
**SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY**

Master's Degree in Information Systems Development and Security

March 2026

# Data Sovereignty in Data Spaces

An Analytical Evaluation of Sovereignty Mechanisms in Distributed Data-Sharing Architectures

Supervisor:
George Xylomenos
Professor

Candidate:
Katerina Basmpa
F3312408

## Acknowledgements

I would like to express my sincere gratitude to my supervisor for his valuable guidance and constructive feedback throughout the preparation of this thesis. His insights were essential in shaping the direction and development of this research.

Special thanks are due to the research team involved in the SeEDS project, whose work on data spaces and decentralized data management served as an important source of inspiration for selecting and exploring the topic of data sovereignty. Their research contributions highlighted the practical relevance and technological potential of data space architectures.

# Abstract

The rapid growth of data-driven technologies has intensified the need for mechanisms that enable secure and trustworthy data sharing, while preserving the rights of data owners. In this context, the concept of *data sovereignty* has emerged as a fundamental principle in modern data governance, emphasizing the ability of organizations and individuals to retain control over how their data is accessed, shared, and utilized. *Data spaces* have been proposed as a promising architectural framework for enabling sovereign data exchange across distributed and heterogeneous environments.

This thesis investigates how data sovereignty is implemented within data-sharing infrastructures, with particular emphasis on architectures aligned with the European data space vision. Based on the design principles defined by the *Data Spaces Support Centre (DSSC)*, an analytical evaluation framework consisting of six dimensions was developed to assess the implementation of sovereignty-related mechanisms.

Using this framework, a set of representative research works and system architectures were analyzed to evaluate their alignment with key sovereignty principles. The results indicate that most existing approaches strongly emphasize identity management, access control, and interoperability mechanisms, while governance structures, regulatory compliance mechanisms, and auditing infrastructures are less consistently implemented.

These findings highlight both the progress made and the remaining challenges in realizing fully sovereign data-sharing environments, emphasizing the need for integrated technical and governance solutions capable of supporting trustworthy and scalable data ecosystems.

# Table of Contents

## Contents

# 1 Introduction

The expansion of digital technologies has resulted in a substantial increase in global data generation. Although large-scale storage and computational technologies are widely available, effective data reuse remains limited in many contexts. This limitation is primarily associated with governance, trust, and strategic considerations, rather than purely technical constraints [1]. In particular, uncertainty regarding data integrity, security, and usage conditions restricts efficient data exchange between data holders and potential data users and prevents data reuse across different applications and contexts.

To address these challenges and support the development of the European digital economy, the concept of *Data Spaces* has emerged as a strategic framework for secure and trusted data sharing [2]. Data spaces are designed to enable controlled data exchange among participants, while maintaining regulatory compliance (especially, privacy) and supporting innovation within governed environments. As part of the European data strategy, data spaces aim to promote data reuse across sectors, while preserving participant control over data assets [2].

Guided by the growing importance of data spaces and data sovereignty in modern data governance, this thesis aims to examine the design and implementation of sovereignty-oriented mechanisms within data space environments. The study focuses on the technical and governance aspects that support controlled data exchange, trust establishment, and policy enforcement.

To achieve this objective, selected research and industrial contributions will be evaluated against a set of design principles derived from the Data Sovereignty and Trust pillar defined by the Data Spaces Support Centre [1]. The evaluation will be conducted using an analytical framework that translates sovereignty principles into structured assessment dimensions, so as to determine the degree of alignment between existing solutions and sovereign data space requirements.

## 1.1 Conceptual Foundations of Data Spaces and Data Sovereignty

Data spaces and data sovereignty are fundamental concepts in contemporary data governance, although their definitions may vary across academic and policy literature. For the purposes of this study, the following interpretations are adopted.

A data space is defined as *an interoperable framework that enables trusted data exchange among participants within a federated data sharing environment governed by shared technical, organizational, and policy rules*. Interoperability must be maintained across technical, semantic, and governance layers to ensure consistent interpretation and controlled use of shared data. Supporting functions typically include participant identification, secure onboarding processes, data discoverability, and regulated access management [3].

Data sovereignty refers to *the ability of data holders to maintain meaningful control over their data throughout the data lifecycle, including storage, sharing, and usage*

*phases.* Rather than implying complete data isolation, data sovereignty represents controlled data utilization that balances protection requirements with opportunities for value creation through regulated data sharing. In practice, data sovereignty is realized through governance policies, contractual usage frameworks, and technical enforcement mechanisms that regulate data transactions [4]. These mechanisms strengthen trust among participants, support regulatory compliance, and enable secure cross-domain data exchange across heterogeneous infrastructures. Design guidance for sovereignty-oriented data space implementation is provided by reference frameworks such as those developed by the Data Spaces Support Centre [1].

Data spaces have been proposed with the express goal of allowing data, especially user-created, to escape from the silos of existing cloud-based applications, such as social networks. The goal is to allow data reuse across different applications, without locking in users to a specific data silo and provider. In this context, data sovereignty is critical, as user-created data must be shared in a manner consistent with the desires of the user who owns them, as opposed to the current state of affairs, where data are effectively controlled by the owners of the silos where they are stored. As a result, data sovereignty is possibly the most critical aspect of data spaces.

# 2 Design Principles for Data Sovereignty in Data Spaces

The *Data Spaces Support Centre* (DSSC), established under the Digital Europe Program, provides a structured framework for the development of interoperable and sovereign data spaces. Within this framework, six conceptual pillars are defined, among which *Data Sovereignty and Trust* is a foundational element [2].

Although the DSSC identifies thirteen design principles across all pillars, this thesis focuses exclusively on those principles that implement the *Data Sovereignty pillar*. This approach ensures adherence to the central research objective: the technical and governance realization of sovereign control in data space environments.

Data sovereignty, in the DSSC context, refers to the capability of participants to retain control over their data assets, define usage conditions, enforce access policies, and ensure trustworthy and compliant interactions across federated ecosystems. The following principles are therefore considered directly relevant [1].

## 2.1 Design Principles Realizing Data Sovereignty

**Ensuring Participants' Rights in Data Sovereignty**

This principle constitutes the normative core of the sovereignty pillar. It guarantees that data owners maintain control over access conditions, usage constraints, and consent management. Sovereignty must be technically enforceable rather than merely contractually declared. Governance structures must therefore support fine-grained access control, policy specification, and continuous monitoring of compliance.

**Establishing Trust and Security in Data Spaces**

Sovereign control presupposes a secure and trustworthy infrastructure. This principle requires robust identity management, authentication mechanisms, secure communication channels, and cryptographic safeguards. Trust is not treated as an abstract value, but as an operational requirement supported by verifiable credentials, secure execution environments, and compliance validation services.

**Transparency**

Transparency supports sovereignty by enabling traceability, accountability, and auditability of data transactions. Governance processes, policy rules, and decision-making mechanisms must be clearly defined and accessible. Technically, transparency requires logging mechanisms, provenance tracking, and the provision of structured metadata, so as to ensure that data usage remains verifiable.

**Compliance with EU Legal Frameworks and Norms**

Sovereign data spaces must operate within the European regulatory landscape, including instruments such as the Data Governance Act, the Data Act, and the General Data Protection Regulation (GDPR). This principle ensures that sovereignty is not isolated from legal accountability. System design in a data space must incorporate organizational and technical safeguards that support lawful and fair data processing.

**Adaptable Data Space Governance Framework**

Effective sovereignty requires clearly defined roles, responsibilities, and decision-making procedures. Governance frameworks must remain adaptable to regulatory and technological evolution, while ensuring structured oversight. Modular governance architectures and periodic review processes support long-term sustainability and policy alignment.

**Data Interoperability in Data Spaces**

Sovereignty must be maintained across heterogeneous systems and organizational boundaries. Interoperability at the syntactic, semantic, and policy levels enables a consistent interpretation of usage conditions and enforcement mechanisms. Standardized data models, shared vocabularies, and interoperable protocols are therefore essential for scalable sovereignty.

## 2.2 Technical Realization of Data Sovereignty and Trust

While design principles define normative requirements, technical building blocks translate sovereignty into actual capabilities. Within the DSSC Blueprint, the technical realization of the Data Sovereignty and Trust pillar is supported by specific operational components [5]. This thesis focuses on the building blocks that directly enable sovereign control mechanisms. Data Sovereignty and Trust encompass the technical capabilities required to:

- Identify participants and their digital assets,
- Establish verifiable trust relationships between them,
- Define, negotiate, and enforce access and usage policies.

The following building blocks constitute the core operational infrastructure.

### 2.2.1 Identity and Attestation Management

Identity and attestation management provides the foundation for trusted participation within a data space. Sovereign control presupposes the reliable identification of actors and the verifiable confirmation of compliance with governance rules.This building block enables:

- Issuance and validation of verifiable credentials.
- Authentication of organizations, individuals, and services.
- Management of membership and compliance attestations.
- Lifecycle control including renewal and revocation.

Interoperability is achieved through standardized identity frameworks such as the W3C *Verifiable Credentials* (VC) data model and *Decentralized Identifiers* (DIDs). Credential exchange protocols support secure and machine-readable verification across distributed ecosystems. Identity management therefore transforms abstract trust assumptions into verifiable and auditable digital evidence.

### 2.2.2 Trust Framework

The trust framework supports governance principles through structured conformity assessment and credential validation mechanisms. It ensures that participants comply with the data space rulebook before and during participation. Core components include:

- Machine-readable governance rulebooks.
- Compliance verification services.
- Registries of trust anchors and accredited trust service providers.
- Digitally signed attestations that can be used to verify conformity.

Trust anchors establish foundational authority, while trust services issue and validate credentials on their behalf. By integrating governance rules with technical validation mechanisms, the trust framework enables scalable and automated trust enforcement across federated data spaces.

### 2.2.3 Access and Usage Policy Enforcement

Access and usage control is the most direct technical expression of data sovereignty. It enables participants to define, negotiate, and enforce machine-readable policies governing data sharing and processing. This building block supports:

- Translation of legal and business constraints into formal policy languages.
- Automated validation of policy consistency,
- Distributed enforcement across data transaction lifecycles.
- Real-time compliance verification during data exchange.

The *Open Digital Rights Language* (ODRL) is recommended for expressing permissions, prohibitions, and obligations. Policy negotiation and exchange can be supported by interoperable protocols such as the Dataspace Protocol. Operationally, participant agents act as enforcement points, while trust services validate associated credentials and policy compliance. Together, these mechanisms ensure that sovereignty extends beyond initial access control and remains enforceable throughout data usage.

## 2.3 Analytical Framework for Evaluating Data Sovereignty in Data Spaces

In order to systematically assess the extent to which existing research and industrial implementations realize data sovereignty within data spaces, this thesis adopts an analytical evaluation framework derived from the Data Sovereignty and Trust pillar of the DSSC [1].

While the DSSC defines normative design principles and associated technical building blocks, the academic literature often varies in scope, maturity, and implementation depth. Consequently, a structured evaluation model is required to determine the degree of alignment between theoretical sovereignty claims and their realization in a specific architecture or implementation. The framework developed in this thesis translates

DSSC sovereignty requirements into measurable evaluation dimensions. These dimensions function as assessment criteria that are applied consistently across all selected papers as part of this thesis.

### 2.3.1 Evaluation Dimensions

The analytical framework is structured around six core dimensions corresponding to the realization of data sovereignty.

#### *Sovereign Control and Usage Enforcement*

This dimension evaluates whether participants retain technically enforceable control over their data assets. Key assessment questions include:

- *Are access and usage policies formally defined*?
- *Are policies machine-readable and interoperable*?
- *Is enforcement automated and active throughout the data lifecycle*?
- *Can data providers revoke or modify usage conditions dynamically*?

A strong implementation requires distributed enforcement mechanisms rather than static contractual agreements.

#### *Identity and Verifiable Trust*

This dimension examines whether the system ensures reliable identification of participants and verifiable compliance with governance requirements. Evaluation criteria include:

- *Presence of federated identity management*.
- *Use of verifiable credentials or attestations*.
- *Defined trust anchors or certification authorities*.
- *Lifecycle management of identity credentials*.

Sovereignty presupposes authenticated and accountable actors; anonymous or weakly identified participation weakens sovereign guarantees.

