried out on the UI, this was carried out in parallel with the work carried out in the back office of the module. Initially, requirements were specified, and mock-ups were developed to lead the user interface design process. Furthermore, Chaincodes were also refined to better support log handling for connectors, components, and policies.

With the Blockchain Core and Manager installed on the DRM virtual machine, their operational readiness was confirmed through testing. In addition, the work carried out on integration tasks expanded to include the Ledger Extractor, enabling the retrieval and formatting of ledger data for use in the UI.The first version of the DRM UI was completed shortly after, providing DRM users access to Connector Logs, Component Logs, and Policy Logs. Therefore, this first iteration of integration between the DRM'S backend, Ledger, and UI ensured a consistent and uninterrupted data flow that could be demonstrated anytime.

Moreover, a cross-component verification mechanism was introduced to confirm the identity of organizational users, supporting consistency and trust across the DRM Blockchain Manager, Ledger Extractor, UI backend, and DRM UI. Final testing across the entire system showed strong stability and interoperability, setting the groundwork for production deployment of the current iteration of the DRM module.

6.3.2.5 Use Case Validation Plan

	KEY:	Default install, UC = Validation	Specific To Case UC = Validation
Module ID	Precision Agriculture	GreenDeal	CityScape
DRM		UC2.2	UC1.1, UC1.2

7 KPI, RISKS, AND PRIMARY ISSUES

KPI Status:

KPI ID	Description	Status
KPI 3.1	Data owner respondents confirm that the modelling tools enables them to understand the risks of sharing data and to better control their sovereignty over it when shared.	Working with the UC is ongoing. The systems have not yet been modelled, so it is not yet possible to assess the KPI.
KPI 3.2	KPI 3.2: Practitioner respondents (eg data users) confirm that the regulatory guidance better enables them to understand and comply with the relevant legislation in complex data sharing situations.	Working with the UC is ongoing. A first workshop has confirmed that the proposed data-product based methodology is relevant and support a better understanding. This has not yet been quantified.
KPI 3.3	KPI 3.3. Data owner respondents confirm that the blockchain DRM system give them greater confidence in sharing data over its complex lifecycle.	Working with the UC is ongoing. The software have not yet been released to be used by the ned-users, so it is not yet possible to assess the KPI.
KPI 3.4	KPI 3.4. Respondents (primarily data owners) confirm that the barriers identified are correct and that the suggested means of lowering them are effective.	Working with the UC is ongoing. A first workshop has confirmed that the barriers and perceived threats identified in DS are consistent with the thinking of the stakeholders.

Table 3. KPI status overview.

Primary Risks (Those in DS2 Risk Register whose likelihood/impact are both > Low):

- (R790) Understanding and agreeing the complex policy landscape and pipeline and mapping to relevant DS2 modules
- (R360) Findings and the knowledge extracted interacting with the stakeholders need to be encoded in the Spyderisk knowledge base.
- (R380) The usage of blockchain technology should be simple enough to be usable in context where non-technical people are involved.

Primary Issues:

- N/A

8 CONCLUSION

This deliverable provides a comprehensive overview of the progress made in WP3 during the first 18 months of the project, with a particular emphasis on data governance, risk assessment and methodology development. WP3 is a multifaceted work package that encompasses both software and non-software components, such as identifying sovereignty threats and preparing a manual for creating data products. All technical documentation and source code are available in the DS2 GitHub repository.

The software modules included in WP3 have been described in detail; they are largely on track and are delivering the expected functionalities. Additionally, the deliverable presents a summary of the KPIs and includes a risk table outlining potential challenges identified during the early development stages.

WP3 adheres to the DS2 agile methodology, working in two-week sprints supported by regular coordination meetings and milestone reporting.

Looking forward, WP3 will continue to address governance issues, with a particular focus on policy management, advancing both methodological and software development activities. At this stage, WP3 remains on schedule and is already making significant contributions to the overall success of the project.

.

