Semiconductor-X White Paper

Initiative for the Digitalization of Supply Chains in the Semiconductor Industry



Data Space Security Concept based on the Project Semiconductor-X

Haris Avgoustinos, EXPO21XX GmbH
Florian Bieck, EXPO21XX GmbH
Gerd Brost, Fraunhofer AISEC
Monika Kamhuber, Fraunhofer AISEC
Hendrik Meyer zum Felde, Fraunhofer AISEC
Franz Wenninger, Fraunhofer EMFT
Tobias Kutzler, Fraunhofer IFF
Pascal Birnstill, Fraunhofer IOSB
Christian Haas, Fraunhofer IOSB
Paul Georg Wagner, Fraunhofer IOSB
Hans Kittelmann, Hochschule für Technik und Wirtschaft Dresden
Ingolf Gehrhardt, Hochschule für Technik und Wirtschaft Dresden
Paul Patolla, Hochschule für Technik und Wirtschaft Dresden
Harald Gossner, Intel
Per-Henrick Addicks, SAP SE
Niranjana Subrahmanyam, Siemens AG

Funded by the European Union
NextGenerationEU

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag

Executive Summary

The Semiconductor-X project addresses the secure exchange of industrial data across the semiconductor value chain, including equipment manufacturers, material suppliers, semiconductor manufacturers, and downstream customers. Its use cases require the exchange of sensitive production data, digital twins, manufacturing information, and intellectual property across organizational and technical boundaries. As Semiconductor-X becomes interconnected with other Manufacturing-X initiatives, such as Catena-X and Factory-X, security must also support cross-dataspace supply chains with different participants, governance frameworks, and protection requirements.

This white paper presents a **holistic, risk-based security concept for industrial data spaces**, developed from the Semiconductor-X perspective and based on the Catena-X architecture and Eclipse Tractus-X technologies. It addresses two complementary areas: **security and governance of the federated data space** and **technical and operational security of the data exchange gateways**. The objective is to establish a scalable and interoperable security framework that preserves data sovereignty, supports industrial collaboration, and enables participation by companies of different sizes and security capabilities.

At the data space level, the concept defines the roles and responsibilities of Data Space Authorities, Operators, Federation Service Providers, participants, component operators, and independent certification bodies. It addresses governance, identity and trust management, secure onboarding and offboarding, certification, and continuous operations. Decentralized Identifiers (DIDs), Verifiable Credentials (VCs), identity wallets, and the Dataspace and Decentralized Claims Protocols (DSP/DCP) provide the basis for federated trust establishment. Cross-dataspace interoperability requires mutual recognition of identities, security requirements, and governance rules, supported by coordinated security monitoring and incident response.

At the gateway level, the paper proposes a **three-level security architecture: Basic, Trust, and Critical**, corresponding to the criticality of the exchanged data based on the previous Manufacturing-X Paper about End2End Security. Basic provides cost-efficient protection for low-risk data exchanges. Trust introduces enhanced security measures for confidential business information, including usage control and platform integrity. Critical extends protection to strictly confidential information, such as process recipes, machine configurations, and semiconductor intellectual property, through stronger isolation, hardware-backed attestation, end-to-end encryption, and confidential computing. This graduated approach enables SMEs to start with limited implementation effort and extend their security capabilities as required.

The security architecture follows **Zero Trust, defense-in-depth, and security-by-design principles**. It combines the Eclipse Dataspace Connector (EDC) with complementary security mechanisms, including WAF/API protection, network segmentation, identity and access control, policy enforcement, monitoring, and secure IT/OT integration according to IEC 62443. A threat model based on STRIDE evaluates relevant assets, trust boundaries, data flows, and attacker capabilities, including network attackers, internal attackers, and honest-but-curious infrastructure providers.

We define three security levels for gateways:

To reduce complexity and simplify onboarding, we base our security architecture on a **three-tiered data criticality classification**. Each data criticality is associated with a security level, which defines the security measures necessary to protect data assets of this criticality.

Data Criticality	Examples	**Security Level**	Security Characteristics
Semi-public	Eco Pass Kit, Digital product passports (metadata	**Basic**	Lightweight protection, high interoperability
Confidential	Purchase Orders, Commercial Documents.	**Trust**	Policy agreements, data encryption, system integrity verification, auditability

Data Criticality	Examples	Security Level	Security Characteristics
Strictly confidential	IP (blueprints, recipes, algorithms)	**Critical**	Hardened nodes, technical policy enforcement, secure tunnels, hardware-backed attestation, continuous monitoring

We propose a **security architecture** for data spaces that uses the security levels corresponding to the data criticality levels defined in the previous table. This layered approach balances security effort with functionality, performance, cost, and ease of use, depending on data sensitivity and use-case criticality. The resulting requirements are mapped to established security frameworks and regulations.

The concept further addresses the **complete security lifecycle**, from secure development and certification to deployment, operations, maintenance, and incident response. This includes secure software supply chains, hardened Kubernetes environments, vulnerability and patch management, continuous security monitoring, and coordinated SOC/CSIRT processes. Certification combines technical component assessments with operational evidence and supports different deployment models, including self-hosted, cloud-based, and managed gateways.

To extend security beyond the connector, the proposed **Smart Industrial Data Hub** introduces a common platform for secure industrial data provisioning. It integrates heterogeneous IT/OT systems and cross-dataspace information using Asset Administration Shells (AAS), fine-grained Zero Trust authorization, and contract-based usage control. AI-assisted criticality assessments, data forensics, integrity verification, and contract interpretation support the detection of manipulated data, unauthorized transfers, privacy risks, and malicious content. The platform enables rapid implementation of industrial applications while reusing common integration, security, and governance mechanisms.

Finally, the paper investigates **AI-assisted threat modelling and security requirements engineering** to address the growing number of data space use cases. AI can support architecture analysis, threat identification, security control mapping, and gap analysis. Based on VDE-AR-E 2802-20, the proposed approach establishes reusable control catalogues for the three security levels. New use cases are assessed against these baselines, additional controls are derived where necessary, and residual risks are documented and reviewed. Human validation remains essential for criticality classification, risk acceptance, and security decisions.

The Semiconductor-X security concept thus provides a **scalable foundation for secure industrial data ecosystems**. By combining federated governance, interoperable trust mechanisms, risk-based gateway security, secure data provisioning, and AI-assisted continuous improvement, it aims to reduce implementation complexity while maintaining appropriate protection for industrial data and intellectual property. The approach is intended to support future Manufacturing-X applications and cross-dataspace collaboration without requiring a new security architecture for every use case.