#### *Governance and Rule Enforcement*

This dimension assesses whether governance structures are formally embedded in the architecture. Key elements include:

- *Clearly defined participant roles and responsibilities.*
- *Machine-readable rulebooks or governance specifications.*
- *Automated conformity assessment mechanisms.*
- *Conflict resolution or compliance monitoring processes.*

Sovereignty must operate within a structured governance environment to ensure enforceability and fairness.

#### *Transparency and Auditability*

This dimension evaluates the system's ability to ensure traceability and accountability. Assessment indicators include:

- *Logging mechanisms for data transactions.*
- *Provenance tracking capabilities.*
- *Audit services enabling verification of policy compliance.*
- *Visibility of decision-making processes.*

Sovereignty without observability cannot be verified in practice, therefore it becomes an set of claims.

#### *Legal and Regulatory Alignment*

This dimension measures whether sovereignty mechanisms are aligned with applicable legal frameworks. Evaluation aspects include:

- *Explicit consideration of GDPR, Data Act, or other regulatory instruments.*
- *Technical support for consent management.*
- *Organizational safeguards ensuring lawful data processing.*
- *Integration of compliance checks within system workflows.*

This ensures that sovereignty extends beyond technical control to legal accountability.

#### *Interoperability and Scalability*

This dimension examines whether sovereignty mechanisms function beyond isolated implementations. Evaluation criteria include:

- *Use of standardized policy languages (e.g., ODRL).*
- *Adoption of interoperable identity frameworks.*
- *Compatibility with cross-domain data exchange protocols.*
- *Scalability across multiple organizations or sectors.*

Localized implementations that cannot scale across federated ecosystems demonstrate limited sovereignty maturity and are only a first step towards the real concept of data spaces.

### 2.3.2 Sovereignty Maturity Classification

Based on the six dimensions presented above, each evaluated paper will be categorized into one of three maturity levels:

#### *High Alignment*

The evaluated system demonstrates strong and explicit implementation of the majority of the six dimensions. Sovereignty-related mechanisms are not only conceptually acknowledged but are clearly implemented through technical, organizational, or procedural means. Since all six dimensions are treated as equally important in this thesis, a work is classified as High Alignment only when sovereignty is implemented in a consistently strong manner *across the framework as a whole*.

#### *Partial Alignment*

The evaluated system addresses data sovereignty in a meaningful way, but the implementation remains incomplete or uneven across the six dimensions. Some sovereignty-related mechanisms are clearly realized, while others are only partially

developed, indirectly supported, or insufficiently specified. This classification applies when the solution demonstrates some sovereignty support, but does not achieve a balanced and strong implementation across all equally weighted dimensions.

*Low Alignment*

The evaluated system refers to data sovereignty only at a conceptual or declarative level, without providing substantial technical or architectural mechanisms to accomplish it. In such cases, sovereignty is weakly embedded overall, and most of the six dimensions remain largely unsupported.

The final sovereignty maturity level of each study is derived through a qualitative aggregation of the six evaluation dimensions. The classification is therefore not based on numerical scoring, but on an interpretative assessment of the overall implementation pattern. Since all six dimensions are considered equally critical, no single dimension is treated as inherently more important than another in determining the final maturity level.

This classification provides a consistent basis for comparing heterogeneous research contributions while preserving the qualitative character of the evaluation.

# 3 Evaluation of Data Sovereignty Implementations

## 3.1 Evaluation Methodology

The objective of this chapter is to evaluate selected research and industrial contributions concerning the realization of data sovereignty within data space environments and, more generally, in distributed data-sharing environments that are similar to data spaces. While some works explicitly propose architectures for data spaces, others focus on specific technical mechanisms that enable sovereign data control, such as consent management, access control, identity verification, and privacy-preserving data processing, which are the cornerstones of data spaces proper.

The final selection of the papers was based primarily on the extent to which they address the concept of data sovereignty within data spaces. However, due to the limited number of studies focusing exclusively on data spaces, the scope was extended to cover a broader range of work in data spaces or similar, that nonetheless places clear emphasis on the implementation of data sovereignty. In addition, the selected works provide sufficient conceptual and technical details to enable their evaluation against the design principles and analytical dimensions proposed in this thesis.

The evaluation was conducted using the analytical framework defined in Section 2.3, which is derived from the design principles of the Data Sovereignty and Trust pillar established by the Data Spaces Support Centre[1]. The framework translates sovereignty requirements into six assessment dimensions: *sovereign control and usage enforcement*, *identity and trust mechanisms*, *governance embedding*, *transparency and auditability*, *legal and regulatory alignment*, and *interoperability and scalability*.

Each selected paper is analyzed individually, following a two-stage assessment process. *First*, a concise summary of the paper's main contribution is presented, focusing on its technical or governance approach to data sharing and management. The summary avoids a detailed reproduction of the original work, emphasizing instead sovereignty-relevant aspects.

*Second*, the evaluation focuses on identifying the existence and maturity of sovereignty-supporting mechanisms within each work. Particular attention is paid to distinguishing conceptual discussions of sovereignty from concrete technical or architectural implementations.

Based on the *overall assessment* across dimensions, each paper is classified into one of three alignment levels: High Alignment, Partial Alignment, or Low Alignment. The purpose of this classification is not comparative ranking but rather the analysis of sovereignty realization maturity solutions. The evaluation is conducted in a sequential manner, presenting and analyzing each selected work independently to maintain methodological consistency.

## 3.2 Scope of the Evaluation

The evaluation conducted in this thesis focuses on the realization of data sovereignty principles within selected research and industrial contributions. The analysis is limited to sovereignty-related technical and governance mechanisms that support controlled

data sharing, trust establishment, and policy enforcement in data space environments as well as in other cases.

The study does not evaluate performance metrics such as computational efficiency, network latency, or system scalability under actual load. Additionally, verification of source code correctness, security penetration testing, or experimental benchmarking is explicitly outside the scope of this work.

Furthermore, the evaluation does not aim to rank the selected papers or determine their practical chances of deployment success. Instead, the objective is to analyze the degree to which each contribution aligns with sovereignty-oriented design principles derived from the Data Sovereignty and Trust pillar of the DSSC framework [1].

## 3.3 Limitations of the Evaluation Framework

While the analytical framework provides a structured method for assessing the realization of data sovereignty principles, certain limitations must be acknowledged.

First, the evaluation relies exclusively on information presented within the selected papers. In cases where implementation details are limited or selectively reported, the assessment may underestimate the actual maturity of the proposed solutions; for example, since publication, the authors may have created additional prototypes, which are not considered. The framework does not incorporate any external validation of the technical claims made or independent verification of system functionality.

Second, the framework is derived from the Data Sovereignty and Trust pillar of the DSSC definitions and therefore reflects the normative orientation of this reference model. Alternative data sovereignty formulations existing in the broader academic literature are not explicitly integrated into the evaluation criteria.

Third, the maturity classification (High, Partial, Low Alignment) is based on qualitative analysis rather than quantitative scoring. Although the use of clearly defined analytical dimensions enhances consistency, a degree of interpretative judgment remains inherent in the classification process.

Finally, the framework focuses specifically on sovereignty-related design aspects. Broader system qualities such as economic viability, user adoption, security robustness beyond sovereignty mechanisms, or large-scale operational performance are outside the scope of this assessment.

Despite these limitations, the framework provides a coherent and consistent basis for examining how selected works realize data sovereignty mechanisms within distributed data-sharing architectures and related data management environments.

## 3.4 Evaluation of Selected Works

### 3.4.1 “Enabling Compute and Data Sovereignty with Infrastructure-Level Data Spaces”

#### *3.4.1.1 Contribution Overview*

The paper [6] proposes an infrastructure-level approach for implementing data spaces that supports both secure data exchange and distributed computation across clusters

managed by different organizations. The architecture is designed to address the challenge of enabling collaboration between data producers and data consumers, while maintaining control over sensitive data assets. The authors argue that existing application-level data space solutions focus mainly on data exchange, but do not adequately support computation close to the data source, which may lead to unnecessary data transfers and increased risks of data leakage.

To address this issue, the paper introduces a multi-cluster architecture based on a cross-cluster federation framework (Liqo), which enables container orchestration clusters (Kubernetes) to interconnect and share resources within a distributed computing environment. In the proposed model, the data consumer's computational workload can be offloaded to the data producer's cluster, rather than transferring the raw data to the consumer's infrastructure for processing. This approach supports the concept of data gravity, by moving computation closer to the data source, while allowing the data provider to maintain control over the execution environment.

The architecture incorporates several security mechanisms designed to protect sensitive data during inter-cluster interactions. These include restricted network connectivity between pods, firewall rules enforced through a cross-cluster networking gateway, auxiliary containers that monitor traffic, and container orchestration network policies that limit communication between services. These mechanisms aim to prevent unauthorized access and data exfiltration, while allowing controlled processing of the data within the provider's cluster.

The authors demonstrate how this infrastructure-level data space can be integrated with existing data space initiatives such as the *International Data Spaces* (IDS) framework and Gaia-X. Through this integration, the proposed approach aims to combine secure data exchange mechanisms with distributed computing capabilities that support sovereign control over data processing.

#### *3.4.1.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The proposed architecture enforces sovereign control through infrastructure-level mechanisms that restrict both data access and processing behavior. Specifically, instead of transferring raw data to external environments, the system enables the execution of consumer workloads within the data provider's cluster, via pod offloading. This ensures that data remains within the administrative domain of the provider, with computation brought to the data.

In addition, the architecture applies multiple enforcement layers, including Kubernetes Network Policies, auxiliary container based traffic control, and gateway-level filtering using iptables rules, which collectively regulate how offloaded workloads interact with data resources. The use of auxiliary proxy containers further enables monitoring and control of outgoing traffic, preventing unauthorized data exfiltration.

This directly satisfies the requirement for technically enforceable usage control, as the restrictions are applied at runtime through system-level controls, rather than relying solely on contractual agreements. Furthermore, the data provider retains the ability to dynamically revoke access, by terminating inter-cluster connections or removing offloaded workloads, ensuring continuous control throughout the data lifecycle.

**Assessment:** *High* alignment, as the system demonstrates strong and technically enforceable controls over both data access and data usage.

### Identity and Verifiable Trust

The architecture assumes a trusted environment where participating clusters authenticate each other before establishing communication channels. The integration with the IDS connector architecture provides mechanisms for authentication and authorization between participating entities, including the use of certificates or authentication tokens during connector interactions. These mechanisms ensure that only authorized actors can access shared data services within the data space. However, the implementation primarily assumes a trusted environment and does not explicitly describe advanced trust mechanisms, such as decentralized identity frameworks, verifiable credentials, or formal attestation processes.

**Assessment:** *Partial* alignment with identity and verifiable trust requirements, relying mainly on existing trust mechanisms provided by the IDS ecosystem.

### Governance and Rule Enforcement

Governance mechanisms in the proposed architecture are primarily implemented through technical constraints embedded in the infrastructure configuration. The system enforces restrictions on communication privileges, network connectivity, and execution privileges for offloaded workloads. These constraints define how external workloads may interact with services inside the data producer's cluster and ensure that the execution environment complies with predefined security rules.

**Assessment:** *Partial* alignment with governance and rule enforcement, with governance primarily implemented through infrastructure-level constraints.

### Transparency and Auditability

The architecture incorporates monitoring mechanisms through auxiliary containers and gateway components that inspect traffic generated by offloaded workloads. These components observe communication flows and ensure that only authorized interactions occur between pods and services inside the cluster. However, the paper provides limited discussion of audit logs or provenance tracking mechanisms that would allow full traceability of data usage across the entire processing lifecycle.

**Assessment:** *Low to partial* support for transparency and auditability

### Legal and Regulatory Alignment

The paper discusses data sovereignty in the context of increasing regulatory requirements related to data protection and data sharing. However, the implementation focuses primarily on technical enforcement mechanisms rather than explicit integration with regulatory frameworks. Legal compliance is therefore implicitly supported through technical safeguards, rather than through explicit regulatory enforcement mechanisms.