ANNEX A REFERENCES

Abbas, A.E., van Velzen, T., Ofe, H. et al. (2024). Beyond control over data: Conceptualizing data sovereignty from a social contract perspective. Electron Markets 34, 20. https://doi.org/10.1007/s12525-024-00695-2.
Agencia Española de Protección de Datos (AEPD) & European Agency for Cybersecurity (ENISA). (2024, April). Report of conclusions of AEPD-ENISA's event on data spaces — data spaces in EU: Synergies between data protection and data spaces, EU challenges and experiences of Spain. Available at: https://www.aepd.es/documento/report-conclusions-aepd-enisa.pdf (Accessed 29 May 2025).
Bartsch, J., Dehling, T., Lauf, F., Meister, S., Sunyaev, A. (2022). Let the Computer Say NO! The Neglected Potential of Policy Definition Languages for Data Sovereignty. In: Friedewald, M., Kreutzer, M., Hansen, M. (eds) Selbstbestimmung, Privatheit und Datenschutz . DuD-Fachbeiträge. Springer Vieweg, Wiesbaden. https://doi.org/10.1007/978-3-658-33306-5_22.
Brost, G.S., Huber, M., Weiß, M., Protsenko, M., Schütte, J., & Wessel, S. 2018. An Ecosystem and IoT Device Architecture for Building Trust in the Industrial Data Space. In Proceedings of the 4th ACM Workshop on Cyber-Physical System Security (CPSS '18). Association for Computing Machinery, New York, NY, USA, 39–50. https://doi.org/10.1145/3198458.3198459.
Carmichael, L., Taylor, S., Senior, S., Surridge, M., Erdogan, G., & Tverdal, S. (2025). Systematisation of Security Risk Knowledge Across Different Domains: A Case Study of Security Implications of Medical Devices. In Proceedings of the 11th International Conference on Information Systems Security and Privacy - Volume 1: ICISSP; ISBN 978-989-758-735-1; ISSN 2184-4356, SciTePress, pages 337-348. https://doi.org/10.5220/0013306100003899.
Data Spaces Support Centre (DSSC). (2024, October 11). Access & Usage Policies Enforcement. Data Spaces Blueprint v1.5. Available at: https://dssc.eu/space/bv15e/766069027/Access+&+Usage+Policies+Enforcement (Accessed 29 May 2025).
Data Spaces Support Centre (DSSC). (2024, May 5). Glossary v3. Available at: https://docs.google.com/document/d/15x6WHHGSoG4ZuXQw8u3AinpJrgbydriL/edit
Gil, G., Arnaiz, A., Higuero, M., & Diez, F.J. (2023, February). Assessment Framework for the Identification and Evaluation of Main Features for Distributed Usage Control Solutions. ACM Transactions on Privacy and Security. 26(1): 10, 1-28. https://doi.org/10.1145/3561511.
Hellmeier, M., Pampus, J., Qarawlus, H., & Howar, F. 2023. Implementing Data Sovereignty: Requirements & Challenges from Practice. In Proceedings of the 18th International Conference on Availability, Reliability and Security (ARES '23). Association for Computing Machinery, New York, NY, USA, Article 143, 1–9. https://doi.org/10.1145/3600160.3604995.
Kelbert, F., & Pretschner, A. 2018. Data Usage Control for Distributed Systems. ACM Trans. Priv. Secur. 21, 3, Article 12 (August 2018), 32 pages. https://doi.org/10.1145/3183342.
Lazaro, C., & Le Métayer, D. (2015, June). Control over Personal Data: True Remedy or Fairy Tale? SCRIPTed 12(1), 3. https://doi.org/10.2966/scrip.120115.3.
Lohmöller, J., Pennekamp, J., Matzutt, R., & Wehrle, K. 2022. On the Need for Strong Sovereignty in Data Ecosystems. In Proceedings of the 1st International Workshop on Data Ecosystems (DEco '22). Vol. 3306. CEUR-WS, Sydney, Australia. Available at: https://ceur-ws.org/Vol-3306/paper6.pdf (Accessed 25 March 2025).
Pitkänen, O., Turpeinen, M., & Lähteenoja, V. (2025, February 6). Rulebook model for a fair data economy (version 3.0): The rulebook model 3.0 is a guide for creators of data spaces. Sitra. Available at: https://www.sitra.fi/en/publications/rulebook-for-a-fair-data-economy/ (Accessed 29 May 2025).
Steinbuss, S. et al. (2021). Usage Control in the International Data Spaces. International Data Spaces Association. https://doi.org/10.5281/zenodo.5675884.
W3C. (2018). Open Digital Rights Language (ODRL): ODRL Information Model 2.2. W3C Recommendation 15 February 2018. Available at: https://www.w3.org/TR/odrl-model/ (Accessed 29 May 2025).

W3C. (2024, August). Data Catalog Vocabulary Version 3. Available at: https://www.w3.org/TR/vocab-dcat-3/
Zrenner, J., Möller, F.O., Jung, C., Eitel, A., & Otto, B. (2019). Usage control architecture options for data sovereignty in business ecosystems. Journal of Enterprise Information Management. 2019;32(3):477-495. https://doi.org/10.1108/JEIM-03-2018-0058.

ANNEX B – SURVEY DETAILED ANALYSIS

This annex presents the detailed analysis of the survey and the follow up workshop about data sovereignty and barriers to data sharing.

Current Attitudes and Understanding

Data Sharing

The survey results indicated that there is a varied understanding of what inter-organisational data sharing actually is. Just 40% correctly identified it as "the on-going exchange of business-relevant data between collaborative partners," aligning with the accepted definition (e.g. Malysh et. al., 2024[2]). In contrast, over a third (36%) associated it solely with sharing research data, which is a subset of inter-organisational data sharing but not its full scope. While almost a quarter (24%) mistakenly linked it to sharing consumer personal information with third parties, which is more related to B2C data practices. These findings suggest that there is significant confusion about the concept of data sharing even among experts.

Equally, the survey results revealed varying levels of understanding among data sharing professionals regarding key legal aspects. Firstly, there was a patchy understanding of the main components of Data Sharing Agreements (DSAs). The highest level of comprehension is in the area of purpose and scope, with 66% of respondents claiming good understanding. Data types and formats, data ownership, and access limitations are also moderately well understood by 55% of respondents. However, there are significant knowledge gaps in other crucial areas. For example, only 41% report a good grasp of security and privacy aspects, as well as roles and responsibilities. More concerning is that just 31% feel confident in their understanding of intellectual property issues, compliance and governance, and agreement duration and termination clauses.

Secondly, this patchy understanding extended to the knowledge of EU and UK regulations (e.g. The Data Governance Act (2023); The Data Act (2023); GDPR (2016); UK GDPR (2018)…etc). Just under a quarter of respondents were confident that they were familiar with the details of all the relevant legislation (24%). In contrast, 60% of respondents were only familiar with the key clauses or broad principles of some of the legislation, while 16% were not familiar with any of them. Despite this relative lack of understanding of the details of the legislation, over half of the respondents (56%) felt that existing regulations were either definitely not good enough (16%), or were only sometimes good enough (40%), to protect their business interests when undertaking data sharing. In contrast, only 40% felt the legislation was mostly (24%) or definitely (16%) good enough to protect their interests.

These results suggest a need for a DataSpace to support stakeholders to improve their understanding of the various different laws and regulations that govern the domain, to overcome the somewhat sceptical perception among data experts that they are not sufficiently fit for the purpose of adequately protecting self-interest. At a regulatory level beyond the DataSpace, there may also be a need for on-going refinement of legal frameworks in order to address these concerns, because rightly or wrongly, these concerns are likely to constitute a significant factor when considering whether or not to begin data sharing activities.

[2] Malysh, K., Ahmed, T., Linåker, J. and Runeson, P., 2024, August. Inter-Organizational Data Sharing Processes-An Exploratory Analysis of Incentives and Challenges. In *2024 50th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)* (pp. 80-87). IEEE.