Executive Summary i
1 Introduction 1
2 Holistic Approach to Securing Data Spaces 2
2.1 Securing the Data Space 2
2.2 Securing the Data Space Gateway 3
2.3 Case Study: Semiconductor-X 5
3 Data Space Security Approach 5
3.1 Design and Development Phase 5
3.2 Deployment Phase 19
3.3 Operations Phase 20
3.4 Maintenance Phase 22
4 Gateway Security Approach 25
4.1 Design and Development Phase 27
4.2 Deployment Phase 54
4.3 Operations Phase 58
5 Case Study: Semiconductor-X 59
5.1 Use Case-Based Requirements Elicitation 59
5.2 Threat Model 64
5.3 Additional Requirements for Semiconductor-X 72
5.4 Usage Control Enforcement in Semiconductor-X 73
5.5 Gateway Design and Architecture Based on Use Case Requirements 74
5.6 Deployment Requirements 78
6 Smart Industrial Data Hub – Secure Data Provisioning and AI-Assisted Security 78
6.1 Architecture and Components 80
6.2 Data Fabric – Secure Data Provisioning 80
6.3 AI Criticality Proof – AI-Assisted Security Assessment 81
6.4 Orchestration and Applications 83
6.5 Scalable Security and Continuous Improvement 83
7 Case study: LLM Applications for secure gateway lifecycle improvements 83
7.1 AI-enhanced architecture understanding 85
7.2 AI-enhanced threat modelling 86
7.3 AI-enhanced security requirements and evolving control catalogues 88
8 Conclusion and Future Work 89
Appendix A. Standards and Regulations 93
Appendix B. Control Catalogues 105

Data Space Operator 105
Data Space Component Operator 106
Data Provider and Data Consumer 109
All Roles 110
Appendix C. Kubernetes Security Architecture Model 111
Appendix D. Miscellaneous 114
Exemplary Evidence Catalogue 114

1 Introduction

The project Semiconductor-X developed multiple use cases that depend on the exchange of sensitive data related to the production of electronic components. These use cases illustrate the need for the exchange of sensitive data across trust domains. The project uses data space technology to exchange data, facilitating the Tractus-X Eclipse Data Space Components Connector. The exchange of valuable data calls for a heightened level of security for data space components. The data exchange gateway is a central data exchange and trust negotiation component. Federation and compatibility for trust engagement call for a structured approach regarding the security model.

Manufacturing-X as a whole tries to bridge the gaps between industry sectors, using a standardized gateway pattern. One of the major challenges here is to find common ground regarding identities, protocols, and semantics. Security must be an integral part of the architecture right from the start. A solid security model helps to create and uphold trust in the ecosystem but also helps to ease legislative burdens such as conformance to NIS2 or CRA, where applicable, giving guidelines and orientation for implementers and operators.

The security concept is built on many elements, some highlighted ones are:

- We propose a three-level security concept for data exchange gateways, mapping these to data criticality levels.
- We give guidance regarding secure operations, secure deployment, development, and management of data processing components.
- We provide insights into risk management and mitigations.

We embed all of this into a holistic approach that covers all the necessary facets to create a secure data space. We will not be able to solve all problems and answer all questions, but we try to give a comprehensive overview. We define these concepts and evaluate them in a case study, taken from three use cases from the now-completed project, Semiconductor-X.

This Paper enhances the previous Paper “E2E Security Architecture of Industrial Dataspaces” of Jun 2026[1].

[1] Avgoustinos, Haris; Gossner, Harald; Jochem, Michael; Goergen, Dominik; Moser, Mathias B. et.al. E2E Security Architecture of Industrial Dataspaces, Manufacturing-X (06.2026)

2 Holistic Approach to Securing Data Spaces

A data space is a federated, decentralized ecosystem in which participants from different companies exchange data under agreed-upon rules. Only a secure data space is a successful data space – its trustworthiness is essential for broad adoption, industrial collaboration, and long-term competitiveness. A holistic security approach is therefore a technical necessity. Unlike centralized platforms, a data space has no single trusted authority that can enforce security by architectural control alone. Security must therefore emerge from a combination of verifiable technical controls, binding organizational rules, and continuous governance. Participants in a data space have access to the data space (e.g., transferring data between companies) by using dedicated gateways. Gateways are the only way to send or receive data in a data space, implementing a secure gateway is therefore an important part of securing a data space. Our holistic security approach splits into two main parts: securing the data space and securing the gateway.

2.1 Securing the Data Space

Our approach to securing a data space considers the complete life cycle of a data space and is shown in Figure 1. We divide the life cycle of a data space into four distinct phases. Different actors are involved in each phase. We describe these actors using the roles they fulfill in a data space. Section 3.1.1 describes these data space roles and their responsibilities in more detail.

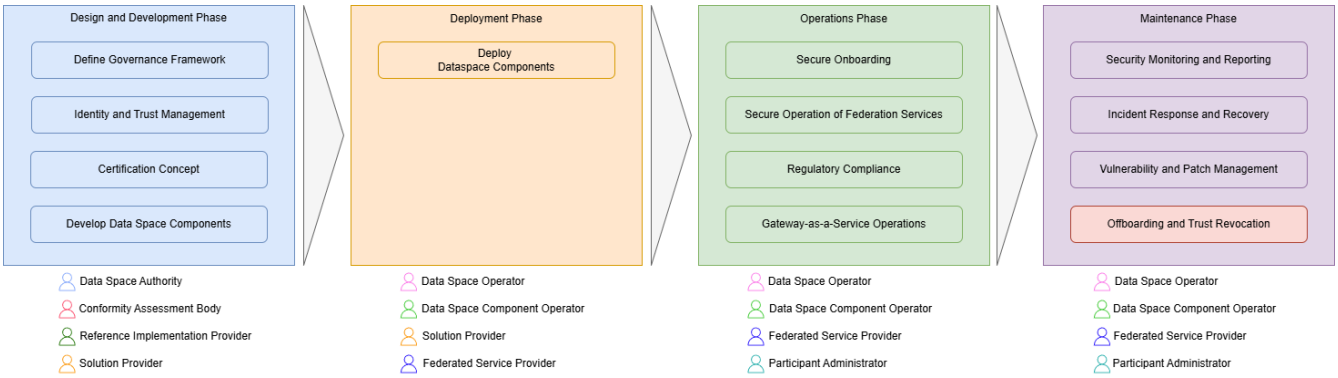


Figure 1: Data Space Security Lifecycle

The first phase is the design and development phase. In this phase, fundamental concepts need to be defined. It begins with the description of the Data space roles, their core responsibilities, typical tasks, defining characteristics, and governance relevance. These roles also serve as an introduction to governance as a set of responsibilities rather than a purely technical matter, framed around the guiding question of who may participate, under which conditions, and who controls what.