**Assessment:** *Low* alignment with explicit legal and regulatory compliance mechanisms.

#### Interoperability and Scalability

The proposed architecture achieves interoperability via Kubernetes-based multi-cluster orchestration and the Liqo framework, which enables seamless interconnection of heterogeneous infrastructures. This allows different organizations to participate in a shared data space, while maintaining independent control over their resources. Furthermore, the architecture is explicitly designed to integrate with established data space initiatives such as those from the *International Data Space Association* (IDSA) and Gaia-X, ensuring compatibility with existing standards and enabling participation in federated data-sharing environments.

This satisfies the requirement for cross-domain interoperability, as the system supports communication across heterogeneous infrastructures and aligns with widely adopted data space standards. Additionally, the ability to dynamically establish and terminate inter-cluster connections supports scalability across multiple participants and domains.

**Assessment:** *High* alignment, as the architecture demonstrates strong interoperability and scalability across federated environments.

#### *3.4.1.3 Overall Sovereignty Level*

The proposed architecture demonstrates a significant degree of technical capability for enforcing sovereign control over data access and processing. By enabling computation to be executed directly within the data provider's infrastructure, the architecture reduces the need for transferring raw data and allows the data provider to regulate how data is accessed and processed.

However, some aspects of the sovereignty framework remain only partially addressed. In particular, the architecture relies mainly on infrastructure-level mechanisms for enforcing governance and security rules, while transparency mechanisms and explicit regulatory compliance features are less developed.

Considering the six evaluation dimensions defined in this thesis, the examined solution demonstrates partial alignment with the principles required for the implementation of data sovereignty in data spaces.

**Sovereignty Maturity Classification:** *Partial Alignment*

### 3.4.2 "Data Sovereignty for AI Pipelines: Lessons Learned from an Industrial Project at Mondragon Corporation"

#### *3.4.2.1 Contribution Overview*

The examined paper [7] investigates the role of data sovereignty in collaborative AI pipelines, through an industrial case study conducted at Mondragon Corporation. The authors focus on the challenges that arise when multiple organizations cooperate in AI development processes and need to share data, while maintaining control over how this data is accessed and used.

The study extends an existing AI pipeline, in which sensor data collected by industrial systems is transmitted to an external data quality service provider. To address concerns related to data governance and partner collaboration, the authors introduce a data sovereignty component that regulates how shared data can be accessed and processed within the collaborative pipeline. The goal of this component is to ensure that data is

used only under agreed conditions, while still enabling cooperation between organizations participating in the AI workflow.

The research adopts an action research methodology conducted over a twelve-month industrial project. Based on the experiences gathered during the deployment of the sovereignty component, the authors derive a set of lessons learned, benefits, and barriers associated with implementing data-sovereign AI pipelines. The results highlight how sovereignty mechanisms can facilitate controlled data sharing between organizations and reduce governance obstacles that typically hinder collaborative data science initiatives.

Rather than proposing a new technical architecture for data spaces, the paper primarily contributes empirical insights derived from the practical deployment of sovereignty mechanisms in an industrial setting. The study therefore focuses on organizational and governance aspects of implementing data sovereignty in collaborative AI systems.

#### *3.4.2.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The paper emphasizes the importance of ensuring that data shared within collaborative AI pipelines is used according to predefined agreements between participating organizations. The introduced sovereignty component allows data providers to specify conditions under which their data may be accessed and processed by external partners. This mechanism aims to prevent uncontrolled data usage and to ensure that data remains under the governance of the original provider during collaborative processing. However, the implementation focuses primarily on governance procedures and organizational agreements, rather than on technically enforced usage policies or machine-readable policy frameworks.

**Assessment:** *Partial* alignment. The study acknowledges the need for sovereign control but relies largely on organizational governance mechanisms, rather than fully automated policy enforcement.

##### Identity and Verifiable Trust

Trust between collaborating organizations is treated as an essential prerequisite for collaborative AI pipelines. The study describes how participating partners must establish trust relationships before data sharing can occur. These relationships are supported by contractual agreements and defined collaboration structures within the project. Nevertheless, the paper provides limited discussion regarding the implementation of federated identity management, verifiable credentials, or formal trust infrastructures that could validate participant identities within the system.

**Assessment:** *Low* alignment. Trust is mainly established through organizational collaboration structures rather than through formal identity and credential mechanisms.

##### Governance and Rule Enforcement

Governance mechanisms constitute a central component of the proposed approach, particularly in the context of coordinating data exchange and responsibilities across multiple organizations involved in the AI pipeline. The implementation explicitly requires the definition of roles, responsibilities, and coordination procedures between

participating actors, as well as clearly specified conditions under which data can be accessed and shared.

More specifically, the study highlights the need for a common definition of user roles, standardized governance procedures, and structured coordination processes to manage interactions between data providers and external service providers. These elements are reflected in the formulation of lessons learned, such as the "Need for a Common Definition of User Roles" and the "Need for a Separation of Control and Data Plane," which emphasize the importance of formally organizing responsibilities and control structures within the system.

This satisfies the requirement for defined participant roles and governance structures, as responsibilities and interaction rules are explicitly identified and systematically integrated into the design. However, governance mechanisms are primarily implemented at an organizational and procedural level, and the paper provides limited evidence of machine-readable rulebooks or automated enforcement of governance rules within the system architecture.

**Assessment:** *High alignment,* as governance is explicitly defined, systematically structured, and recognized as a critical requirement for enabling coordinated and controlled data sharing, even though automation of rule enforcement remains partially specified.

### Transparency and Auditability

The study acknowledges the importance of transparency when organizations share data within collaborative AI workflows. In particular, the sovereignty component aims to ensure that partners are aware of how shared data is used within the pipeline. However, the paper does not provide a detailed description of logging systems, provenance tracking, or auditing infrastructures that would allow systematic monitoring of data usage. As a result, transparency mechanisms appear to rely primarily on organizational communication and process coordination, rather than on dedicated technical monitoring solutions.

**Assessment:** *Low to partial* alignment.

### Legal and Regulatory Alignment

Legal and contractual aspects of data sharing are strongly emphasized in the study. The introduction of the sovereignty component is partly motivated by the need to ensure that collaborative AI workflows comply with legal requirements related to data governance and organizational responsibilities. The study highlights the importance of defining data usage agreements and governance procedures that regulate how partners may access shared datasets. However, the paper does not present detailed mechanisms for integrating specific regulatory frameworks such as the GDPR into the technical architecture of the AI pipeline.

**Assessment:** *Partial* alignment. Legal governance is recognized but primarily addressed at the organizational level.

### Interoperability and Scalability

The examined AI pipeline supports collaboration between multiple organizations involved in industrial data analysis. The study demonstrates how data sovereignty

mechanisms can facilitate data sharing between organizations participating in a joint AI development process. However, the solution is designed primarily for a specific environment and does not explicitly describe standardized interoperability mechanisms that would enable large-scale integration across multiple sectors or data space infrastructures. As a result, the scalability of the proposed approach beyond the examined industrial context remains limited.

**Assessment:** *Low* to *partial* alignment.

#### *3.4.2.3 Overall Sovereignty Level*

While the paper demonstrates that sovereignty components can improve trust and coordination between participating organizations, the implementation focuses primarily on governance procedures and collaboration practices rather than on fully automated technical enforcement mechanisms. Several elements associated with sovereign data spaces, such as machine-readable policy enforcement, federated identity infrastructures, and standardized interoperability mechanisms, are only partially addressed.

Considering the six evaluation dimensions defined in this thesis, the examined solution demonstrates partial alignment with the design principles supporting data sovereignty in data spaces.

**Sovereignty Maturity Classification:** *Partial Alignment*

### 3.4.3 "A Reference System Architecture with Data Sovereignty for Human-Centric Data Ecosystems"

#### *3.4.3.1 Contribution Overview*

This paper [8] develops an architecture for human-centric B2C data ecosystems, explicitly designed to support data sovereignty for individuals (data subjects) within a multi-actor data sharing and data utilization setting. It defines a decentralized, role-based architecture consisting of data suppliers (Data Owners, Data Providers), data demanders (Data Consumers, Workbench Providers), and several intermediary roles (e.g., Broker, Registrar, Fiduciary, Vocabulary Curator, Data Quality Curator) that enable data sharing, monetization, and compliant processing while aiming to preserve transparency, fairness, and trust.

A central architectural concept is the use of *Personal Data Storages* (PDS) offered by Data Providers as decentralized storage endpoints connected via a Data Resource Port; Data Owners access and manage their data through PDS accounts and are intended to exercise data subject rights in a sovereignty-preserving manner. The architecture also integrates a hybrid consent model: while Data Owners retain control, certain "activity types" (e.g., importing/storing/offering data) can be delegated to Data Providers through broad consent ("admin policies") to address usability constraints, whereas specific consent is required for responding to data orders.

To standardize and enforce usage conditions, the architecture requires Data Owners (or Providers) to specify usage policies that are translated into standardized data licenses stored and managed through a License Repository. Consent-related records (licenses, admin policies, data order responses) are consistently recorded in a Consent Registry, which is evaluated by the Registrar to authorize processing requests. Transparency of

processing is supported through a Processing Information Directory, which records the identity of the responsible actor and processing purposes/activity types for authorized and performed processing activities.

#### *3.4.3.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The architecture provides sovereign control through a set of explicit architectural mechanisms that regulate how personal data can be accessed and processed. These include user-managed PDSs, which allow data subjects to directly manage their data and exercise their rights, formalized usage policies that specify permitted processing conditions, and standardized data licenses that encode these policies into machine-readable and enforceable rules.

A key implementation element is the Consent Registry, where consent artifacts (e.g., licenses, admin policies, and data order responses) are persistently stored and used by the Registrar to evaluate processing requests before any data access occurs. This ensures that data processing is only executed when it complies with predefined usage conditions. Additionally, the architecture explicitly supports revocation mechanisms, allowing data subjects to withdraw consent and trigger deletion or termination of data access.

These mechanisms collectively demonstrate that control is embedded into the system architecture through enforceable policy evaluation and consent tracking.

**Assessment:** *High* alignment, as the architecture provides concrete, technically enforceable mechanisms (e.g., consent registry, license evaluation, revocation support) that enable data owners to define, monitor, and dynamically control data usage.

##### Identity and Verifiable Trust

The architecture defines explicit roles and assigns responsibilities to actors (Data Provider/Consumer as controllers, Registrar as evaluator of processing requests), and it records actor identities alongside their processing purpose in the Processing Information Directory, supporting accountability. However, the paper does not focus on a specific federated identity stack (e.g., verifiable credentials, trust anchors) as a primary technical substrate; trust is largely achieved through the intermediary structure and registries rather than explicit VC-style infrastructures.

**Assessment:** *Partial* alignment. While the system defines identifiable ecosystem roles and registries supporting accountability, it does not implement a full federated identity or verifiable credential infrastructure**.**

##### Governance and Rule Enforcement

Governance is structurally embedded in the architecture through a well-defined role model and intermediary services, including the Broker, Registrar, Fiduciary, and various curator roles. Each actor has clearly specified responsibilities, contributing to a coordinated governance framework across the ecosystem.

Crucially, governance is not limited to organizational definitions, but is implemented through technical enforcement mechanisms. In particular, the Registrar evaluates processing requests by verifying them against stored consent records (e.g., licenses and

admin policies) in the Consent Registry. This ensures that all data processing activities comply with predefined governance rules before execution.

Furthermore, the use of standardized data licenses transforms governance rules into machine-readable policies, enabling consistent and automated enforcement across the system. This is a shift from purely contractual governance toward architecture-level enforcement of rules and responsibilities.

**Assessment:** *High* alignment, as governance is explicitly defined through roles and responsibilities and is technically enforced via automated validation of processing requests, ensuring compliance at the system level.

### Transparency and Auditability

Transparency is explicitly supported through dedicated system components that ensure traceability of both consent and processing activities. The Consent Registry records all consent-related artifacts (e.g., licenses, admin policies, and data order responses), providing a consistent and verifiable record of user decisions.

In parallel, the Processing Information Directory records key metadata about executed processing activities, including the processing purpose and the identity of the responsible actor. This enables reconstruction of data usage history and supports accountability requirements.

These mechanisms ensure that all processing actions are documented, traceable, and auditable, allowing both system participants and regulators to verify compliance with declared usage conditions. The architecture explicitly aligns these features with GDPR documentation requirements, particularly regarding purpose limitation and accountability.