Finally, only a fifth of all respondents felt that they and their organisation had a clear understanding of the monetary value of the data they were sharing (21%). The remaining 79% either had no clear understanding (46%) or no understanding at all (33%) of the value of the data being shared. This is important for several reasons. Firstly, without a clear assessment of the value of the data it becomes difficult to quantify the potential business benefits of data sharing and to allocate appropriate resources that justify the investment costs and strategic decisions. Secondly, it becomes harder to negotiate equitable DSAs because the value of the data can provide a quantifiable foundation for those negotiations. Thirdly, it highlights a potential tension within a DataSpace between open data and data that is provided with an associated cost, and the impact that may (or may not) have on the business model of the DataSpace and its longer-term financial sustainability. Finally, risk assessment is also impacted, as it is more difficult to evaluate the potential risk of, for example, data breaches or misuse when there is no clear notion of the value of the data that has been lost or misused.

Against this backdrop of a relative lack of understanding regarding what data sharing actually is; the key components of a Data Sharing Agreement; the details of existing legislation; and the monetary value of data assets, the survey also revealed that a meaningful minority (20%) consider the transition to a data sharing ecology too complex and resource-intensive.

Despite all this, there remains a largely positive attitude towards inter-organisational data sharing among European and UK data experts. A significant 72% see it as beneficial, with 36% viewing it as a source of competitive advantage and another 36% considering it worth the transition despite risks. Only 8% perceive it as primarily risky or high-risk. This suggests that while most professionals clearly recognise the value of data sharing, there may be some challenges and knowledge gaps when it comes to implementation.

Willingness to share data with partners in different industry sectors were very equal, where the industry sector itself does not impact that willingness. However, what does seem to matter is altruistic data sharing outside partner networks for the overall benefit of the industry sector. Nearly two thirds (60%) were not willing to share data with state agencies, umbrella organisations, industrial hubs…etc, while just 8% were extremely comfortable doing so. This indicates that indirect benefits, such as sector growth, are less motivational than direct benefits to the immediate organisation.

The results suggest that perhaps the overall positive attitude to inter-organisational data sharing is somewhat tempered by an awareness of challenges, including legal weaknesses, complexity and a limited set of preferences for with whom to share data. Perhaps as a result of this, the responses indicated that there was an exactly 50 / 50 split between those who are in some way likely to increase the amount of inter-organisational data sharing in the future, and those who are undecided or unlikely to do so.

Data Sovereignty

As with the lack of consensus regarding the definition of inter-organisational data sharing, the survey results revealed a further range of understanding of what data sovereignty actually means. The responses highlighted four distinct interpretations, each emphasizing different aspects of data governance and control.

The most prevalent view, held by 38% of respondents, agreed with the survey definition that data sovereignty is "the ability of the data owner to decide how to share and use its data". This perspective emphasises individual or organizational control over data, aligning with principles of data ownership and self-determination. Closely following this view, 33% of respondents associated data sovereignty with "a set of core principles and actions with the main objective of establishing trust in data collection and use". This view emphasises the ethical and trust-building aspects of data sharing, suggesting a strong awareness of the importance of responsible data practices in establishing trust relationships. Again, closely behind this view, a quarter of the respondents (25%) understood data sovereignty as "the idea that data is geolocated and therefore subject to the ethics, laws and regulations of a particular nation or jurisdiction". This aligns more closely with the traditional definition of data sovereignty found in academic and legal contexts, where the geographical and jurisdictional aspects of data sharing and the role of national laws in regulating data sharing

activities is recognised. Finally, only 4% of respondents viewed data sovereignty as "a spectrum of approaches adopted by different nation states to control data generated in or passing through national internet infrastructure". This low percentage suggests that fewer professionals associate data sovereignty with specific national strategies for data control.

The diversity in responses indicates a lack of consensus on the definition of data sovereignty among data professionals. This fragmented understanding could lead to challenges for DataSpaces in implementing consistent data sharing practices across organisations and borders. The results also suggest a shift in perception from purely legal and geographical definitions towards more user-centric and trust-based interpretations of data sovereignty. Equally, the low percentage associating data sovereignty with national control strategies might indicate a gap between policy-level discussions and practical understanding among professionals. It could also suggest that professionals are more focused on operational aspects of data sharing rather than broader geopolitical considerations.

These varied interpretations highlight the complex and multifaceted nature of data sovereignty in today's digital landscape. They underscore the need for a clear understanding within a DataSpace of data sovereignty concepts, perhaps detailed within the DataSpace Rulebook and clearly explained during the on-boarding process, to ensure consistent and effective implementation of data sharing practices.

In contrast, there was a strong consensus among data professionals that data sovereignty complexities and a lack of clarity significantly impact their willingness to share data. A substantial majority (75%) agreed to some extent with the statement: “Data sovereignty complexities significantly impact my willingness to share data” (29% strongly agreed, 46% somewhat agreed). Equally, an overwhelming majority (88%) agreed with the statement: “A lack of clarity over what data sovereignty really means for my organisation significantly impacts my willingness to share date” (25% strongly agreeing, 63% somewhat agreeing). This high level of consensus indicates that data sovereignty issues are a major obstacle to data sharing in organisations. The complexities and perceived lack of clarity associated with data sovereignty appear to create hesitation and caution among professionals when considering data sharing initiatives, which highlights the need for DataSpaces to support stakeholders in better understanding data sovereignty regulations and practices. The importance of this, coupled with the confusion around what data sovereignty actually is, meant that data sovereignty was identified as an important area to explore in more depth during the Deep Dive Workshop.

Current Practices

The survey results have shed light on current data sharing practices. Firstly, 67% of respondents reported that their organisation does currently share data with another organisation, mainly as a data provider (31%) or both a data provider and consumer (28%). The infrastructure for data sharing is present (68%), with cloud-based platforms or services being the most common. Data sharing processes are also fairly widely established, with 60% reporting having one or more of DSAs, protocols, compliance checks or secure transfer methods in place.

Despite this initially positive set of current practices, there are also areas of weakness. More than half or organisations (58%) do not have, nor have ever had, any data sharing training programmes for senior management teams and the relevant staff. Perhaps as a result, 62% of respondents reported either not knowing (50%) or not having (12%) clear data governance and ownership structures within their organisations. Equally, there is a lack of knowledge concerning what data licenses are currently being used, with over half (54%) not knowing what licenses are being used. Also of note was that where the licenses were known, many of them were open licenses as opposed to commercial licenses.