The governance framework defines and describes the four hierarchical governance levels from law and regulation down to technical enforcement, and details Level 0 as the non-negotiable legal and normative foundation. It then explains how frameworks, standards, and regulations are mapped to roles, covering data space-specific frameworks, sector standards, international norms, EU and national regulation, and extraterritorial legislation. A closing section addresses data residency and data sovereignty, including processing location, cross-border data flows, and exposure to non-EU legislation.

Identity Management and Trust Establishment discusses identity and access management and the shift toward Self-Sovereign Identities. It covers the architecture and core components (DIDs, Verifiable Credentials and Presentations, wallets, verifiable registries, DSP and DCP), the entities and identifiers used

in Semiconductor-X, and the limitation of organization-centric identity together with the delegated-issuance approach for employees, technical users, and AI agents. Further sections examine the procedural interaction of DSP and DCP, including delegation chains, trust anchors and the identity lifecycle from onboarding to revocation, cross-border legal recognition, and security considerations such as DNS dependencies, key protection, and the limits of token-based authorization and usage control.

The Certification Concept for Gateways in the context of data spaces details the procedure of how either the organizational layer or the technical components can be certified. Popular certification criteria catalogues, common security standards and best practices from other data space initiatives can be used to certify data spaces. These certification standards address roles and separation of duties, the certification lifecycle with process flow, scoping, change management, suspension and withdrawal, and software validation and deployment variants for on-premises/OT, cloud and PaaS, and operator-hosted setups.

Develop data space components comprises all steps for the implementation of software components (including the specification of required hardware to run on. These are deployed in the deployment phase, enabling deployment on dedicated server hardware, cloud infrastructures or other setups.

The operations phase comprises of the steps necessary to operate the components that, together, create a common data space. Secure onboarding includes partner registration and identity management. Other steps are the validation of the regulatory compliance of these components as well as operating gateways in a service-based model. Secure operations also includes federation services, most importantly the Wallet services.

Data Space Operations describes how the Data Space Operator fulfils its roles in daily practice across the full participation lifecycle. It covers secure onboarding and secure offboarding, including emergency containment and staged operational modes, the operation of federation services (identity provider, DID management, wallet, discovery and catalog), central security monitoring and incident response with SOC and CSIRT responsibilities, incident reporting obligations, cross-data space security coordination, and regulatory compliance under NIS2, the Cyber Resilience Act, and supply chain security requirements.

The phase model is concluded by the maintenance phase, which includes security monitoring of data space components, incident response, patch management, etc. The final step is the offboarding of participants and components.

2.2 Securing the Data Space Gateway

The second part of our holistic security approach focuses on the data space gateway, which provides a controlled interface between an organization's internal environment and the data space, enabling sovereign data exchange. The gateway is not a single component but a logical composition of four elements: Decentral Discovery Service, Identity Wallet, Connector (Tractus-X EDC), and data processing applications, running on a trusted platform with perimeter security outside its boundary. An abstract gateway security architecture is shown in Figure 3 and discussed in Chapter 4. The approach to securing the data space gateway over its life cycle is shown in Figure 2. Again, the life cycle is divided into four distinct phases. In each phase, different data space roles participate in the activities, and the active roles are listed below each phase.

The first phase, the Design and Development phase, does not start from scratch but uses as input a baseline security architecture that depends on the data space the gateway is supposed to be used for. For our approach, we use as input an abstract gateway security architecture introduced at the beginning of Chapter 4.

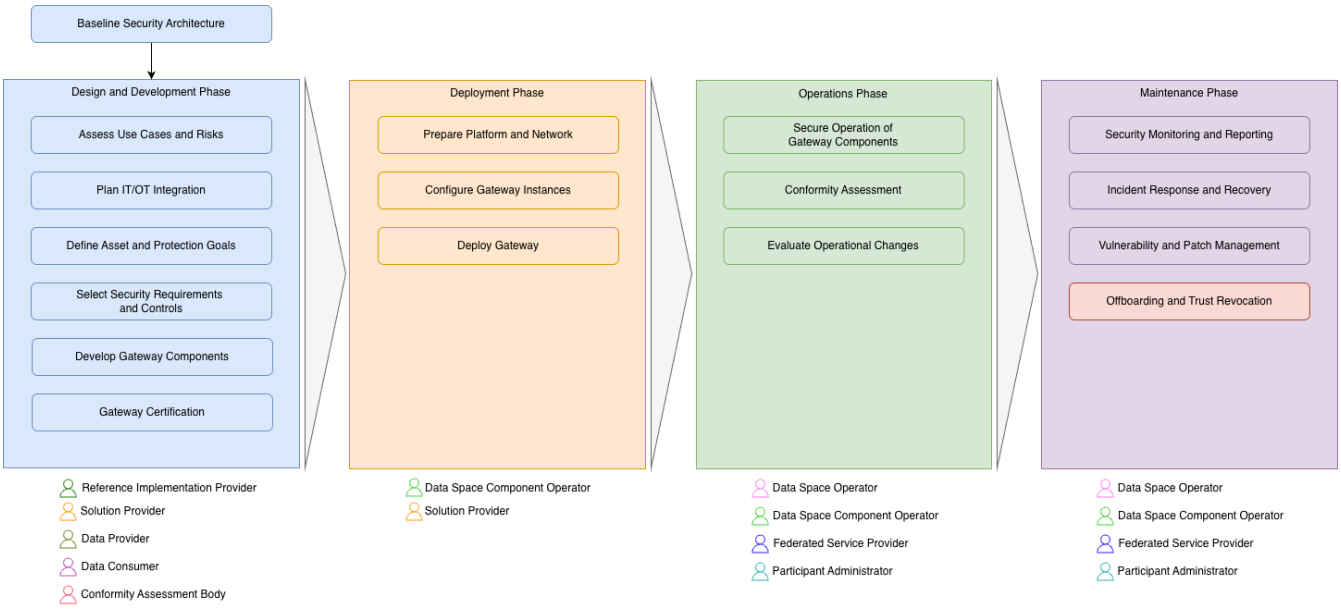


Figure 2: Gateway Security Lifecycle: From Risk-Based Deployment to Continuous Secure Operation

The threat model that was used defines three attacker classes: network attackers, honest-but-curious cloud providers, and internal attackers. Assets are grouped into data (three criticality levels: semi-public, confidential, strictly confidential), credentials, identities, infrastructure, federation services, and supporting assets. Protection goals cover confidentiality, integrity, availability, authenticity, accountability, and non-repudiation.