**Assessment:** *High* alignment, as the system provides concrete logging and registry-based mechanisms that enable full traceability and accountability of data processing activities.

### Legal and Regulatory Alignment

Legal compliance, particularly with the GDPR, is explicitly integrated into the architectural design. The architecture incorporates mechanisms that support core regulatory principles, including:

- Consent management and revocation, implemented via licenses and the Consent Registry
- Purpose limitation, enforced through predefined processing purposes linked to licenses and data apps
- Data subject rights, such as access, deletion, and portability, supported through Personal Data Storages and standardized data models

Additionally, the architecture ensures that processing legitimacy can be demonstrated through recorded consent and purpose documentation, aligning with GDPR requirements. At the same time, the paper acknowledges legal ambiguities, particularly regarding broad consent and admin policies, which are identified as a “grey zone” requiring careful implementation and revocability safeguards.

Thus, legal compliance is not only considered conceptually, but is also realized through concrete system components and processes.

**Assessment**: *High alignment*, as GDPR principles are systematically embedded into the architecture through enforceable mechanisms, while limitations and legal uncertainties are explicitly recognized.

#### Interoperability and Scalability

The architecture promotes interoperability through standardized data import, metadata extraction, controlled vocabularies (data models, usage policy ontologies) in a Vocabulary Catalogue, and unique metadata identifiers that underpin communication across actors. Scalability is supported conceptually through a decentralized infrastructure and modular components, and the authors explicitly relate design commonalities to broader initiatives (e.g., IDS, Gaia-X) at the level of design principles and architectural orientation.

**Assessment**: Partial alignment. Although the architecture introduces standardized vocabularies and modular components, practical interoperability with external identity frameworks and large-scale data space infrastructures is not fully specified.

#### *3.4.3.3 Overall Sovereignty Level*

The proposed architecture presents a comprehensive and structured approach to enabling data sovereignty within human-centric data ecosystems. Rather than treating sovereignty as an abstract principle, it operationalizes it through a combination of technical enforcement mechanisms, architectural roles, and governance processes that collectively regulate how personal data is accessed, shared, and processed.

At its core, the architecture achieves sovereignty through enforceable usage control. This is realized via standardized data licenses and ex ante authorization procedures, where every data access request is validated against predefined consent conditions before execution. This design ensures that control over data is not merely declarative but is technically embedded and continuously enforced within the system.

In parallel, the architecture establishes strong accountability guarantees through registry-based mechanisms. The use of a Consent Registry and a Processing Information Directory enables persistent recording of consent decisions and processing activities, making all data interactions traceable, auditable, and verifiable. This supports both internal governance and external regulatory oversight.

Furthermore, the architecture is explicitly aligned with legal and regulatory requirements, particularly the GDPR. Key aspects, such as consent management, purpose limitation, and data subject rights, are not only acknowledged but are directly implemented through system components, ensuring that compliance is an inherent property of the architecture rather than an external obligation.

While the model places less emphasis on advanced identity infrastructures (e.g., verifiable credentials), it compensates through a well-defined role-based ecosystem and intermediary services that facilitate trust, coordination, and rule enforcement across participants.

Overall, the architecture demonstrates a high level of sovereignty alignment, as it integrates control, governance, transparency, and compliance into a cohesive system design. Its primary limitation lies in the partial specification of interoperability with

external identity and data space frameworks, which may affect its applicability in broader, cross-ecosystem deployments.

**Sovereignty Maturity Classification:** *High Alignment*.

### 3.4.4 "Declarative Policy Control for Data Spaces: A DSL-Based Approach for Manufacturing-X"

#### *3.4.4.1 Contribution Overview*

The paper [9] proposes a *Domain-Specific Language* (DSL) for specifying and managing data usage policies within industrial data spaces. The authors focus on a key challenge in federated data ecosystems: enabling domain experts to define and enforce data governance policies without requiring extensive software engineering expertise. The proposed approach is developed within the context of the Manufacturing-X initiative, which aims to enable secure and sovereign data exchange across industrial production networks.

To address the complexity of configuring data space connectors, the authors examine existing implementations of discovery and access-control mechanisms based on a range of industrial communication, data exchange, and digital asset representation technologies. Based on this analysis, they derive a unified metamodel that captures the essential configuration elements required to implement discovery and usage control layers in data space connectors.

The metamodel is implemented as a DSL that allows engineers to define connector configurations declaratively. Through this DSL, data governance rules, including access permissions, usage policies, contract constraints, and identity provider configurations, can be specified in a human-readable format that can later be processed by machines. The authors demonstrate the feasibility of the approach through a prototype implementation, applied to an industrial testbed representing a software-defined value network of interconnected factories.

Overall, the paper contributes a modeling approach intended to simplify the specification of sovereignty-related policies in industrial data spaces by translating complex connector configurations into structured and declarative policy definitions.

#### *3.4.4.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The proposed DSL focuses on enabling the declarative specification of data usage policies within data space connectors. It allows stakeholders to define access permissions, contractual conditions, and usage constraints (e.g., restricting usage to monitoring purposes or limiting access to specific roles). These policies are integrated into connector configurations via access policies and contract offers, making them machine-readable and interoperable across different technologies.

However, the enforcement of these policies is primarily delegated to the underlying data space connectors, rather than being directly implemented or validated within the DSL itself. The paper does not provide detailed mechanisms for runtime enforcement,

continuous monitoring, or dynamic revocation of usage conditions, which are key aspects of full sovereign control.

**Assessment**: *Partial* alignment, as the approach enables structured and machine-readable policy definition, but relies on external systems for enforcement and does not fully demonstrate lifecycle-level control over data usage.

### Identity and Verifiable Trust

The system incorporates identity-related configuration elements within the connector specification. The unified metamodel includes components for identity provider configuration, authentication credentials, and OAuth-based authentication mechanisms that regulate secure interactions between data space participants. These mechanisms ensure that data access requests are associated with identifiable actors and verified through authentication services. However, the architecture does not implement a comprehensive decentralized identity or verifiable credential framework.

**Assessment:** *Partial* alignment. The architecture integrates identity provider configurations and authentication mechanisms but does not fully implement federated identity infrastructures or verifiable credential frameworks.

### Governance and Rule Enforcement

Governance is addressed through the DSL's ability to encode access rules, roles, and contractual conditions into a unified configuration model. These governance elements define which participants can access specific data assets and under what conditions. By translating governance rules into machine-readable configurations, the approach supports their consistent interpretation across heterogeneous data space technologies.

Nevertheless, governance enforcement is not implemented as an independent architectural layer but is instead embedded within connector configurations and dependent on external execution environments. The paper does not introduce mechanisms such as automated conformity assessment, rulebook validation, or continuous compliance monitoring, which are central to fully operational governance frameworks.

**Assessment:** *Partial* alignment, as governance rules are formally modeled and integrated into configurations, but enforcement remains indirect and lacks dedicated compliance verification mechanisms.

### Transparency and Auditability

The architecture includes mechanisms that support traceability of data usage through structured policy definitions and metadata associated with data assets. The unified metamodel also records descriptive metadata and configuration parameters that define how data resources are discovered and accessed within the data space. However, the paper provides limited discussion of logging infrastructures, provenance tracking systems, or independent audit mechanisms that would allow systematic monitoring of policy compliance.

**Assessment:** *Partial* alignment. While policy definitions and metadata support transparency of governance rules, the implementation of detailed auditing and monitoring mechanisms remains limited.

#### Legal and Regulatory Alignment

The paper emphasizes the importance of policy-driven data governance in enabling trusted data exchange across organizational boundaries. By enabling declarative specification of usage policies and contractual constraints, the DSL provides a mechanism through which organizations can encode legal conditions governing data usage. Nevertheless, the study does not explicitly implement or analyze compliance with specific regulatory frameworks such as GDPR or the EU Data Act.

**Assessment:** *Partial* alignment. Legal governance principles are acknowledged through policy-based controls, but explicit integration of regulatory compliance mechanisms is not extensively addressed.

#### Interoperability and Scalability

The unified metamodel integrates configuration elements from multiple industrial data space technologies, combining capabilities from industrial communication protocols, data exchange connectors, digital asset representation models, and linking mechanisms for distributed data resources. By abstracting common elements across these technologies into a single DSL, the approach enables interoperable configuration of heterogeneous connectors and facilitates integration across industrial environments.

The modular structure of the metamodel allows extension with technology-specific components, supporting adaptability and reuse across different implementations. This abstraction layer reduces complexity and promotes consistent configuration practices across federated systems.

**Assessment**: *High alignment*, as the architecture provides a concrete mechanism (unified metamodel and DSL) that enables interoperability across multiple standards and supports scalable deployment across heterogeneous environments

#### *3.4.4.3 Overall Sovereignty Level*

The proposed DSL-based approach contributes a structured and practical mechanism for modeling and managing sovereignty-related policies within industrial data spaces. Its primary strength lies in enabling domain experts to define machine-readable governance and usage policies that can be integrated into data space connectors.

The architecture demonstrates particularly strong support for interoperability and scalability, achieved through the unification of multiple industrial standards within a single metamodel. In addition, it provides consistent support for policy specification and governance modeling.

However, the implementation of several key sovereignty dimensions is limited. Policy enforcement is delegated to external systems, while identity management, auditability, and regulatory compliance mechanisms are only partially addressed. The absence of explicit runtime enforcement, comprehensive monitoring, and formal compliance verification limits the extent to which sovereignty is implemented within the proposed approach.

Considering the six evaluation dimensions defined in this thesis, the solution provides consistent but incomplete support across most dimensions, with only one dimension demonstrating full implementation.

**Sovereignty Maturity Classification:** *Partial Alignment*

## 3.4.5 "Secure Computation and Trustless Data Intermediaries in Data Spaces"

#### *3.4.5.1 Contribution Overview*

The paper [10] investigates how advanced cryptographic technologies can support secure and sovereign data processing within data spaces. Specifically, it focuses on the integration of secure *Multi-Party Computation* (MPC) and *Fully Homomorphic Encryption* (FHE) to enable privacy-preserving computation across distributed participants. The authors argue that these technologies allow data to remain encrypted during processing, thereby preventing unauthorized access to sensitive information, while still enabling collaborative data analysis.

A central concept introduced in the paper is the notion of trustless data intermediaries. Unlike traditional intermediaries that may access or process plaintext data, trustless intermediaries rely on cryptographic methods to facilitate secure computation without being able to view the underlying data. Through MPC and FHE, computations can be performed on encrypted inputs, ensuring that only the end results are revealed to authorized parties, while preserving the confidentiality of the original datasets.

To demonstrate the feasibility of this approach, the paper analyzes several application domains, including air traffic management, manufacturing platforms, and secondary data markets. These case studies are used to identify key technical challenges associated with integrating secure computation into data spaces. The authors organize these challenges into several categories, including identity management, data usage policies, node selection, access control, and the operation of trustless intermediaries.

Building on this analysis, the paper proposes an integration model for secure computation within data spaces. The model includes processes for participant onboarding, asset publication, policy negotiation, and execution of secure computation services within a federated infrastructure. The goal of this architecture is to support scalable, privacy-preserving collaboration across multiple organizations while maintaining control over data usage and processing conditions.

#### *3.4.5.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The architecture implements sovereign control primarily with secure MPC and FHE, which ensure that data remains protected during computation and is not disclosed in plaintext to other participants or intermediaries. In the proposed deployment model, input data is encoded or encrypted before computation, and only the result is reconstructed by authorized result parties. This means that the system does not merely restrict initial access to data but technically constrains how data can be processed throughout the computation phase itself. In addition, the transaction phase requires that requested computations comply with predefined policies attached to assets before execution can begin, linking data usage directly to agreed conditions. These mechanisms satisfy the core requirement of this dimension, namely that control over data usage is technically enforced rather than left to contractual trust alone.

**Assessment**: *High alignment*. The architecture provides concrete technical enforcement of data usage control through MPC and FHE, ensuring that sensitive data remains protected during processing and that computations are executed only under predefined conditions.