The survey indicated that there was a fairly even mix of data types that are currently shared, with company demographics (29%), other – including weather, geospatial, energy, agricultural, and research data (29%), production data (25%), and customer data (25%) being the most prevalent. In contrast, strategic data (12%)

and sales data (4%) are the least shared data types. This suggests that there remains a concern about sharing data which may impact competitive advantage.

In general, the simpler data formats are most commonly used for the datasets that are being shared, with text (46%), PDF (46%), and Excel (42%) files proving most popular. XML (25%) and SQL or other database specific formats (12%) are the least common file types used.

Turing to current risk assessment and security practices, just over half (52%) reported conduction often (28%) or always (24%) thorough risk assessment before undertaking data sharing. However, over a third of organisations sometimes (16%) or rarely (20%) carry out a risk assessment before data sharing, which indicates quite diverse attitudes and practices concerning risk management. This suggests that there is considerable value in a DataSpace providing support for organisations in conducting cybersecurity risk assessment.

Perhaps linked to this relative paucity in risk assessment, the survey reveals that by far the most commonly deployed security control is to apply data access controls to datasets (76%). However, no other type of security control is routinely used by more than half of the organisations, with controls such as the use of blockchain/DLT tools (24%), regular cybersecurity audits (24%) and data pseudonymisation (16%) being the least deployed. Of note, fewer than half (48%) report providing employee training in cybersecurity, which when coupled with the lack of training provided in data sharing generally, highlights a need for comprehensive training programmes in this domain.

Aligned with this lack of security controls (and training), access restrictions are by far the most common restrictions applied to datasets that are shared (70%). Equally though, no other type of restriction is used in more than 35% of cases, with storage restrictions (17%) and non-aggregation/enrichment restrictions (13%) being the least common. Indeed, 13% of all datasets have no restrictions at all placed on them since, most likely, they are public in nature. These restrictions and security controls will impact the Data Access and Usage Policies attached to each dataset that is made available through a DataSpace, which in turn will affect the terms of the DSAs, so are important to understand. Furthermore, the inclusion of data restrictions resonated with the earlier findings concerning the use of open licenses, where there may be a tension between open licenses and restricted datasets, as well as between open/free and paid-for datasets. As a result, this tension was deemed worthy of further exploration during the Deep Dive Workshop with the Use Case partners that followed the survey.

Perceived Challenges of Data Sharing

The survey next turned to an exploration of the perceived challenges facing stakeholders when undertaking inter-organisational data sharing in four main areas – strategic, operational, technical and network/partnership challenges.

The three most important strategic challenges that an effective DataSpace should support users in overcoming are the complexities in establishing trust relationships / collaboration networks (64%); the potential for data misuse by recipients (64%); and the complexities in placing a monetary value on the data (56%). The first two of these have at their core the idea of trust, and consequently a DataSpace must place trust mechanisms at its core, and must itself be trustworthy for its participants (clearly this is the raison d'etre for a DataSpace in the first place).

Alongside that, a DataSpace should also support its users in addressing the three most important operational challenges of the cost (time / effort / financial) of establishing data sharing processes (64%); the complexities of negotiating and finalising DSAs (60%); and the lack of knowledge, skills and expertise within the organisation regarding all aspects of data sharing (48%). Again, DataSpaces are designed to help reduce the time and effort costs of data sharing, which can include supporting DSA activities with the provision of templates and guidelines…etc. Perhaps less common, but clearly of value, would also be the provision of training resources on aspects of data sharing within the DataSpace service provision.

Technically, the survey indicates that the three most important challenges that a DataSpace should help overcome are the transparency of data collection, processing and use (56%); the quality, integrity, consistency, and timeliness of the data (52%); and the ability to deploy strict access controls (40%). This helps to inform the design of the metadata and information that should be attached to a dataset at the point that it is made available through the DataSpace, to ensure that transparency and informative dataset descriptions are appropriately attached to a dataset.

Finally, in relation to the challenges faced when working with others in the data sharing domain, either in a network or as direct partners, the three most important challenges to help address are the need for clear governance and ownership structures (64%); an assessment of the reputation and reliability of partners (52%); and the ability to establish shared risks and mutual value co-creation (40%). As before, one of the primary functions of a DataSpace is to ensure that its members are reliably who they say they are, and to support effective data sharing structures. However, the notion of systems, tools, or methods for mutually sharing risk and value creation may be an area worthy of further thought and research. A deeper exploration of some of these challenges were undertaken in the Deep Dive Workshop that followed the survey.

Deep Dive Workshop

The survey results raised three main areas that it was felt merited some further and deeper exploration – Data Sovereignty; Data Licensing and Restrictions; Overcoming Datasharing Challenges. Initially a follow-up spreadsheet was disseminated to all WP7 partners in M15 seeking further detail and information on various aspects related to these topics as they specifically relate to the three DS2 use cases. The responses were then briefly presented and lightly discussed during the WP7 bi-weekly meeting in preparation for the full Deep Dive session in M16. To ensure good attendance and enable sufficient preparation a detailed flyer for the workshop was provided to the WP7 lead, who then disseminated it across the work package. As a result in April 2025 at an extended WP7 full team meeting involving representatives of all three use cases, including both technical and user partners (18 participants total), the Deep Dive Workshop was delivered.

For each of the three topics, participants were provided with a short reminder presentation of the relevant survey and spreadsheet results and clarifications over the issues arising from the survey. Then participants were encouraged to consider how those issues impact on and/or would be addressed in each of the use cases. Next, participants were invited to add their thinking to a Miro board that had been specifically set up for the workshop. The board provided space for the representatives of each use case to present their opinions and ideas within use-case-specific spaces on the board. After this, a whole group guided discussion was held (and recorded) on each of the topics during which the most pertinent of the Miro board comments were explored in greater depth.

Data Sovereignty Workshop Discussion

Participants were referred to the SITRA rulebook for reference after a clarification of what data sovereignty actually means had been presented. Participants were then asked to consider the 'order of priority' between the DataSpace Rulebook (the documentation for the DataSpace governance framework, including membership requirements, obligations and general terms and conditions) and the Data Policies attached to individual Data Products within their specific use case. There was a universal agreement that all use case DataSpaces should have, as far as possible, a similar DataSpace Rulebook (most likely based on the SITRA rulebook).