We define three security levels for gateways:

To reduce complexity and simplify onboarding, we base our security architecture on a **three-tiered data criticality classification**. Each data criticality is associated with a security level, which defines the security measures necessary to protect data assets of this criticality.

Data Criticality	Examples	**Security Level**	Security Characteristics
Semi-public	Eco Pass Kit, Digital product passports (metadata	**Basic**	Lightweight protection, high interoperability
Confidential	Purchase Orders, Commercial Documents.	**Trust**	Policy agreements, data encryption, system integrity verification, auditability
Strictly confidential	IP (blueprints, recipes, algorithms)	**Critical**	Hardened nodes, technical policy enforcement, secure tunnels, hardware-backed attestation, continuous monitoring

The graduated model provides a "one-fits-all" solution where participants can start at Basic and upgrade later. For strictly confidential data, protection must extend across all three states (at rest, in transit, in use) through confidential computing, TEEs, Transparent Memory Encryption, and remote attestation.

The EDC alone cannot cover all security aspects — it requires compensating controls such as firewalling/network segmentation, and monitoring.

Secure development follows an OSS-specific secure development lifecycle (SDL) with build provenance, reproducible builds, signed commits/releases, and hardened CI/CD pipelines. Risk assessment prioritizes the STRIDE framework for threat modeling.

Nine formal security requirements are defined, independent of criticality level: authentication, usage control, data protection in transit/at rest/in use, platform integrity, proof of integrity, key management, and availability.

Deployment uses a five-tier network architecture with Zero Trust, no direct OT access (via OT-DMZ), and control/data plane separation. The gateway platform requires a hardware trust anchor (TPM/HSM). Kubernetes deployments require signed images, read-only root filesystems, default-deny network policies, and RBAC.

Gateway constellations scale with data sensitivity: shared deployment (low), separate connectors (medium), separate physical platforms (high), with downward compatibility only.

Monitoring and patching are mandatory and a condition of data space membership, with timeframes derived from CRA/NIS2. A reference stack includes kube-bench, kubescape, and kube-hunter for Kubernetes hardening.

Operations cover the control plane (OAuth2/JWT, tamper-evident audit logs, multi-tenancy) and data plane (no persistent storage, short-lived tokens, vault secrets with ≤90-day rotation, per-connector distinct identities with HSM/TPM keys).

Federation services include the Identity Provider (DNSSEC-protected did:web), wallet with non-export principle, rate-limited Discovery Service, 24/7 SOC, and CSIRT authorized to block malicious entities.

2.3 Case Study: Semiconductor-X

To illustrate how our approach to securing the data space gateway can be applied, we include a case study on Semiconductor-X. We show how each activity in the lifecycle was applied in Chapter 5 to secure a gateway that supports representative use cases from the semiconductor industry.

3 Data Space Security Approach

Security in a data space that spans many independent organizations cannot be achieved by technical measures alone. Clearly defined roles and responsibilities are needed as a basis to address governance issues, identity and trust management, and define a certification process for data space components. This chapter details our approach to securing the data space lifecycle. In the following, we discuss all the activities in our approach during the individual lifecycle phases.

3.1 Design and Development Phase

We begin with the definition of the roles and responsibilities in a secure data space. These roles are used later in the discussed activities, i.e.:

- governance and its impact on the data space
- identity management and trust establishment
- certification concepts for data space components

While developing data space components is an important activity in this phase, we will not discuss this activity here, but refer to Section 4.1.6. There, the secure development of a data space gateway is discussed, which serves as a representative for data space components in general.

3.1.1 Data Space Roles

Security in a data space depends on a clearly defined set of responsibilities. Governance therefore establishes the rules, roles, and responsibilities that ensure the data space remains trustworthy, that all participants act according to the same rules, and that data is shared in a sovereign, legally compliant, and interoperable manner. This raises the central question: "Who is allowed to participate, under which conditions, and who controls what?"

In Semiconductor-X, these responsibilities are mapped onto the governance levels: the overarching legal and normative framework, federation/data space governance, participant governance, and technical/operational governance. Derived from these levels, the collaboration in the data space is structured by a defined set of roles, each bundling concrete responsibilities, typical tasks, and defining characteristics. The roles span three levels of responsibility: the data space/federation level, the participant level, and the technical/operational level (see Table 1).

It is worth emphasizing that a single organization may assume several of these roles simultaneously in practice. For instance, an Operating Company (e.g., Cofinity-X) may assume multiple roles within the data space simultaneously: 1. Data Space Operator, 2. Federation Service Provider, 3. Data Space Component Operator, and 4. Solution Provider / Customizing Operator.

Regardless of such consolidation, every role is described separately in the remainder of this section, because each carries its own distinct responsibilities, trust requirements, and potential security implications. The standards, norms, regulations, and legal frameworks that must be considered within this governance structure are catalogued in Appendix A.

Table 1: Data space roles organized by responsibility level

Responsibility Level	Roles	Short Description
Data Space / Federation	Data Space Authority	Defines governance rules, participation conditions and compliance requirements for the data space.
	Data Space Operator	Operates and coordinates core data space services and enforces governance rules operationally.
	Conformity Assessment Body	Independently audits and certifies participants and components against data space rules.
	Federation Service Provider	Operates federated trust services such as identity, catalog, policy and monitoring infrastructure.
Participant Level	Data Provider	Shares data and defines usage policies while retaining full data sovereignty.
	Data Consumer	Uses shared data strictly according to agreed contracts and usage policies.
	Participant Administrator	Manages onboarding, certificates, wallets and governance contact for an organization.
Technical / Operational	Data Space Component Operator	Operates technical data space components (e.g. gateway) ensuring security, availability and (partially) policy enforcement.

	Reference Implementation Provider (e.g. Tractus-X)	Develops and maintains standard-compliant reference implementations for data space components.
	Solution Provider	Integrates and customizes reference implementations for specific organizational deployments.

3.1.1.1 Data Space Authority

The Data Space Authority constitutes the regulatory and normative instance of the data space (e.g., Catena-X in the automotive domain). It defines the overarching order and the trust framework that every participant and every operator must adhere to. The Authority does not act as a service provider or data holder; its influence is exercised exclusively through rule-setting and decision-making.

The Data Space Authority may be organized in different governance forms. It can be a single central entity (e.g., a dedicated operating company or association) or a consortium/federation comprising all participating organizations. In the federated model, every participant holds a stake and a voice in the governance of the data space.