### Identity and Verifiable Trust

The architecture includes identity-related mechanisms in the onboarding phase, where participants are authenticated before being allowed to join the data space. The paper proposes that this validation may rely on external trust anchors, such as established electronic identification and trust service frameworks, domain-validated digital certificates, or globally recognized legal entity identification systems and it discusses the use of DIDs and VCs to support secure and privacy-preserving authentication. It also considers more advanced mechanisms, such as attribute-based credentials and selective disclosure, to enable participants to prove eligibility without necessarily revealing their full identity. These elements show that the architecture recognizes identity verification and credential-based trust as essential conditions for secure participation in data spaces. However, these mechanisms are presented primarily as part of the proposed integration approach and the challenge analysis, rather than as a fully implemented and operational trust framework within the system itself.

**Assessment:** *Partial* alignment. The architecture incorporates identity verification and credential-based trust as important design elements, but it does not fully implement a comprehensive and operational identity and verifiable trust infrastructure.

### Governance and Rule Enforcement

Governance is realized through the combination of asset-specific policies, contract negotiation, and transaction validation. During the setup phase, participants publish assets together with associated policies describing how these assets may be used. In the transaction phase, consumers select combinations of data, compute nodes, and functions, after which an automatic contract negotiation process checks whether the requested computation satisfies the policies of all involved asset providers before the contract is signed and the computation is triggered. This constitutes more than a conceptual governance discussion: it is a concrete mechanism for enforcing usage conditions at the point of transaction execution. Governance rules must be embedded in the architecture and applied before data processing begins. Although the paper does not implement continuous governance monitoring across the full lifecycle, it does provide a clear and technically grounded mechanism for rule enforcement at the operational level.

**Assessment:** *High alignment*. Governance rules are explicitly attached to assets and enforced through automated contract negotiation and transaction validation, providing a clear architecture-level mechanism for rule enforcement.

### Transparency and Auditability

The architecture incorporates mechanisms for monitoring and validating computation processes. The system supports logging, verification of computation results, and the potential use of cryptographic techniques such as zero-knowledge proofs to ensure correctness and integrity of results. These mechanisms enhance transparency by

allowing stakeholders to verify that computations have been executed according to agreed policies and protocols.

**Assessment:** *Partial alignment*. While the architecture provides mechanisms for verifying computation correctness, the implementation of comprehensive audit infrastructures is only partially addressed.

#### Legal and Regulatory Alignment

The paper explicitly references European regulatory frameworks such as the GDPR and the Data Governance Act, emphasizing that data spaces must comply with these legal requirements when enabling data sharing and processing. The proposed architecture also highlights the role of regulated data intermediaries in facilitating compliant data exchange across organizations. However, the paper focuses primarily on technical architecture rather than detailed implementation of legal compliance mechanisms.

**Assessment:** *Partial alignment*. Regulatory considerations are acknowledged and integrated conceptually but are not extensively implemented in the proposed architecture.

#### Interoperability and Scalability

The framework is explicitly designed for data space integration and supports operation within federated ecosystems by building initiatives such as IDS, Gaia-X, and broader European data space efforts. Its architecture enables different participants to publish assets, negotiate policies, and execute secure computations across distributed environments rather than within a single closed system. In addition, the paper treats dynamic participation, flexible node selection, and interoperable policy and identity mechanisms as core integration requirements, which shows that the system is conceived for deployment across heterogeneous and evolving infrastructures. This directly satisfies the criterion that sovereignty mechanisms must function beyond isolated implementations, supporting distributed, cross-organizational collaboration. Even though some standardization gaps remain, interoperability and scalability are not merely future aspirations in the paper; they are foundational design assumptions of the proposed framework.

**Assessment:** *High* alignment. The architecture is explicitly designed for federated and distributed data space environments and supports interoperable, scalable collaboration across multiple participants and infrastructures.

#### *3.4.5.3 Overall Sovereignty Level*

The paper presents a comprehensive architectural approach for integrating privacy-preserving computation technologies into data spaces. Its main contribution lies in showing how secure MPC and FHE can support collaborative data processing while ensuring that sensitive data remains protected during computation. In addition, the framework incorporates policy-driven transaction validation and is explicitly designed for operation within federated data space environments, thereby strengthening technical support for controlled and interoperable data exchange.

At the same time, the analysis reveals that several dimensions of data sovereignty remain only partially accomplished. Specifically, identity and verifiable trust mechanisms are discussed as essential architectural elements, but they are not implemented as a complete and fully operational trust framework. Likewise, transparency and auditability are supported only indirectly through verifiability features rather than through a comprehensive monitoring and auditing infrastructure. Legal and regulatory alignment is also acknowledged, especially in relation to European data governance initiatives, but explicit compliance mechanisms are not fully developed.

Considering the six evaluation dimensions defined in this thesis, the examined approach demonstrates strong support for Sovereign Control and Usage Enforcement, Governance and Rule Enforcement, and Interoperability and Scalability, while Identity and Verifiable Trust, Transparency and Auditability, and Legal and Regulatory Alignment remain only partially addressed.

**Sovereignty Maturity Classification**: *Partial* alignment

### 3.4.6 "You Shall Not Pass (Without Consent): Enforcing Data Sovereignty with Solid Pods"

#### *3.4.6.1 Contribution Overview*

This paper [11] proposes a decentralized architecture for enforcing data sovereignty in privacy-preserving data platforms, through the integration of Solid Pods as personal data stores. The approach builds upon the Solid protocol to enable individuals to store and manage their personal data in decentralized repositories, while maintaining direct control over data sharing and access permissions. The authors aim to address the limitations of centralized data platforms, which often restrict user control and transparency over how personal data are processed.

The proposed system introduces two primary architectural components: the Request Portal and the Solid Gateway. The Request Portal allows individuals acting as data contributors to review incoming data requests and selectively approve or reject them. Once consent is granted, the Solid Gateway retrieves the authorized data from the user's Solid Pod and forwards it to a privacy-preserving analysis environment. This architecture supports request-based data sharing while maintaining user control over access permissions.

A key innovation is the implementation of a granular data-sharing strategy. Instead of granting access to complete data files, the system extracts only the data elements required for a given request and generates a minimal dataset for sharing. This approach aims to reduce unnecessary data exposure and align data-sharing processes with the principle of data minimization defined in the GDPR.

To evaluate the proposed architecture, the authors developed a prototype implementation involving multiple simulated data contributors and conducted experiments using synthetic datasets. The evaluation examines system performance, data exposure levels, and the effectiveness of the granular sharing mechanism compared with alternative approaches such as centralized repositories and file-level sharing strategies.

#### *3.4.6.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The architecture implements sovereign control through personal Solid Pods, where data contributors remain the direct custodians of their data and decide whether a request is accepted or rejected. Access is not granted globally or permanently, but through request-specific authorization implemented via Web Access Control lists, which bind permissions to dedicated request identities. In addition, the proposed Granular strategy restructures data into request-specific files containing only the minimal subset needed for the approved analysis, rather than exposing the entire data. This means that sovereign control is enforced through three concrete mechanisms: user-controlled storage, explicit consent before access, and technically restricted access to minimal data subsets. The architecture also supports dynamic revocation, since contributors can withdraw consent by deleting or modifying the corresponding access control entry, after which the Solid Gateway will no longer retrieve the data. These mechanisms demonstrate that control over access and usage is not merely declared, but implemented directly in the data-sharing workflow.

**Assessment:** *High* alignment. The system provides concrete technical mechanisms for user-controlled authorization, minimal data exposure, and dynamic revocation, thereby strongly implementing sovereign control over data access and use.

##### Identity and Verifiable Trust

Identity and trust are implemented through standardized web-based identification and authentication mechanisms, which enable consistent identification of both users and services within the ecosystem. In addition, the system generates a dedicated identity reference for each request, ensuring traceability and controlled access, ensuring that permissions are not granted to broad or reusable identities but to narrowly scoped request-specific identities. This strengthens trust by isolating permissions across analytical tasks and reducing the risk of uncontrolled reuse of access rights. Because authentication and authorization are tied to these standard identity mechanisms, the architecture satisfies the requirement that only authenticated and specifically authorized actors can access requested data resources. The design therefore embeds identity directly into the access-control workflow.

**Assessment**: *High* alignment. The architecture uses standardized and request-specific identity mechanisms that enable verifiable authentication and fine-grained authorization for both users and services.

##### Governance and Rule Enforcement

Governance is embedded in the system through a structured request-based workflow that requires data consumers to specify the purpose of the request, the requested data attributes, and the intended processing method before access can be granted. These requests are coupled with consent artifacts structured using the Data Privacy Vocabulary and enforced through access control lists in the contributor's pod. As a result, governance rules are translated into operational constraints that determine which request may access which data and under what declared conditions. However, while the architecture clearly realizes request-level consent and access governance, it does not fully implement broader governance structures, such as a complete machine-readable

rulebook, automated conformity assessment, or ecosystem-level compliance monitoring.

**Assessment:** *Partial* alignment. The architecture provides concrete governance support through request-based consent workflows and enforceable access policies, but it does not fully implement the broader governance and compliance mechanisms required for a high classification**.**

### Transparency and Auditability

The architecture incorporates auditability through the Solid Gateway's provenance log, which records each data retrieval event together with the accessed URL, timestamp, HTTP status code, and a cryptographic hash of the retrieved content. These records allow data access operations to be reconstructed retrospectively and enable integrity checks over retrieved resources. In parallel, consent is documented in structured files stored in the contributor's pod, including acceptance time, expiry time, and withdrawal time when applicable. This creates two complementary traces: one for user decisions, and one for actual retrieval operations. As these traces are directly linked to requests and data accesses, the system provides more than general transparency; it offers concrete evidence that data access events can be inspected and verified after the fact.

**Assessment:** *High* alignment. The system includes explicit provenance logging and structured consent records that provide traceability, integrity verification, and accountability for data-sharing operations.

### Legal and Regulatory Alignment

Legal and regulatory alignment is explicitly reflected in the architecture through its implementation of data minimization, purpose-specific consent, and consent withdrawal, all of which are central requirements of the GDPR. The Granular strategy directly supports Article 5 GDPR by restricting access to only the data elements necessary for a specific approved request, rather than transferring full files. Consent is not assumed at registration time, but is managed per request, with the possibility of revocation at any time and immediate effect on subsequent retrieval attempts. In addition, consent artifacts are structured and stored in a way that documents the legal basis and scope of the approved sharing action. The architecture therefore does not merely reference regulatory principles abstractly but incorporates them into concrete technical mechanisms that shape how data is exposed and processed.

**Assessment:** *High* alignment. The architecture concretely implements regulatory principles such as data minimization, purpose-bound consent, and consent withdrawal through technical mechanisms embedded in the system design.

### Interoperability and Scalability

The system is based on open web and semantic technologies that enable structured data representation, standardized communication, decentralized identity management, and privacy-aware data modeling, providing a strong foundation for interoperability with broader web-based and privacy-preserving data infrastructures. The architecture also demonstrates integration with external data analysis environments, showing that decentralized storage can interoperate with existing processing frameworks rather than requiring a closed, standalone solution. However, although the paper provides strong evidence of interoperability, its evaluation is conducted in a controlled environment and

does not validate scalability across larger multi-organizational or cross-sector deployments. The paper itself identifies real-world deployment complexity, heterogeneous pod providers, and distributed performance as issues for future work.

**Assessment:** *Partial* alignment. The architecture demonstrates strong interoperability through open standards and integration with existing privacy-preserving platforms, but scalability across broader federated deployments remains only partially demonstrated.

#### *3.4.6.3 Overall Sovereignty Level*

The examined architecture presents a strong and technically grounded approach to implementing data sovereignty within decentralized data-sharing platforms. By combining Solid-based personal data stores, consent-driven access control, and granular data-sharing mechanisms, the system enables individuals to retain direct control over how their data is accessed and processed. In addition, the architecture incorporates concrete mechanisms for decentralized identity, provenance logging, and structured consent management, which together provide a robust basis for accountable and privacy-aware data sharing.

The evaluation shows particularly strong alignment in the dimensions of Sovereign Control and Usage Enforcement, Identity and Verifiable Trust, Transparency and Auditability, and Legal and Regulatory Alignment. These dimensions are supported by explicit technical mechanisms embedded in the architecture rather than by abstract design intentions alone. At the same time, Governance and Rule Enforcement and Interoperability and Scalability are only partially developed, since broader ecosystem-level governance structures and large-scale federated deployment are not yet fully demonstrated.