However, considerable discussion and disagreement occurred when discussing the priority of the terms and conditions (T&Cs) laid out in the Rulebook and those laid out in the Data Product Policy (and subsequent DSA). Some use cases felt that the Data Policy T&Cs should always take precedence over the rules of the DataSpace, as advised in the SITRA Rulebook, while others put forward examples where this should not be the case (such

as a DataSpace rule that data should not be shared with entities from certain countries versus a Policy condition that the data can be freely shared with any interested entity). No clear consensus across use cases was reached on this matter and it remains a subject for further discussion.

There were also calls for systems that 'translate' Policy T&Cs for non-expert / non-tech users; tools which can support the creation of Data Product Policies (such as PCR); and processes by which access to datasets is no longer permitted for entities which have left the DataSpace (i.e. are no longer members), even if the dataset has been previously shared.

Data Licensing and Restrictions Workshop Discussion

Again, participants were referred to the relevant sections of the SITRA rulebook for reference after a brief presentation on the tension between open licensing and restricted data had been delivered. There was a clear link to the previous topic, as many data restrictions are defined in Data Access and Usage Policies (the constituent parts of a Data Product Policy).

Use cases called for templates that support different types of restrictions and licensing, and authentication mechanisms to verify the role of every DataSpace participant. One use case reported widespread use of public data, which has to be provided under open licenses, but which also logged who was accessing that data, and this concept of data access traceability was echoed in a second use case. The third use case was able to define access restrictions relating to the role of DataSpace participants, for example, certain roles should only be able to access certain datasets and not other similar datasets – although on discussion it was not clear how this was intended to be practically implemented. This focus on access restrictions reinforced the survey findings, where cybersecurity controls (and associated risk assessment) is most frequently narrowly understood as access controls, and further highlighted the need for a broader appreciation of the importance of cybersecurity and risk assessment as a method of generating trust between entities, rather than just as a method of controlling and restricting actions.

An in-depth discussion also took place concerning the importance of having methods, tools and/or processes to monitor compliance with the agreed Rules, Data Product Policies (and DSAs). There was also a discussion of the process by which separate DataSpaces with different Rules can negotiate a set of jointly agreed Rules thereby enabling data to be shared across DataSpaces effectively. Effective monitoring of Rules and Policies and inter-DataSpace negotiation were universally recognised as critical to the effective implementation of the DataSpaces in all three use cases.

Overcoming Data Sharing Challenges Workshop Discussion

Again, participants were referred to the relevant sections of the SITRA rulebook for reference after a presentation of the key challenges that the survey had illuminated. In relation to overcoming the strategic challenge of putting a monetary value on the data, use cases suggested that there could be a DataSpace Rule regarding curated versus raw data, with curated data being able to have a price attached but raw data remaining free (or open).

This led to a broader discussion of business models. The use cases reported that developing the most effective business model for monetising the DataSpace for long-term sustainability were on-going and active discussions, with ideas being explored concerning a price for the data itself; a reliance on public funding (as the datasets being made available were required to be open); an annual subscription fee for membership of the DataSpace; and/or taking a percentage of the payments made by DataSpace members for services and tools.

In addition, one use case identified the importance of the DataSpace Rulebook in overcoming the strategic and network/partnership challenge of the complexity of establishing trust relationships. Furthermore,

concerning the complexity of negotiating DSAs, another use case also called for contact templates to be available.

Summary

The survey and Deep Dive Workshop indicate a broadly positive attitude to data sharing and a reasonable amount of data sharing activity, infrastructure and processes already in place. Cross-sector data sharing is considered desirable, although altruistic data sharing with state agencies, umbrella organisations and industry hubs to grow the sector is less favoured. There is also a broad perception that data sharing is complex and has multiple challenges, which results in just half of all stakeholders expressing a likelihood to increase data sharing activities in the future.

The results indicated that a number of important specific gaps were present in the understanding or implementation of:

7. data sovereignty
8. legislation
9. data sharing agreements
10. licensing vs restrictions
11. security / risk assessment
12. training

In addition, a range of important strategic, operational, technical and inter-organisational challenges to data sharing were identified. The most significant of which are:

5. Strategic – establishing trust; potential for misuse; establishing monetary value of data
6. Operational – associated costs; complexities of Data Sharing Agreements; lack of skills
7. Technical – transparency; quality; access controls
8. Inter-organisational – governance and ownership; trustworthy partners; access controls

A number of important topics also remain open for future discussion, including business models / monetisation, and the relationship between Data Space Governance Frameworks (aka Rulebooks) and the Data Policies found in Data Products.

The results of this research are recommended to inform WP3 and wider DS2 development in the ways outlined in the table below.

Topic	Finding	Action
Data Sovereignty	Lack of agreement over definition and understanding of the term	Decide on a comprehensive definition, incorporating both the jurisdictional and control aspects of data sovereignty, and use it to inform Knowledgebase extension (WP3)
Data Sovereignty	Data Space Governance Frameworks (aka Rulebook) should be as similar as possible in order to facilitate sharing across data spaces	All Use Cases to base their Data Space Governance Framework on the SITRA rulebook reference (WP7)
Data Sovereignty	Lack of agreement concerning the priority between the Data Space Governance Framework and the Data Policies found in Data Products	Active design consideration for the PCR, Portal/IDM and PAE modules (WPs 4 & 6)
Legislation	Lack of in-depth understanding of the various pieces of relevant legislation	Explore the possibility of inclusion of legislative compliance within Knowledgebase extension and its potential for inclusion in risk assessment (WP3)