Its responsibilities cover the full normative spectrum of the data space:

- Definition of the data space rules and governance principles
- Definition of the participation conditions for organizations
- Definition of security and compliance requirements
- Definition of minimum requirements for data space components and of the derived certification rules
- Overall responsibility for the compliance and integrity of the data space
- Further development of the governance rules and strategic steering of the data space
- Approval of new use cases, standards, and extensions

The Authority's mandate materializes most visibly in the rulebook: it is issued and continuously maintained by the Data Space Authority, which also defines the permitted use cases and standards of the data space. Beyond this standing normative work, the Authority serves as the final escalation and decision-making instance for governance-related matters.

By design, the Data Space Authority deliberately remains "data-free": it does not hold any data and does not operate data processing services. This separation is a deliberate safeguard against conflicts of interest. The rule-setter must never become a party to the traffic it regulates. The credibility of the Authority rests solely on its rule-setting and decision-making authority, which itself must remain within the boundaries of the overarching legal and normative framework.

3.1.1.2 Data Space Operator

While the Data Space Authority defines what must apply in the data space, the Data Space Operator ensures that it does. It is responsible for the operational execution of the data space and thereby translates the normative framework into daily practice. Its core responsibilities are the operation or coordination of the central data space services, the implementation of the rules defined by the Data Space Authority, and the assurance of a stable, secure, and reliable operation.

Typical tasks arising from this mandate include the technical operation of the federated core services, the onboarding of new participants, and the operational enforcement of the governance requirements. It is thus the Operator through whom governance decisions become tangible for participants.

Like the Data Space Authority, the Operator is deliberately restricted: it is bound by the rules defined by the Data Space Authority and has no access to the participants' business data. This separation of rule-setting and rule-execution prevents a concentration of power within the data space and preserves the data sovereignty of its participants.

3.1.1.3 Conformity Assessment Body

The Conformity Assessment Body is the overarching, independent, and neutral entity responsible for verifying and confirming that the entire data space complies with the rules and frameworks established by the Data Space Authority. It thereby closes the governance control loop: the Data Space Authority defines the rules, the Data Space Operator implements them, and the Conformity Assessment Body independently verifies that both occur as intended.

In this capacity, the Conformity Assessment Body is responsible for certifying participants and data space components, verifying compliance with the minimum requirements for data space components, and ensuring adherence to the data space's security, governance, and compliance requirements. To avoid conflicts of interest, specifically, an entity certifying itself, this function is structured within the scheme into two strictly separated sub-roles with mutually exclusive tasks:

Certification Body

Owns the catalogue of criteria, issues and revokes certificates, and maintains the trust list. It makes the certification decision based on evidence provided by the evaluation facility but is not permitted to host gateways itself.

Evaluation Facility

Conducts the actual evaluations: audits, penetration tests, conformity checks, spot checks, and maturity assessments. It decides on exceptions and signs off on evaluation reports but is not permitted to supply the evaluated software itself.

Certification is not to be viewed as a one-time action, but rather as a lifecycle instrument: conformity must be continuously maintained and can be revoked in the event of non-compliance, technically implemented within the scheme via suspension and withdrawal, whereby a gateway is immediately removed from the data space (rather than merely administratively). The governance relevance of this role is fundamental: without a neutral entity to independently verify that the operating company and data space have been implemented as intended by the Data Space Authority, trust cannot be reliably established. Verifying conformity strengthens this trust and forms the basis for effective governance enforcement across the entire data space.

3.1.1.4 Federation Service Provider

The Federation Service Provider, which may, but does not have to, be identical to the Data Space Operator, operates the federated trust services on which the everyday interactions in the data space rely. To ensure the reliable development and operation of these federated services, additional roles are involved, including

Developers, Maintenance Providers, and Operators. Developers are responsible for the initial software development, including risk analysis and ensuring compliance with relevant standards and requirements. Maintenance Providers support the lifecycle of the services by managing updates, handling change management processes, and responding to incidents. Operators, finally, are responsible for the deployment and day-to-day operation of the services, including monitoring, operational risk management, and incident response to guarantee stable services and secure service availability.

The Federation Service Provider is responsible for:

- Identity Provider / Root of Trust
- Discovery / Service Catalog
- Federated Monitoring Services

In governance terms, the role is characterized by neutrality, the absence of any data access, and the establishment and operation of the technical trust anchors of the data space. This neutrality is essential: the identity, discovery, and monitoring infrastructure determines who is who and what exists within the data space, so a compromised or partial federation service would silently undermine every subsequent trust decision made on top of it.

3.1.1.5 Data Provider

The Data Provider is the role through which value actually enters the data space. It shares data and defines the access and usage policies governing who may access this data and how it may be used, while retaining full data sovereignty over everything it publishes. Sovereignty is thereby not a secondary effect but the guiding principle. Every further processing step by other participants remains subject to the conditions set by the provider.

Operationally, the Data Provider decides which data is shared with which participants, defines the corresponding policies and usage conditions, and operates its own data sources on the IT or/and the OT side. In addition, it operates or consumes data space components such as connectors, Digital Twins/Asset Administration Shells, or the Data Hub, whether as self-managed instances or as SaaS offerings.

This role carries clear governance obligations: compliance with the data space rules, correct semantic usage of the agreed standards, and reliable data quality. A provider that publishes data in a non-conformant semantic model or of poor quality undermines not only its own credibility but also the credibility and ultimately the interoperability of the entire data space.

3.1.1.6 Data Consumer

The Data Consumer is the counterpart of the Data Provider. It uses shared data strictly in accordance with the agreed contracts and usage policies, and bears responsibility for any onward use, transformation, or disclosure of that data staying within those bounds.

From a governance perspective, this role is bound by a set of obligations that operationalize data sovereignty on the consumer side: adherence to purpose limitation, meaning that data may only be used for the purposes agreed with the provider; the need-to-know principle, which restricts data access to what is genuinely required; and the obligation to enable audits, so that compliance with the usage policies can actually be verified.

Technically, these obligations are not voluntary commitments. The technical enforcement of policies takes (mostly) place via the connector, which ensures that data can only be accessed and used within the boundaries agreed by its provider. The Data Consumer therefore experiences governance not as an abstract rulebook but as a concrete technical gate at every data exchange.

3.1.1.7 Participant Administrator

Within each organization registered in the data space, at least one Participant Administrator must exist. The Administrator is a person who acts as the designated point of contact for this company. The role bundles everything that keeps an organization's data space presence operational: the onboarding of the organization, the management of certificates, wallets, and connector configuration, and the function as the central point of contact for governance-related topics.