Considering the six evaluation dimensions defined in this thesis, the proposed solution demonstrates strong overall support for the implementation of data sovereignty in decentralized data-sharing infrastructures, even though some aspects remain only partially developed.

**Sovereignty Maturity Classification**: High Alignment

### 3.4.7 "Securing Data Sovereignty and Data Security for Independent Participants in Supply Chains"

#### *3.4.7.1 Contribution Overview*

The paper [12] presents an approach to enabling secure information exchange among independent companies participating in industrial supply chains. The primary objective is to support collaborative data analysis across organizations, while preserving data sovereignty and protecting sensitive industrial information. The authors point out that companies are often reluctant to share production data, due to concerns related to intellectual property protection and cybersecurity risks. To address these concerns, the proposed concept replaces traditional raw data exchange with a model based on selective information sharing. Instead of transferring complete datasets, participating organizations exchange only predefined information derived from their internal datasets. This approach allows companies to retain control over their original data while still enabling collaborative analysis across supply chain partners.

The architecture is developed within the context of a regional innovation initiative focused on advancing digital transformation in manufacturing environments, targeting small and medium-sized enterprises. The system connects distributed data sources from multiple companies, using a linking platform that facilitates coordinated data analysis without centralizing raw data.

The architecture integrates several technologies to support decentralized data processing and collaborative analytics. These include a distributed machine learning approach that enables model training across datasets held by different parties, a coordination framework that orchestrates the learning process among participants, and a workflow integration platform that facilitates information exchange between organizations. The prototype implementation is structured into three main layers: a data management layer built on scalable distributed storage and processing infrastructure, an intermediary agent layer that enables autonomous coordination and communication, and a workflow integration layer responsible for connecting and synchronizing information across supply chain partners.

By combining federated analytics with selective information exchange, the proposed architecture aims to enable cross-company data analysis, while ensuring that sensitive raw data remains under the control of the originating organization.

#### *3.4.7.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The architecture supports sovereign control primarily by ensuring that raw production data remain within the originating organization. Instead of transferring full datasets across the supply chain, participants exchange only selected derived information, while collaborative analysis is performed through federated learning techniques that avoid direct sharing of original records. This design gives each company substantial control over what information is disclosed to partners and reduces the exposure of sensitive internal data.

However, the paper does not clearly specify whether access and usage policies are formally defined in machine-readable form, nor does it demonstrate automated enforcement mechanisms throughout the full data lifecycle. In addition, the ability of data providers to dynamically revoke or modify usage conditions after sharing is not explicitly accomplished.

**Assessment:** *Partial* alignment. The architecture provides strong data-locality and selective-disclosure mechanisms, but it does not fully demonstrate the policy-based, lifecycle-level, and dynamically revocable control required for a high level of sovereign usage enforcement.

##### Identity and Verifiable Trust

The paper discusses the importance of secure communication and trusted interactions between supply chain partners. It highlights the need for secure system integration, authenticated communication channels, and adherence to established cybersecurity frameworks for industrial and operational environments. However, the architecture does not present a detailed identity management framework or mechanisms such as verifiable credentials or decentralized identity infrastructures.

**Assessment:** *Partial* alignment. The architecture emphasizes secure communication and trusted collaboration but does not provide a comprehensive identity management infrastructure.

### Governance and Rule Enforcement

Governance within the proposed architecture is primarily implemented through predefined agreements between participating organizations regarding the types of information that may be exchanged. The workflow connector platform facilitates these agreements by linking datasets across partners based on established relationships within the supply chain. However, the governance framework is largely organizational and contractual rather than implemented through automated rule enforcement mechanisms.

**Assessment:** *Partial* alignment. Governance mechanisms exist through predefined agreements and platform coordination, but automated rulebook enforcement mechanisms are limited.

### Transparency and Auditability

The architecture acknowledges several risks related to data misuse, incorrect interpretation of shared information, and potential insider threats. To mitigate these risks, the authors discuss the need for monitoring mechanisms, secure communication protocols, and system-level security controls. Nevertheless, the system description does not cover detailed auditing infrastructures, provenance tracking mechanisms, or comprehensive logging systems for verifying data transactions.

**Assessment:** *Partial* alignment. The architecture recognizes the importance of monitoring and security controls but provides limited implementation details regarding auditing and traceability.

### Legal and Regulatory Alignment

The paper emphasizes the importance of safeguarding intellectual property and aligning the architecture with established cybersecurity frameworks relevant to industrial environments. It also discusses the need to consider legal aspects of cross-company data exchange within supply chains. However, the architecture focuses primarily on technical security and does not explicitly address broader regulatory frameworks such as GDPR or the European Data Act.

**Assessment:** *Partial* alignment. Legal considerations related to industrial security and intellectual property protection are acknowledged but not comprehensively integrated.

### Interoperability and Scalability

The architecture is designed for heterogeneous industrial environments, where organizations rely on diverse enterprise management systems, production control systems, and data storage infrastructures. It supports interoperability through standardized communication protocols, distributed learning approaches, and modular integration components, enabling effective collaboration across multiple supply chain partners. However, it does not clearly demonstrate standardized policy languages or interoperable identity frameworks, and scalability is addressed mainly within supply chain settings rather than broader federated ecosystems.

**Assessment:** *Partial* alignment. The architecture supports interoperability across heterogeneous industrial systems and multi-organizational collaboration, but it does not fully satisfy all interoperability and scalability criteria of this evaluation framework.

#### *3.4.7.3 Overall Sovereignty Level*

The proposed architecture provides a practical approach for enabling collaborative data analysis in supply chains while preserving organizational control over sensitive production data. By replacing raw data exchange with selective information sharing and integrating federated learning techniques, it enables cross-company insights without requiring centralized storage or disclosure of raw datasets.

At the same time, the analysis shows that sovereignty support is only partial when assessed against the full set of evaluation dimensions defined in this thesis. While the solution introduces strong technical mechanisms for keeping raw data local and enabling collaboration across heterogeneous industrial systems, several other dimensions remain insufficiently developed. In particular, the architecture does not clearly implement machine-readable usage policies, lifecycle-level enforcement, dynamic revocation of usage conditions, interoperable identity frameworks, or comprehensive auditing and compliance mechanisms.

Considering the six evaluation dimensions defined in this thesis, the proposed solution demonstrates relevant but incomplete support for the implementation of data sovereignty in distributed industrial environments. Its main strength lies in protecting proprietary production data during collaborative analytics, while governance, identity, auditability, legal alignment, and broader interoperability mechanisms remain only partially addressed.

**Sovereignty Maturity Classification:** *Partial* Alignment.

### 3.4.8 “Implementing Data Sovereignty: Requirements & Challenges from Practice”

#### *3.4.8.1 Contribution Overview*

The paper [13] investigates the practical challenges and requirements associated with implementing data sovereignty in industrial data-sharing environments. Rather than proposing a specific technical architecture, the study focuses on understanding how organizations currently approach data sovereignty and which obstacles they encounter when sharing data across company boundaries. The authors conduct a qualitative study based on semi-structured interviews with eleven industry experts, applying grounded theory methods to identify patterns in real-world implementations of sovereign data-sharing systems.

The interviews reveal several motivations for data sharing within industrial ecosystems, including production optimization, regulatory compliance, sustainability monitoring, and data monetization. These use cases illustrate that companies increasingly rely on cross-organizational data exchange, while simultaneously seeking mechanisms to maintain control over their data assets.

From the qualitative analysis, the authors derive seven core requirements for implementing data sovereignty in practice. These include organizational requirements such as legal compliance, confidentiality protection, and classification of data services,

as well as technical requirements including policy enforcement, data security and integrity, data visibility, and transparency.

In addition to identifying requirements, the study highlights thirteen practical challenges encountered during the implementation of data sovereignty solutions. These challenges are grouped into organizational, technical, and socio-cultural categories. Key technical challenges include the difficulty of enforcing data usage policies, managing heterogeneous infrastructures, handling the full data lifecycle, and establishing reliable identity management frameworks. Organizational and human factors such as lack of standards, trust issues, and limited organizational readiness also significantly affect the implementation of sovereignty mechanisms.

The paper concludes that data sovereignty is not solely a technical problem, as it also involves governance structures, legal frameworks, organizational processes, and trust relationships among participating organizations. The findings therefore provide a practical perspective on the requirements and obstacles involved in developing sovereign data-sharing infrastructures.

As the paper does not propose itself an architecture, in the following we instead evaluate how the use cases studied via the questionnaires fit within our assessment framework.

#### *3.4.8.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The study identifies policy enforcement as one of the most critical technical requirements for implementing data sovereignty. Interview participants emphasized the importance of controlling data usage even after data has been shared, including mechanisms such as anonymization, data deletion policies, and access authorization frameworks. However, the study also highlights that existing implementations largely rely on extended access control mechanisms rather than comprehensive usage control enforcement.

**Assessment:** *Partial* alignment. The paper recognizes the importance of usage control mechanisms, but indicates that fully enforceable policy-driven control is still limited in practice.

##### Identity and Verifiable Trust

Identity verification and trust relationships between organizations are discussed as essential prerequisites for sovereign data sharing. Interview participants raised questions regarding the identification of organizations, certification authorities, and trust anchors responsible for verifying participants within data-sharing ecosystems. The study emphasizes that unresolved identity verification challenges remain a significant barrier to implementing reliable data sovereignty infrastructures.

**Assessment:** *Partial* alignment. Identity management is recognized as an important requirement but remains insufficiently implemented in existing industrial solutions.

##### Governance and Rule Enforcement

The study highlights the importance of governance frameworks and organizational agreements for managing data-sharing relationships. Companies commonly rely on legally binding contracts, internal policies, and organizational processes to regulate data access and sharing conditions. However, these governance mechanisms are often

implemented through manual or organizational processes rather than automated policy enforcement infrastructures.

**Assessment:** *Partial* alignment. Governance structures exist primarily in contractual and organizational forms rather than through automated rule enforcement systems.

#### Transparency and Auditability

Transparency and traceability are identified as important requirements for data-sharing environments. Interview participants emphasize the need for transparent processes, logging mechanisms, and audit capabilities to verify how data is processed and accessed across organizations. Nevertheless, the study indicates that many existing solutions still lack comprehensive auditing infrastructures capable of monitoring full data lifecycles.

**Assessment:** *Partial* alignment. Transparency mechanisms are recognized as necessary but remain only partially implemented in practice.

#### Legal and Regulatory Alignment

The paper identifies legal and regulatory compliance as a central requirement for implementing data sovereignty in practice. Interview participants refer to frameworks such as the GDPR and the European Data Act, especially in relation to data deletion, consent, and legal responsibility. However, the study does not itself implement concrete technical mechanisms for compliance or integrate compliance checks into a system architecture.

**Assessment:** *Partial* alignment. Legal and regulatory requirements are clearly recognized, but they are not implemented through explicit compliance mechanisms within the proposed contribution.

#### Interoperability and Scalability

The study identifies infrastructure heterogeneity as a major technical challenge when implementing data sovereignty across organizations. Differences in IT landscapes, legacy systems, cloud platforms, and applications complicate the deployment of interoperable data-sharing solutions. Technologies such as dataspace connectors based on the International Data Spaces architecture are mentioned as potential approaches for addressing these interoperability challenges.

**Assessment:** *Partial* alignment. Interoperability challenges are acknowledged, but fully scalable solutions remain under development.

#### *3.4.8.3 Overall Sovereignty Level*

The study provides a practice-oriented analysis of data sovereignty, highlighting how organizations approach data sharing and where implementation gaps emerge. Its findings confirm that sovereignty is a multi-dimensional concept, requiring the alignment of technical, organizational, and legal mechanisms.

Across the evaluated dimensions, the results indicate consistent but partial implementation. Sovereign control is mainly realized through extended access control rather than enforceable usage control. Identity and trust mechanisms are recognized as essential but remain insufficiently developed. Governance is primarily implemented through contractual agreements and organizational processes, with limited automation. Transparency and auditability are acknowledged, yet comprehensive lifecycle

monitoring is often lacking. Legal and regulatory requirements are well understood, but not technically embedded. Finally, interoperability challenges persist due to heterogeneous infrastructures and limited standardization.