Data Sharing Agreements (DSAs)	DSAs are too complex and many aspects are not well understood	Use the least understood aspects to focus development efforts in the PCR module to ensure adequate simplification and automation in these areas (WP6)
Licensing & Restrictions	Tension between datasets that are required (or chosen) to be Open Access and the data usage and access restrictions that may be present in the Data Space Governance Framework and the Data Policies found in Data Products	Inform the development of the Data Product description process and content (WP3) Ensure PCR module accounts sufficiently for open datasets (WP6) Use the finding to inform on-going discussions concerning Data Space business models and sustainability (WP1 & WP3)
Security & Risk Assessment	Risk assessments only sporadically conducted Main focus of security controls is only on Access Controls Despite this focus, controlling Access remains an important concern when sharing with another organisation	Ensure the SDS module(s) supports more frequent risk assessment by simplifying and automating the process (WP3) Ensure the Knowledgebase extension accounts for the complexities of Access Controls as defined in multiple places within and between Data Spaces (see 4.1 below) (WP3) Ensure the DS2 enabled risk assessments provide information about other important controls and control strategies that should be implemented beyond Access Controls (WP3)
Monetary Value of Data	Placing a monetary value on a dataset (that is not open by default) is extremely difficult	Understand how the 'value' of a dataset can be estimated sufficiently well that it can be accounted for in the Knowledgebase extension to contribute to the risk assessment calculations (WP3) Use the finding to inform on-going discussions concerning Data Space business models and sustainability (WP1 & WP3)

ANNEX C – A MANUAL FOR BUILDING DATA PRODUCTS FOR DATA SPACES

This section is a methodology on how to build data products for data spaces:

- STEP 1: Understanding data products in the context of data spaces
- STEP 2: Specifying intermediate services and data products
- STEP 3: Identifying data product flows

STEP 1: Understanding data products in the context of data spaces

Background information on data products

Detailed information on data products in the data spaces can be found as part of the Data Spaces Support Centre's (DSSC) building block Data Space Offering: https://dssc.eu/space/BVE2/1071253231/Data+Space+Offering. This document focuses and summarizes the information on the data products from said building block.

What is a data product:

A Data product is a data sharing unit, packaging data and metadata, and any associated license terms. The data product may include, for example, the data products' allowed purposes of use, quality and other requirements the data product fulfils, access and control rights, pricing and billing information, etc.

Data products are collected under offerings containing data product(s), service(s), and the offering description that provides all the information needed for a potential consumer to make a decision whether to consume the data product(s) and/or the service(s) or not.

Whilst the content of data products may vary, a data product typically includes:

- The data,
- The description of the data,
- Tts allowed purposes of use,
- Quality, format, frequency, duration and other requirements the data product fulfills,
- Access and control rights (e.g., attribution, Intellectual Property Rights, liabilities, geographical limitations, usability for training AI models),
- Delivery options (e.g., APIs, SMTP, web interface, mapping tools),
- Information about data provenance and lineage,
- Pricing and billing information,
- Other information (e.g., ethical considerations),
- Metadata describing all these above.

Why are data products needed:

Perceiving data as a product will, for many organisations, require a change of mindset: the idea of producing data with reuse in mind.

From the data product owner's perspective, it is beneficial that data products containing the same data can be delivered to multiple use cases. Therefore, it is recommended that data product owners consider developing multiple data products with various options of the associated information (e.g., different license terms, delivery options, commercial and technical requirements).

Practical tips for the generation of data products:

1. *Identification of the need and purpose of the data product in the context of a use case (s).*

Data products should satisfy consumers' needs with reuse in mind. They should have a "meaning," and a purpose of use. Therefore, when creating data products, it is recommended to start on the consumers' side to clarify the expected value of the data product. A key consideration is the type of value the data products deliver, which can be both monetary and non-monetary. Another consideration is the data space(s) and the use case(s) through which the data product will provide value to the data product consumers.

2. *Developing data products.*
   Developing data products might require developing several data product candidates and then choosing the one that gets the best score after evaluation (see the next step). The data product candidate development consists of the following steps:
   - When the data product owner defines the purpose of a data product, there is a need to identify the data source(s) that provide the data ingredients for the data products.
   - All licensing terms, commercial aspects, delivery options, consumption pattern, information about data provenance and lineage, requirements and various policies needs to be decided.
   - The necessary metadata needs to be created using a suitable standard.
   - Finally, all required information needs to be bundled into a consumable form, using a suitable standard.
3. *Quality assurance, evaluation and validation of the data product.* The data product needs testing, as well as validation that it fulfills its requirements and intended purposes.
4. *Publishing the data product via catalogue services.* As a last step, the offering should be made available for the potential data product consumers of the data space. efer to the Publication and Discovery building block for the technical details.
5. *Maintenance and supporting services.* The data product provider is responsible for maintaining and supporting services throughout the whole lifecycle of the data product.

Example of a data product:
https://data.marine.copernicus.eu/product/ANTARCTIC_OMI_SI_extent_obs/description