The role hence acts as the link between technology, organization, and governance. Because administrative tasks have an immediate security dimension (they touch identities and trust anchors), the Participant Administrator is effectively the first line of accountability for keeping the organization's technical footprint in line with its governance obligations.

3.1.1.8 Data Space Component Operator

At the technical/operational level, the Data Space Component Operator describes who actually operates the data space components. This is governance-relevant precisely because this is where rules become technically effective. The role may be assumed by the participating organization itself or by an external provider, for example, a SaaS provider or the operating company, and it typically combines the operator and maintenance provider functions: deployment and operation of the components, including monitoring, operational risk management, update and change management, and incident response.

The range of components is extensive and includes the data space gateway (see Chapter 4), which is the entry point to the data space for Data Providers and Data Consumers. Across all of these components, the Component Operator is responsible for ensuring availability, security, and scalability; for implementing and enforcing policies and security requirements; and for logging, monitoring, and auditability. This also includes the secure operation of the underlying platform itself, meaning a securely hardened deployment (see Section 4.3.1*)* controlled change, update, and patch processes, as well as continuous security monitoring and coordinated incident response, which can be federated at data space level through a Security Operation Center.

For participants whose components are operated by external providers, such federated monitoring is deliberately offered as a voluntary service rather than a mandatory obligation. Federated monitoring is limited to metadata, such as which endpoints communicate with each other, connection timing, or traffic volume, rather than the content of the exchanged data itself. Still, making it compulsory would require insight into the operational patterns of third-party systems and could expose intellectual property, for example details of proprietary equipment and processes, thereby eroding the data sovereignty the data space is meant to protect. Where participants decline federated monitoring, a minimum security level is nonetheless ensured through the baseline mechanisms of this chapter, in particular certification and conformity assessment, which mirrors the responsibilities of the Certification Authority.

The governance significance of this role can hardly be overstated: policies are technically enforced here, security controls become measurable and verifiable, and misconfigurations have direct governance impact.

A wrongly configured connector is not merely a technical defect but a governance breach, making the operator the role in which the effectiveness of the entire governance framework is ultimately decided. It should be noted, however, that operators do not hold business data sovereignty: they operate the components through which sovereign data flows, but they do not own this data.

3.1.1.9 Reference Implementation Provider

The Reference Implementation Provider, in practice typically an open-source initiative such as Tractus-X, develops and maintains standard-compliant reference implementations for data space components. Beyond the pure development and maintenance of these implementations, the role is responsible for ensuring standard conformity and interoperability of the software it provides.

Its governance relevance stems from a simple fact: governance rules are translated into software by this role. Reference implementations tend to become de facto standards for the data space, and security and policy concepts are concretized here first. The quality and governance fidelity of these implementations therefore directly shape how the data space behaves in reality.

Note – Cyber Resilience Act (CRA)

With regard to legal responsibility, an important distinction applies under the European Cyber Resilience Act (CRA). Reference implementations in data spaces are typically provided as open-source software, and open-source projects maintained by foundations or communities (e.g., the Eclipse Foundation) act as "Stewards" responsible for maintaining secure development practices and the integrity of the project. However, when a company incorporates such open-source components into a commercial product, platform, or operational environment within the data space, that company becomes the "Manufacturer" under the CRA and is responsible for the compliance and security of the resulting product or system. Responsibility for CRA compliance therefore typically lies with the organization that deploys or commercializes the reference implementation, not with the open-source project itself.

3.1.1.10 Solution Provider

The Solution Provider, for example, the participating company itself, the operating company, or the data space component operator, is responsible for adopting and integrating reference implementations into an organization-specific setup, including the underlying risk analysis and ensuring compliance with relevant standards and interoperability requirements.

Its responsibilities accordingly cover the adoption and integration of the reference implementations and their customization and extension for organization-specific requirements. In doing so, this role bridges the gap between the standard and the real-world implementation: it is where the carefully governed reference implementation meets the heterogeneous realities of individual organizations.

This proximity to the reference implementation is also what makes the role governance-critical. The Solution Provider is bound by the data space rules, must not weaken security controls or policy enforcement mechanisms, and must prove conformity, for example via audits or certification. Deviations from the reference can introduce governance risks, which is why participants may adapt implementations to their specific organizational or technical requirements only within defined limits: any substantial deviation may require a conformity assessment or certification to ensure continued compliance with the data space governance rules. The degree of allowed deviation and the corresponding certification requirements are

typically defined within a dedicated certification framework or trust establishment concept. The CRA explicitly applies to organizations acting as Solution Providers within the data space (see also Section 3.1.1.9).

3.1.2 Governance

While Chapter 3.1.1 described the roles that carry governance in a data space, this chapter outlines what governance itself encompasses: the normative levels that structure the data space, the legal foundations that underpin them, and the frameworks and regulations that participants and operators must align with.

3.1.2.1 What is "Governance" in a Data Space?

Governance comprises the rules, roles, and responsibilities that ensure the data space remains trustworthy, that all participants act according to the same rules, and that data is shared in a sovereign, legally compliant, and interoperable manner. Hence, governance answers the guiding question of every data space: "Who is allowed to participate, under which conditions, and who controls what?"

Governance in a data space is documented in two complementary artifacts. The role model, describing all roles within the data space, including their tasks and responsibilities, is laid out in Section 3.1.1. In addition, Appendix A: Standards and Regulations provides an overview of all relevant standards, norms, regulations, and legal frameworks that must be considered within the data space governance. Together, these artifacts form the normative backbone against which participation, component certification, and operational enforcement are measured.

3.1.2.2 Central Governance Levels in a Data Space

At a high level, the following four governance levels exist:

- Law and Regulation
- Federation / Data Space Governance (cross-participant)
- Participant Governance (per organization)
- Technical / Operational Governance (operation)

These levels build on one another in a strict hierarchy. The overarching legal and normative framework forms the mandatory foundation on which the data space governance is built; the federation-level governance translates these external requirements into data space rules; the participant governance implements them within each organization; and the technical and operational governance ensures that the rules finally become technically effective.

3.1.2.3 Overarching Legal and Normative Framework

This level lies outside the direct control of the data space, yet it is a mandatory prerequisite for its legally compliant operation and for the participation of organizations.

Level 0, the governance level, comprises national and supranational legislation, for example, data protection law (GDPR), competition and antitrust law, product liability, and export control regulations, as well as IT security and cybersecurity requirements. The latter include certifications for standards such as ISO/IEC 27001 or BSI IT-Grundschutz and norms such as NIS2 and the Cyber Resilience Act. In addition,

Level 0 encompasses sector- and industry-specific regulatory requirements, for instance for the automotive, semiconductor, healthcare, or energy industries, which carry their own compliance obligations including certifications for standards such as ISO/IEC 27001, TISAX, or IEC 62443 (see Appendix A: Standards and Regulations).