Overall, the study demonstrates that while the key dimensions of data sovereignty are clearly identified in practice, their technical and operational realization remains incomplete and fragmented.

Considering the six evaluation dimensions defined in this thesis, the work shows that the use cases study can be classified to have partial alignment, as they capture the essential requirements but do not provide fully implemented or enforceable sovereignty mechanisms.

**Sovereignty Maturity Classification:** *Partial* Alignment.

### 3.4.9 "A Reference Architecture for Enabling Interoperability and Data Sovereignty in the Agricultural Data Space"

#### *3.4.9.1 Contribution Overview*

The paper [14] proposes a reference architecture for an *Agricultural Data Space* (ADS), aimed at addressing interoperability challenges and data sovereignty concerns in digital farming environments. The agricultural domain is characterized by a fragmented digital landscape where multiple software systems, machines, and service providers operate within separate digital ecosystems, often using heterogeneous data formats and standards. As a result, integrating agricultural data across platforms is technically complex and frequently leads to data silos. In addition to interoperability challenges, the authors highlight that data sovereignty is a key concern for farmers, who often lose control over their data once it is shared with service providers. Farmers frequently lack transparency regarding how their data are processed and who benefits from it, which reduces their willingness to participate in digital agricultural platforms.

To address these issues, the authors design a reference architecture based on a digital platform called the Twin-Hub, which enables secure data exchange while allowing farmers to retain control over their field data. The architecture is structured around three main solution concepts: field data storage, consent and access management, and field data exchange. These components collectively support controlled data sharing across agricultural stakeholders.

A key architectural concept in the proposed solution is the use of digital twins, specifically "digital field twins", which represent digital versions of physical agricultural fields and store field-related data. The Twin-Hub platform acts as the central system managing these digital field twins while enabling service providers to access data through standardized interfaces and controlled access mechanisms.

The architecture further introduces fine-grained consent management and access monitoring mechanisms, allowing farmers to determine who can access specific data, for what purpose, and when. Data access operations are logged to ensure transparency regarding how field data are used by external services.

Overall, the proposed reference architecture aims to support interoperable data exchange within the ADS, while strengthening data sovereignty through consent-based access control and transparent monitoring of data usage.

#### *3.4.9.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The architecture gives farmers direct control over whether external services may access specific field data and for which purpose through the Consent Manager and the Access Manager. Consent is granted at a fine-granular level, and services can access only the specific data elements and operations covered by a valid consent. The platform also supports revocation through the user interface, and access decisions are checked before each request is executed. However, the paper explicitly states that the current solution provides data access control, not full data usage control, and that once data leave the platform, technical enforcement becomes difficult and future work is needed.

**Assessment:** *Partial* alignment. The architecture provides concrete and fine-grained access control with consent and revocation, but full lifecycle usage control is not technically enforced and is explicitly identified as unresolved.

##### Identity and Verifiable Trust

The architecture implicitly assumes that service providers interacting with the Twin-Hub platform can be authenticated and identified when accessing field data. Access management mechanisms rely on authentication tokens and identity verification to validate requests from external services. The paper does not present a detailed identity federation framework or mechanisms such as verifiable credentials or decentralized identity infrastructures.

**Assessment:** *Partial* alignment. Identity verification mechanisms are considered necessary for platform access control, but a comprehensive trust infrastructure is not fully developed.

##### Governance and Rule Enforcement

Governance is embedded through the consent workflow, where farmers act as data owners and decide which services may access which field data, for what purpose, and with which operations. These rules are realized through the Consent Manager and Access Manager, so they are not left purely to external contracts. However, under our framework, strong governance also requires broader architectural governance elements, such as defined participant responsibilities, machine-readable rulebooks, conformity assessment, and compliance monitoring. The paper mainly implements technical consent-based control, while broader ecosystem governance is not fully developed.

**Assessment:** *Partial* alignment. Governance rules are implemented through consent and access-control mechanisms, but the architecture does not fully implement the broader governance structures and compliance mechanisms required for a high classification.

##### Transparency and Auditability

The architecture implements transparency through a dedicated logging mechanism that records every authorized data access event at the level of the individual field twin. Each access request generates a structured log entry containing the identity of the requesting

service, the granted user, the accessed field and data element, the timestamp, and the type of operation performed. This means that transparency is not treated as a general design principle only, but is implemented as a concrete traceability mechanism embedded in the platform workflow.

From the perspective of this evaluation framework, this directly satisfies the core requirements of transparency and auditability. First, it provides logging of data transactions, since every access operation is recorded systematically. Second, it enables provenance tracking, because each log links the access event to a specific service, user, data object, and action. Third, it supports auditability, as these records allow farmers to retrospectively verify how their data were accessed and by whom. The architecture therefore makes data-sharing actions observable and accountable in practice, rather than leaving them opaque after consent has been granted. Although the paper does not describe a separate external audit service, the implemented logging and monitoring functionality is sufficiently concrete and directly connected to system operation to justify a strong evaluation in this dimension.

**Assessment:** *High* alignment. The architecture includes explicit, structured logging mechanisms that enable traceability of data access events and provide a concrete basis for accountability and auditability in practice.

#### Legal and Regulatory Alignment

The architecture acknowledges that legal frameworks and regulations influence data sharing within the agricultural domain. Farmers' concerns about privacy, contractual terms, and data ownership are recognized as major barriers to adoption of digital farming platforms. However, the paper focuses mainly on architecture design and does not provide detailed mechanisms for integrating regulatory compliance processes.

**Assessment:** *Partial* alignment. Legal considerations are recognized conceptually but are not deeply integrated into the technical architecture.

#### Interoperability and Scalability

Interoperability is a central objective of the architecture. The platform uses standardized access, shared vocabularies, and a generic data exchange mechanism to support interaction between heterogeneous agricultural systems. In addition, the architecture is designed as a federated network of multiple Twin-Hub instances, allowing different organizations to host and connect their own instances. However, the paper itself identifies scalability as a crucial quality attribute that still needs to be addressed in future work, which means that scalability is not yet strongly demonstrated in the current architecture.

**Assessment:** *Partial* alignment. The architecture provides strong support for interoperability through standardized exchange and federation, but scalability is explicitly identified as a remaining limitation rather than a fully realized capability.

#### *3.4.9.3 Overall Sovereignty Level*

The proposed reference architecture provides a structured approach to supporting data sovereignty in the ADS by combining consent-based access control, federated platform design, and explicit logging of data access operations. Through digital twins, the Consent Manager, and the Access Manager, farmers retain decision authority over

which services may access specific field data and for which purpose, while the platform also records access events in a way that supports traceability and accountability.

At the same time, the analysis shows that sovereignty support remains uneven across the evaluated dimensions. The architecture offers a clear and well-implemented mechanism for transparency and auditability, but other sovereignty aspects are only partially realized. In particular, the paper explicitly distinguishes between data access control and full data usage control, meaning that restrictions on usage after data leave the platform are not fully enforceable. Likewise, governance is mainly implemented through technical consent workflows rather than broader ecosystem-level governance structures, and scalability is identified as an important objective that still requires further development.

Considering the six evaluation dimensions defined in this thesis, the proposed solution demonstrates meaningful but incomplete support for the implementation of data sovereignty in agricultural data ecosystems. Its main strength lies in transparency and accountable data access, while sovereign usage control, governance breadth, and full interoperability and scalability remain only partially developed.

**Sovereignty Maturity Classification:** *Partial* Alignment

### 3.4.10 "Secure and Efficient Data Spaces (SeEDS)"

#### *3.4.10.1 Contribution Overview*

The *Secure and Efficient Data Spaces* (SeEDS) project is an open-source data space architecture built on top of the *Named-Data Networking* (NDN) information-centric networking paradigm. The project aims to demonstrate how a distributed networking infrastructure can support secure, efficient, and sovereignty-preserving data exchange among multiple participants [15].

The system implements a fully standards-compliant data space architecture, supporting a standardized interface for data exchange. Through this interface, data providers and consumers can publish, discover, and query data entities represented in a structured, linked data format. The platform introduces several functional capabilities including temporal queries, event-based subscriptions, and content filtering, enabling flexible and efficient data management within the data space environment.

A key architectural feature of SeEDS is the deployment of distributed data intermediaries, eliminating single points of failure and improving system resilience. The architecture also incorporates selective disclosure mechanisms, allowing only specific attributes of a data object to be revealed to authorized data consumers while maintaining cryptographic integrity guarantees.

Security and trust are implemented through a decentralized identity framework based on DIDs and certificate-based authorization mechanisms. These components allow data providers to control which services are authorized to advertise and distribute their data assets.

Overall, SeEDS demonstrates how a distributed networking infrastructure combined with standardized APIs and cryptographic mechanisms can support secure and controlled data exchange in a data space context.

#### *3.4.10.2 Sovereignty Alignment Assessment*

##### Sovereign Control and Usage Enforcement

The SeEDS architecture provides mechanisms that allow data providers to maintain control over how their data is accessed and disclosed. In particular, the system implements attribute-level selective disclosure, enabling data providers to reveal only specific attributes of a data object to authorized consumers while preserving cryptographic integrity guarantees. However, although attribute filtering enables partial disclosure of data, the architecture does not fully implement machine-readable policy frameworks or dynamic policy negotiation mechanisms that would enable complex data usage restrictions or lifecycle policy enforcement.

**Assessment:** *Partial* support for sovereign control through attribute-level disclosure mechanisms.

##### Identity and Verifiable Trust

The SeEDS architecture incorporates decentralized identity and trust mechanisms using Decentralized Identifiers and certificate-based authorization. Content providers can authorize specific SeEDS services or proxies to publish content on their behalf, while consumers can verify provenance and authenticity through cryptographic mechanisms tied to these identities. In addition, the project explicitly positions its security model as fully distributed and self-sovereign, avoiding reliance on centralized public key infrastructures. However, although these mechanisms provide a concrete basis for authenticated and verifiable participation, the reports do not fully specify a broader federated identity framework with explicit lifecycle management of credentials across the ecosystem.

**Assessment:** *Partial* alignment. The architecture provides concrete decentralized identity and trust mechanisms, but it does not fully demonstrate the complete federated and lifecycle-managed trust infrastructure required for a high classification.

##### Governance and Rule Enforcement

The SeEDS architecture implements several operational mechanisms supporting governance within the data space infrastructure, including standardized APIs, distributed service coordination, and controlled content advertisement procedures. The system also relies on trusted services responsible for enforcing access control and content filtering operations. However, the report does not define a formal, machine-readable rule framework or governance model comparable to those found in established data space reference architectures.

**Assessment:** *Partial* governance support through operational mechanisms rather than formal governance frameworks.

##### Transparency and Auditability

The architecture includes several mechanisms that support traceability and verification of data operations. In particular, the cryptographic signature scheme used for selective disclosure allows data consumers to verify the integrity and authenticity of the disclosed content. Nevertheless, the system does not explicitly describe logging infrastructures, audit services, or formal provenance tracking mechanisms for monitoring data transactions within the ecosystem.

**Assessment:** *Partial* transparency and auditability through cryptographic verification but limited auditing infrastructure.

#### Legal and Regulatory Alignment

The project aims to implement a standards-aligned data space based on established data space models and standardized data exchange interfaces. These approaches contribute indirectly to regulatory alignment by supporting interoperability and structured governance interactions. However, the report does not explicitly address legal compliance mechanisms, such as consent management, policy enforcement, or regulatory monitoring capabilities.

**Assessment:** *Low* alignment with legal and regulatory governance frameworks.

#### Interoperability and Scalability

Interoperability is a central design goal of the architecture. The system provides a standardized interaction interface that enables consistent communication between data providers, data consumers, and intermediary data services. At the edge of the network, SeEDS services expose HTTP-based interfaces for conventional IP applications and translate these requests into NDN operations, thereby enabling interoperability between legacy IP-based environments and the SeEDS data space infrastructure. In addition, the architecture distributes brokerage and rendezvous responsibilities across multiple SeEDS services and incorporates a primary-secondary replication scheme to avoid single points of failure. The implementation has also been tested over the worldwide NDN testbed, demonstrating its suitability for deployment across geographically distributed infrastructures. These mechanisms together show that interoperability and scalability are not only intended design goals but are concretely implemented and validated in the system.