- Description
- Notifications
- Data access
- Contact

DOCUMENTATION

- User Manual
- Quality Information Document
- Roadmap
- Licence
- How to cite

DOI

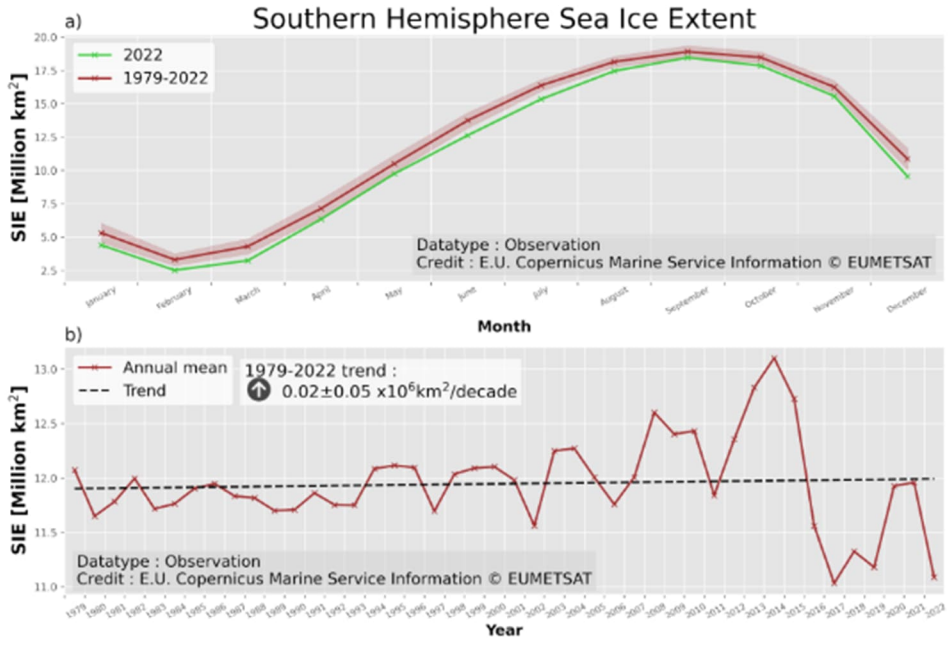
10.48670/moi-00187

Overview

DEFINITION

Sea Ice Extent (SIE) is defined as the area covered by sufficient sea ice, that is the area of ocean having more than 15% Sea Ice Concentration (SIC). SIC is the fractional area of ocean surface that is covered with sea ice. SIC is computed from Passive Microwave satellite observations since 1979. SIE is often reported with units of 106 km2 (millions square kilometers). The change in sea ice extent (trend) is expressed in millions of km squared per decade (106 km2/decade). In addition, trends are expressed relative to the 1979-2022 period in % per decade. These trends are calculated (i) from the annual mean values; (ii) from the September values (winter ice loss); (iii) from February values (summer ice loss). The annual mean trend is reported on the key figure, the September (maximum extent) and February (minimum extent) values are reported in the text below. SIE includes all sea ice, except for lake and river ice. ...

Read more



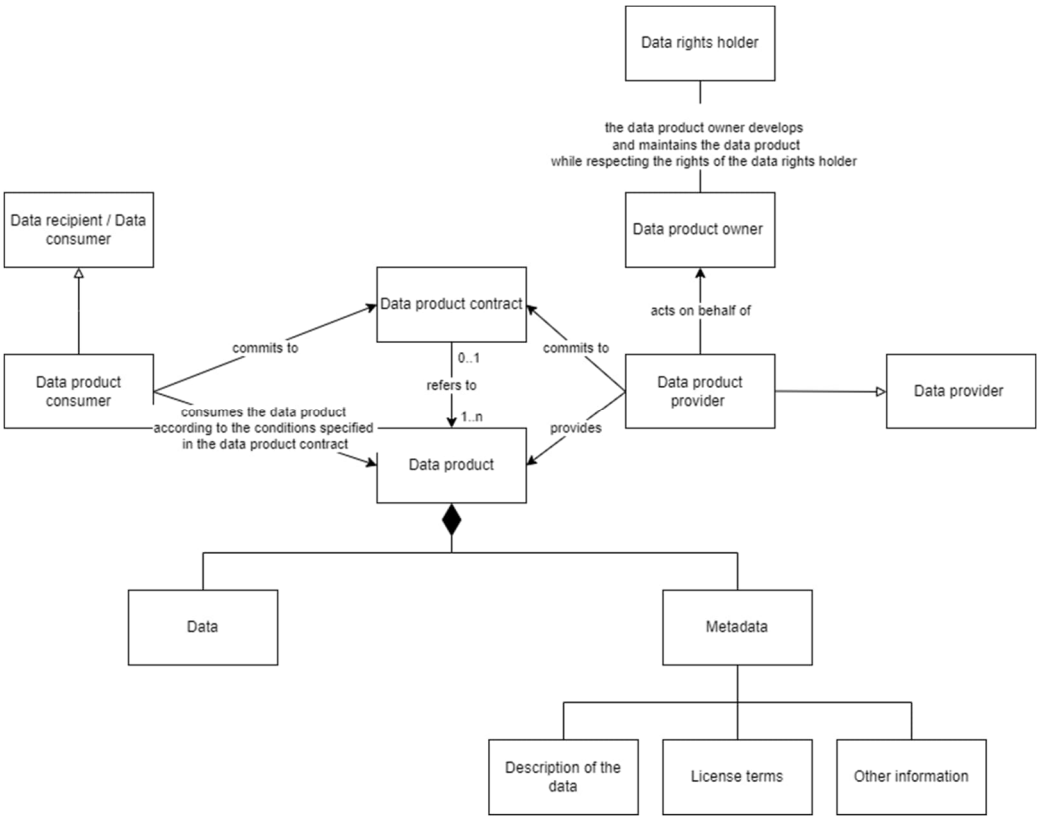
Download image

Classification

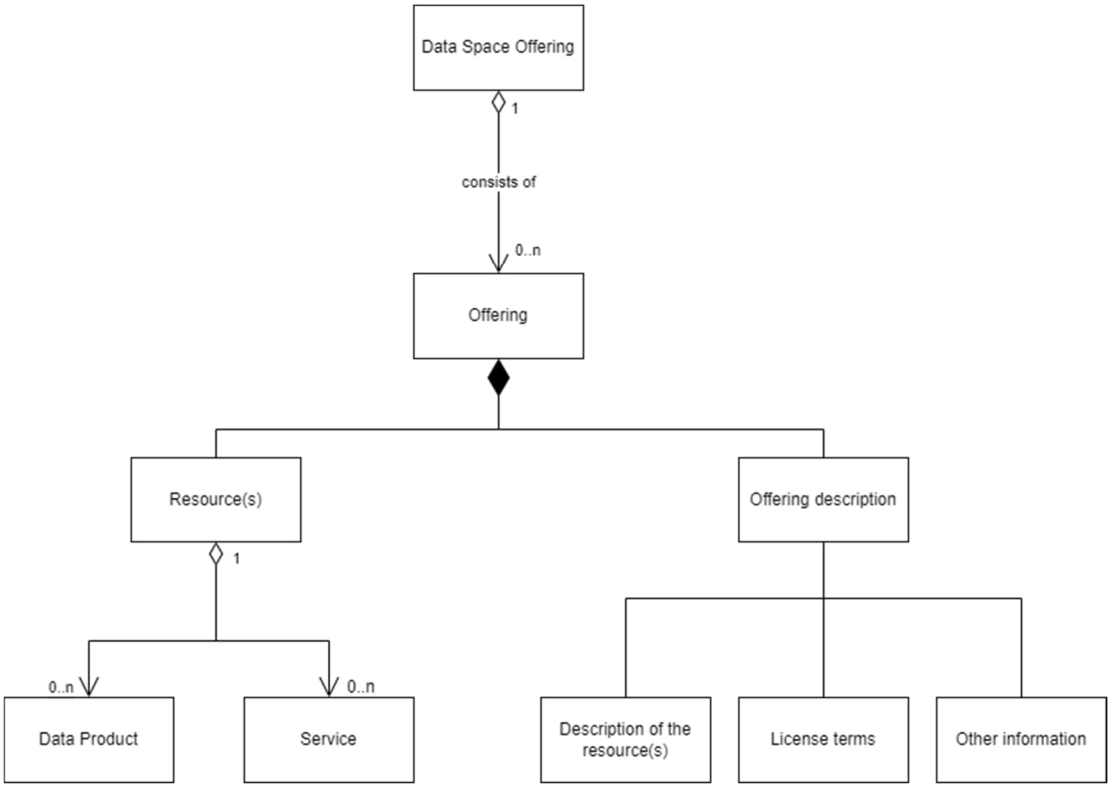
Full name	Antarctic Monthly Sea Ice Extent from Observations Reprocessing
Product ID	ANTARCTIC_OMI_SI_extent_obs
Source	Satellite observations
Spatial extent	Global Ocean · Antarctic Ocean · Lat -90° to 0° · Lon -180° to 180°
Temporal extent	1 Jan 1979 to 31 Dec 2022
Temporal resolution	Monthly
Variables	Sea ice extent
Indicator family	Sea ice change
Feature type	Point series
Blue markets	Polar Environment Monitoring · Science & Innovation
Projection	WGS 84 (EPSG:4326)
Update frequency	Annually
Format	NetCDF-4
Originating centre	MET Norway
Last metadata update	30 November 2023