The defining characteristics of this level follow from its position above the data space: it is non-negotiable, it is not definable by the data space itself, and it must be accepted and complied with by all participants. In this sense, Level 0 forms the legal basis for participation in an industry-specific data space.

3.1.2.4 Overview of Governance & Regulatory Framework

Establishing a secure and trustworthy data space requires alignment with a broad set of standards, regulations, and governance frameworks. These span multiple layers, from data space-specific architectures and industry standards to international norms and EU regulatory requirements, and they apply differently depending on the role a given entity assumes within the data space ecosystem.

Appendix A: Standards and Regulations provides a structured mapping of relevant frameworks, standards, and regulations to the roles defined in Chapter 3.1.1. It covers:

1) Data space-specific frameworks, such as IDSA and Catena-X standards, which define the architectural and governance foundation of the data space
2) Sector-specific industry standards, addressing the particular requirements of semiconductor and electronics manufacturing environments
3) International standards, including ISO/IEC families relevant to information security, supply chain trust, and digital twins
4) EU and national regulatory frameworks, governing data protection, cybersecurity, and digital markets
5) Extraterritorial legislation, which may affect data sovereignty depending on the infrastructure providers involved

The mapping highlights that governance obligations are by no means evenly distributed: data space-level roles such as the Authority and the Operator primarily steer and enforce the rulebook, while participant-level and technical roles such as Data Provider, Data Consumer, Component Operator, and Solution Provider carry the direct compliance obligations in day-to-day operation.

3.1.2.5 Data Residency and Data Sovereignty

The sovereignty of the data space does not end at the technical control plane but extends to where data is processed and stored. Data processing and storage must comply with applicable data protection and sovereignty requirements, which for all roles include GDPR compliance with a lawful basis for processing, privacy by design, data minimization, and purpose limitation, as well as the requirements of the EU Data Act regarding data portability, access rights, and fair data sharing. Technology and data transfers are additionally subject to export control regulations, and the Data Space Operator observes competition and antitrust law with a view to non-discriminatory access to the data space.

Extraterritorial legislation such as the US CLOUD Act and the PATRIOT Act becomes relevant whenever US-based infrastructure providers are involved, since such providers may be legally obliged to disclose stored data regardless of the contractual situation between the data space and its participants. Concrete measures to reduce this exposure include the operation of production data space services in EU data

centers with a preference for EU cloud regions, the documentation of all data flows, especially cross-border ones, the assessment of US-based cloud providers regarding their legal exposure, data processing agreements with all subprocessors, data minimization by design in onboarding and federation services, and defined retention and deletion policies. Country-related risks are addressed during onboarding through the country risk assessment and sanctions and watchlist screening described in Chapter 3.3.1.

3.1.3 Identity Management and Trust Establishment

Identity and Access Management (IAM) is a framework of policies and technologies that enables an organization to manage digital identities and control user access to resources. It ensures that the right people (human users) and things (technical users) have the right access to the right resources at the right time and for the right reasons. A robust and secure IAM system depends on trust, especially on how trust can be established. This section reviews four key challenges of secure identity and trust management in a data space and discusses possible solutions: which types of identities exist, how identities are issued and onboarded, how they are verified, and how they can be revoked when they become invalid.

Currently, data space identity management is shifting towards Self-Sovereign Identity (SSI). SSI is a decentralized approach to digital identity that gives individuals and organizations specific control over their digital identities. Instead of relying on centralized authorities (like governments or corporations) to manage and store identity data, SSI empowers individuals to create, own, and manage their own identifiers and credentials. This approach requires new foundations, including decentralization, so that identities are not controlled by a central trusted provider. It also enables each entity to decide which information to share with other entities by using Verifiable Credentials (VCs) as cryptographically secure attestations of attributes. Also, entities have Persistent Identifiers that they control and can maintain across different services.

3.1.3.1 Architecture and core components

Applying SSI identity concepts in a data space requires the interaction of specific core components. Decentralized Identifiers (DIDs) are unique digital identifiers created using asymmetric cryptography, where the owner generates a private/public key pair and uses the private key to prove control over the DID, while the public key is published in the DID document for others to verify that proof. They are registered and resolved using a specific DID method. For example, using *did:web*, the DID is tied to a web domain controlled by the owner, and the DID document is hosted over HTTPS. Unlike X.509 certificates where the subject may generate the key pair but typically relies on a Certificate Authority (CA) to validate identity and issue a certificate binding the public key to that identity, a DID allows the subject to control the identifier and demonstrate proof of control directly through the DID method without requiring a centralized certificate authority who gives the identifier its trust and binding. Based on DIDs, a VC is a digital credential that contains claims about a person, organization, or entity and are cryptographically signed by an issuer. For presentation, based on the VC a Verifiable Presentation (VP) is generated. A VP is a set of information of a VC, signed by the Holder of a VC, such that a verifier can verify this set of information, by only receiving knowledge for this set of information, but not any other information. A VC is typically issued by the DID of the issuer and issued to the DID of the holder or subject, so the credential is linked to both the issuer and the credential subject(holder). Like an X.509 certificate, a VC is a signed data structure that allows a verifier to check the authenticity and integrity of the contained information. Both can include attributes and claims about a subject. However, the primary purpose of an X.509 certificate is to bind a public key to an identity within a PKI trust hierarchy. Whereas the primary purpose of a VC is to make claims about a subject that can be presented as credentials.

For managing DIDs and VCs, each entity requires an SSI wallet. An SSI Wallet is a digital tool used to store, manage, and present decentralized identifiers and verifiable credentials. It gives users control over their identity data and how it is shared with others. Currently, it is planned to provide three different wallet architectures. As a first step, there will be a white-label wallet as a service provided by the Core Service Provider, or the option to use a self-owned wallet (BYOW). At some point, users will additionally be able to select their wallet from the marketplace. The tasks of the data space wallet are issuing VC, creating Self-Issued ID tokens, and querying VPs.

To verify the decentralized components, a Verifiable Registry (e.g., *did:web* URI) is required. This system is used to publish and access the public information needed to verify decentralized identifiers and credentials. It supports trust by enabling verifiers to check things like public keys, DID documents, and credential status.