**Assessment:** *High* alignment. The architecture provides concrete standardized interfaces and a distributed service design that enables interoperability with heterogeneous systems and supports scalable deployment across geographically distributed infrastructures.

#### *3.4.10.3 Overall Sovereignty Level*

Based on the evaluation across the six analytical dimensions, the SeEDS architecture demonstrates a selective but technically significant realization of data sovereignty. Its strongest contribution lies in the way it combines standardized interoperability mechanisms, distributed service design, and cryptographic integrity and provenance protection within a decentralized data-sharing environment. In particular, the architecture provides a concrete basis for interoperability across heterogeneous systems and geographically distributed infrastructures, which is a major strength in relation to sovereignty in federated data spaces.

At the same time, the evaluation shows that sovereignty support is not equally mature across all dimensions. While SeEDS includes meaningful decentralized identity and authorization mechanisms, these do not yet amount to a fully developed federated trust framework with complete lifecycle management of credentials. In addition, broader governance structures and explicit legal or regulatory compliance mechanisms remain less developed than the technical networking and interoperability layers. This means that the architecture is strongest in its technical infrastructure for sovereign data

*exchange,* but less complete in the institutional and compliance-oriented layers of sovereignty.

Considering the six evaluation dimensions defined in this thesis, SeEDS demonstrates strong technical alignment in core infrastructural aspects of sovereignty, while other dimensions remain only partially implemented.

**Sovereignty Maturity Classification:** *Partial* Alignment

## 3.5 Comparative Discussion of Evaluation Results

The results of the evaluation are summarized in *Table 1* below, which presents the level of alignment of each analyzed study with the six analytical dimensions of the proposed data sovereignty evaluation framework. The evaluation of these studies reveals a diverse range of approaches for implementing data sovereignty mechanisms in distributed data-sharing environments. Although the analyzed works differ in scope, technical maturity, and architectural focus, several common patterns emerge regarding how sovereignty principles are implemented.

A *first* observation concerns the strong emphasis on identity and trust infrastructures across most of the examined works. Many architectures implement mechanisms such as federated identity systems, DIDs, or VCs, to ensure that participants can be reliably authenticated and authorized before accessing data assets. These mechanisms are essential for establishing trust relationships between participants and constitute one of the most consistently implemented sovereignty dimensions.

A *second* recurring theme is the use of technical mechanisms for controlling access to data. Several solutions introduce access control models, cryptographic mechanisms, or selective disclosure techniques that enable data providers to restrict the information that becomes available to other participants. In some cases, such as attribute-level disclosure or policy-based access restrictions, these mechanisms allow providers to retain partial control over the visibility of their data assets. However, only a limited number of works implement fully machine-readable usage policies capable of enforcing complex data usage restrictions across the entire data lifecycle.

The analysis also highlights *varying levels of governance support* among the evaluated systems. Some architectures incorporate explicit governance structures or policy frameworks that regulate participant behavior and define operational rules for data exchange. In other cases, governance mechanisms remain implicit or are assumed to be managed externally by organizational agreements rather than embedded directly in the technical architecture. As a result, the realization of governance-related sovereignty principles appears less consistent compared to identity or access control mechanisms.

Regarding *transparency and auditability*, several solutions provide mechanisms that allow verification of data authenticity or integrity, typically through cryptographic signatures or content verification techniques. These mechanisms enable consumers to validate the origin of data objects and detect potential tampering. Nevertheless, comprehensive auditing infrastructures that support full traceability of data usage events are rarely described in detail within the evaluated works.

The evaluation also indicates that *legal and regulatory alignment is only partially addressed* in most technical implementations. While many works reference regulatory frameworks such as GDPR or emphasize the importance of privacy protection, explicit mechanisms for regulatory compliance, such as consent management systems or automated compliance verification, are often absent or only conceptually discussed.

Finally*, interoperability and scalability emerge as key design objectives* across several studies, particularly those explicitly targeting data space environments. Many architectures adopt standardized data exchange interfaces, common data models, or distributed system architectures in order to support cross-organizational data sharing and integration across heterogeneous infrastructures.

Overall, the comparative analysis suggests that current research and implementation efforts primarily focus on *technical enablers of sovereignty*, particularly identity management, access control, and secure data exchange mechanisms. In contrast, governance frameworks, regulatory compliance mechanisms, and comprehensive auditing infrastructures are relatively underdeveloped in many of the evaluated works.

These findings indicate that achieving full data sovereignty requires not only robust technical infrastructures but also the integration of governance and regulatory mechanisms capable of supporting accountable and transparent data sharing across distributed environments.

**Table 1 Comparative Summary of Sovereignty Alignment Across Evaluated Studies**

| Paper | Sovereign Control & Usage Enforcement | Identity & Verifiable Trust | Governance & Rule Enforcement | Transparency & Auditability | Legal & Regulatory Alignment | Interoperability & Scalability | Overall Sovereignty Alignment |
|---|---|---|---|---|---|---|---|
| **Infrastructure-Level Data Spaces** | High Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Low Alignment | High Alignment | Partial Alignment |
| **Data Sovereignty for AI Pipelines (Mondragon)** | Partial Alignment | Low Alignment | High Alignment | Low Alignment | Partial Alignment | Partial Alignment | Partial Alignment |
| **Human-Centric Data Ecosystem Architecture** | High Alignment | Partial Alignment | High Alignment | High Alignment | High Alignment | Partial Alignment | High Alignment |
| **Declarative Policy Control for Data Spaces** | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | High Alignment | Partial Alignment |
| **Secure Computation and Trustless Data Intermediaries** | High Alignment | Partial Alignment | High Alignment | Partial Alignment | Partial Alignment | High Alignment | Partial Alignment |

| Paper | Sovereign Control & Usage Enforcement | Identity & Verifiable Trust | Governance & Rule Enforcement | Transparency & Auditability | Legal & Regulatory Alignment | Interoperability & Scalability | Overall Sovereignty Alignment |
|---|---|---|---|---|---|---|---|
| **Solid Pods Consent Enforcement** | High Alignment | High Alignment | Partial Alignment | High Alignment | High Alignment | Partial Alignment | High Alignment |
| **Supply Chain Data Sovereignty Security** | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment |
| **Implementing Data Sovereignty – Practice** | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment |
| **Agricultural Data Space Architecture** | Partial Alignment | Partial Alignment | Partial Alignment | High Alignment | Partial Alignment | Partial Alignment | Partial Alignment |
| **SeEDS NGI Sargasso Final Report** | Partial Alignment | Partial Alignment | Partial Alignment | Partial Alignment | Low Alignment | High Alignment | Partial Alignment |

# 4 Conclusion

This thesis examined the concept of *data sovereignty within distributed data-sharing environments,* with particular emphasis on the emerging paradigm of *data spaces*. As data sharing becomes increasingly important for innovation, organizations and individuals require mechanisms that allow them to retain meaningful control over their data assets, while still enabling data-based collaboration across organizational and technical boundaries.

The *first part* of the study presented the conceptual foundations of data spaces and data sovereignty, drawing on definitions and architectural guidance provided by European initiatives such as the *Data Spaces Support Centre* (DSSC) [1]. Within this framework, particular attention was given to the Data Sovereignty and Trust pillar, which defines technical and governance principles required for enabling sovereign data exchange[1], [5]. Based on these principles, an analytical framework consisting of six evaluation dimensions was developed in order to assess how different research works implement data sovereignty mechanisms.

In the *second part* of the thesis, a set of representative studies and implementations were evaluated using this framework [6]-[15]. The analysis included both data space–oriented architectures and related distributed data-sharing solutions that address aspects of sovereign data control. The comparative evaluation revealed that most existing approaches focus primarily on *technical enablers of data sovereignty*, particularly identity management, access control mechanisms, and interoperability infrastructures that support secure data exchange across heterogeneous systems. At the same time, the evaluation identified several areas where the realization of data sovereignty remains incomplete. Governance frameworks, automated policy enforcement mechanisms, and comprehensive auditing infrastructures are often only partially addressed or remain conceptual in many proposed architectures. Similarly, explicit integration of *legal and regulatory compliance mechanisms* is less frequently implemented within the technical designs of the examined systems.

Overall*,* the results suggest that while significant progress has been made in developing technical mechanisms that support sovereign data sharing, achieving full data sovereignty requires the *integration of technical, governance, and regulatory components within a coherent architectural framework*. Data spaces represent a promising approach toward this objective, as they aim to combine interoperable infrastructures, shared governance structures, and standardized mechanisms for controlling data usage.

Future research should further explore the development of *machine-readable policy frameworks, automated governance enforcement mechanisms, and scalable trust infrastructures* capable of supporting large-scale sovereign data ecosystems. In addition, empirical validation through real-world deployments will be essential for evaluating the practical effectiveness of proposed sovereignty mechanisms in operational environments. In conclusion, this thesis contributes to the understanding of how data sovereignty can be implemented within distributed data-sharing systems and

highlights both the progress achieved and the challenges that remain in realizing fully sovereign data exchange infrastructures.

# References

**[1]** B. Peeters et al. 2022. Interim Report Data Space Design Principles. Data Space Support Center.

**[2]** Data Space Support Center. https://dssc.eu/space/Partners/175472674/About+DSSC

**[3]** CEN Workshop Agreement. 2024. Trusted Data Transaction. Reference document from CEN Members National Standard Bodies. CWA 18125.

**[4]** International Data Spaces: https://internationaldataspaces.org/why/data-sovereignty/

**[5]** Data Space Support Center Blueprint: https://blueprint.dssc.eu/?pane=technical#KeyConceptsofDataSpaces-3.2BuildingBlocks%2Cthecapabilitiesyouneed%26specificationsyoucanuse

**[6]** J. Marino, L. Camiciotti, F. Cheinasso, A. Olivero, F. Risso. 2023. Enabling Compute and Data Sovereignty with Infrastructure-Level Data Spaces. ESAAM 2023: 3rd Eclipse Security, AI, Architecture and Modelling Conference on Cloud to Edge Continuum.

**[7]** M. Altendeitering, J. Pampus, F. Larrinaga, J. Legaristi, F.M. Howar. 2022. Data sovereignty for AI pipelines: lessons learned from an industrial project at Mondragon corporation. CAIN '22: 1st Conference on AI Engineering – Software Engineering for AI.

**[8]** S. Scheider, F. Lauf, F. Möller, B. Otto. 2023. A Reference System Architecture with Data Sovereignty for Human-Centric Data Ecosystems. Business & Information Systems Engineering. ISSN 1867-0202, Springer Fachmedien Wiesbaden GmbH. Wiesbaden. Vol. 65. Iss. 5. pp. 577-595. https://doi.org/10.1007/s12599-023-00816-9.

**[9]** J. Pfeiffera et al. 2025. Declarative Policy Control for Data Spaces:A DSL-Based Approach for Manufacturing-X. 35th CIRP Design Conference

**[10]** C. Fabianek, S. Krenn, T. Lorunser, V. Siska. 2024. Secure Computation and Trustless Data Intermediaries in Data Spaces. arXiv:2410.16442v1

**[11]** T. Hajszan, M. Staudinger, T. Miksa. 2025. You Shall Not Pass (Without Consent): Enforcing Data Sovereignty with Solid Pods. https://dl.acm.org/doi/10.1145/3771554.

**[12]** J. Kallisch, K.H. Niemann, C. Wunck, M. Runge, M. Voß. 2025. Securing Data Sovereignty and Data Security for independent participants in supply chains. 21. AALE-Konferenz.

**[13]** M. Hellmeier, J. Pampus, H. Qarawlus, F.M. Howar. 2023. ARES 2023: The 18th International Conference on Availability, Reliability and Security. https://dl.acm.org/doi/10.1145/3600160.3604995.

**[14]** R. Falcão, R. Matar, B. Rauch, F. Elberzhager, M. Koch. 2023. A Reference Architecture for Enabling Interoperability and Data Sovereignty in the Agricultural Data Space. Information 2023, 14, 197. https://doi.org/10.3390/info14030197.

**[15]** Y. Thomas, N. Fotiou, I. Pittaras, G. Xylomenos et al. 2024. Secure and Efficient Data Spaces over Named Data Networking. 2025. IFIP Networking 2025 Conference.