Screenshot of the example "Antarctic Monthly Sea Ice Extent from Observations Reprocessing. *E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS).* DOI: 10.48670/moi-00187 (Accessed on 18-06-2025)")

Data product and related concepts:



Data space offering and related concepts:



STEP 2: Specifying intermediate services and data products

[*Please, fill in your use case information*]

TEMPLATE FOR DATA SPACE USE CASES TO DEFINE VALUE CREATION SERVICES AND DATA PRODUCTS
Data space use case name:
[*Your use case name*]

1. BUSINESS CASE

1A. Business service aimed for the final end customer
Name of the business service:
[*Your business service identification*]

1B. End customer
Identification of the final end customer of the service:

- Who would be paying for the business service?
- What customer segment does the final end customer belong to?
- Who is the final user of the service generated by the data sharing?

[*Your final end customer identification*]

1C. Value proposition
Identification of the value proposition for the final end customer:

- What is the value the final user gets from the data sharing?

[*your value proposition towards the end customer*]

2. KEY ACTIVITIES NEEDED TO DELIVER THE VALUE PROPOSITION
2A. Identification of needed value creation services (intermediate services) to produce data products:
What kind of data related activities (Value Creation Services) are needed in order to deliver the value proposition to the final user?
Categories: [*Choose relevant, mark internal activities with I and external activities with E*]

	Core services		Data visualization
			Data quality management and assessment
			Technical enablers for compliance
			Security
			Monitoring and reporting
	Data handling services		Data selection
			Data extraction
			Data combination
			Data packaging
			Data processing

			Data transformation
			Data delivery
			Data interpretation and reuse
	Value added services		Data fusion and enrichment
			Collaborative data analytics
			ML model hosting and AI driven services
			Federated/distributed learning
			Training and education
			Data innovation labs
			Data ethics, fairness and transparency
			Customizable and on-demand services
	Application integration services		AI integration services
			ERP/CRM integration
			Simulation environments (e.g. digital twins)
			Integration with virtual worlds
	Infrastructure integration services		Infrastructure catalogue
			Orchestrator
			Load balancing
			Different types of provisioning (storage, computing…)
	Business enablement services		Billing
			Smart contracts
			Certifications

https://dssc.eu/space/BVE2/1071257170/Value+creation+services

2B. Identification of resulting data products based on the above activities:

What kind of re-usable data products can be produced based on the above activities?

2C. Interlinking the activities (value creation services) and data products:

[*Collect here the names of your value creation services and resulting data products*]

Value creation service name nr [VCSx]:	Data product name nr [DPx]:
[*Your value creation service name(s)*]	[*Your data product name(s)*]

3. SPECIFY EACH DATA PRODUCT AND IDENTIFY THEIR KEY RESOURCES

3A. Data product specification

Identification of the data product provider:

- Who is going to produce the data product to be used by the data consumer to build further services?

Identification of the contract terms for the data product:

- What kind of contract terms shall be used for the data product?

Maintaining the data product:

- What is needed for maintaining the data product?

3B. Key data resources for the data product

Description of the data resources:

- What are the data resources needed for building a data product?
- Are these data resources available?

Rights management of the data resources:

- Who provides the data needed for the data product?
- Who posesses rights to the data and what kind of rights?
- How are rights management (provenance, lineage) taken care of?
- Are the rights of the data providers compatible with the license terms of the data product?

4. ECOSYSTEM AND SCALING PERSPECTIVES

4A. Understanding the business aspects of the data product users

- Who is going to use the data product to generate further services?
- How will the data product user combine different data products into a further offerings?
- What will be the revenue mechanism for the data product user?
- Do the data product contract terms fit into the business model of the data product user?

4B. Scaling opportunities

Re-use opportunities for the intermediate services and data products:

- How can data products be reused?
- What kinds of further downstream services can be innovated from the data products?
- What kinds of novel customer segments can be found?

STEP 3: Identifying data product flows

The logical flow of data products starts from identifying the chain of actions from different data sources, either 3rd party data or data from a data space participant, continuing to developing data products through intermediate services. These data products are meant to be re-used in a manner adding value to the data. The chain of actions may end by a service provider providing services to the end customer. These actions may be provided by different actors. The flow of actions can be expressed as per the figure on the right.

The purpose of the data product flow is to gain an understanding of the iterativeness, re-use, and value-adding role of the data products. In addition, it increases the understanding of the phases and components involved in the value creation process and requires clear identification of the actors and actions involved in different phases.