For communication inside a data space, two protocols are important. First, the Eclipse Data Space protocol[2] (DSP) is designed to facilitate interoperable data sharing between entities governed by usage control and based on Web technologies. Second, the Eclipse Decentralized Claims Protocol[3] (DCP) serves as a basis for decentralized identity management. The DCP defines an interoperable overlay to the DSP specifications for conveying organizational identities and establishing trust in a way that preserves privacy. This specification defines schemas and protocols required for entities to publish data, negotiate agreements, and access data as part of a federation of technical systems termed a data space.

3.1.3.2 Entities and Identifiers in Semiconductor-X

Table 2 - Entities and identifiers

Identifier	Scope	Remarks
DID	For each company	- Offers cryptographic verifiability - self-issued - Bound to a web domain
Business Partner Number (BPN)	X-ecosystem specific	- Issued by Data Space Authority - ecosystem specific identifier for onboarded legal entities (participants)
Membership Credential	Specific to a particular X data space	- Cryptographic verifiable anchor - Verifiable Credential

Identity management in data spaces is based on the three central identifier concepts shown in Table 2 - Entities and identifiers. The Business Partner Number (BPN) is the primary identifier used across data spaces. Membership Credentials provide proof that the entity associated with a BPN is a participant in a data space and is authorized to operate a connector. The DCP/DSP standards currently focus largely on organization-level identity management. As a result, they primarily support legal entities identifying themselves to one another, but do not enable more fine-grained identity management.

These limitations create potential issues, particularly in the context of usage control and the growing use of AI agents. To enable data sharing that complies with privacy regulations and is more nuanced, there is currently a lack of representation for subordinate entities such as human employees, technical services, or increasingly relevant AI agents. To address this shortcoming, a “delegated issuance” approach must be

[2] https://eclipse-data space-protocol-base.github.io/DataspaceProtocol/2025-1-err1/
[3] https://eclipse-data space-dcp.github.io/decentralized-claims-protocol/v1.0.1/

implemented. Onboarded companies must be enabled to act as their own local issuers of Verifiable Credentials (VCs).

Through this delegation, human users can receive their own VCs that are cryptographically linked to the organizational interface (DID). This enables verifiable proof of fine-grained roles and permissions within the Data space ecosystem. Similarly, technical users and AI agents must also be regarded as machine identities. Equipping them with their own VCs makes it possible to cryptographically and unambiguously authenticate automated data queries on behalf of the company while simultaneously restricting access precisely to specific data records.

3.1.3.3 Protocol Interaction DSP and DCP

The procedural interaction within the data space is largely determined by the interplay of two protocols. The Dataspace Protocol (DSP) governs operational data usage and contract negotiation by defining standardized processes for catalog retrieval, contract negotiation, and data transfer. However, the DSP itself does not establish trust among the participating parties. This gap is filled by the Decentralized Claims Protocol (DCP), which functions as a dedicated trust layer and is built as an overlay on top of the DSP. Based on the Issuer-Holder-Verifier model, the DCP ensures that every DSP communication between connectors is preceded by a DCP interaction between the connector and the wallet. In this step, the consumer presents their Membership VC to cryptographically prove their legitimate participation in the data space. Following successful contract negotiation via the DSP, the consumer receives an access token, called Endpoint Data Reference (EDR), which authorizes the subsequent data retrieval.

Since authentication in this basic model is primarily organization-centric, fine-grained access control at the user or technical level requires extending the DCP flow to include sub-identities. This requires the implementation of a delegation chain: The Data Space Authority (DA) trusts the onboarding organization, while the organization acts as a trusted local issuer for its employees and services. The verifiable credentials (VCs) of these sub-identities must be cryptographically signed with the organization's DID, which in turn must possess a valid membership VC. When accessing specific data, the consumer connector must therefore present not only the organizational membership VC but also, depending on the context, the VC of the technical user or employee actually performing the action. This enhanced procedure enables precise provider-side verification, allowing the provider connector to enforce granular policies.

3.1.3.4 Trust and Trust Anchors

Building trust and implementing robust governance form the foundation for secure collaboration in the data space. Initially, the Data Space Authority (DA) serves as the central anchor of trust, acting as the issuer of Membership Credentials. The onboarding process is conducted through an out-of-band process, in which a new participating company submits all relevant information to the DA via a secure out-of-band channel. The DA verifies the submitted information, assigns a Business Partner Number (BPN) where necessary, and issues the Membership Credential to the new company. Importantly, the participating company must generate and stay in control of its own private keys. The DA only uses the corresponding public key to issue and bind the Membership Credential to the company's identity. The DA must never have access to the company's private keys. The validation of identities and credentials takes place in a decentralized and cryptographic manner, without requiring a real-time query to the central authority during the transaction.

Conceptually, a strict distinction must be made in this context between "trust" and "trustworthiness": While "trust" describes a participant's subjective expectation that a partner will act in accordance with the

contract, “trustworthiness” refers to the objective, measurable, and verifiable fulfillment of these expectations, such as through demonstrated compliance and adherence to governance requirements.

However, a local trust anchor, such as a single Data Space Operating Company, faces limitations when it comes to global scaling. Technical and cryptographic trust mechanisms alone are not legally valid across national borders. Therefore, involving international participants creates legal challenges, such as recognizing identities, assigning liability for misconduct, and enforcing governance rules in different countries.

To establish a globally trusted data space, the trust anchor must be integrated into cross-border, legally binding frameworks. A highly suitable candidate for this is eIDAS 2.0[4], particularly in the form of the European Business Wallet. Since this wallet already contains a great deal of corporate information (such as commercial register extracts or certifications) that is mandatory for the onboarding process, and which has been verified in a legally compliant manner by qualified trust service providers, the Data Space Authority can draw on this pre-validated data. This drastically accelerates the onboarding process and creates a legally binding basis of trust to ensure global interoperability through Mutual Recognition Agreements (MRAs).

A key aspect of the identity lifecycle is the revocation of identities. If a company leaves the Data space or violates governance guidelines, the Data Space Authority places the corresponding Membership VC on the VC Status Lists so that it is automatically rejected during future authentications by verifiers. In this case, the revoked entity is no longer part of this Data space, but the BPN of this entity is still valid and can be used in other Data spaces.

For the management of sub-identities, responsibility for revocation shifts to the participating organization itself. Since the organization acts as a local issuer, it must maintain its own status lists for the sub-VCs it issues. If an employee is terminated or a technical user becomes invalid, the organization revokes the corresponding VC by updating its local status list. During verification, providers must therefore check not only the organizational membership VC but also the validity of the presented sub-VC against the company’s status list. In addition, if the organization’s Membership VC itself is revoked by the DA, the cryptographic trust chain breaks, causing all sub-identities delegated by that organization to automatically lose their verifiability and their access rights within the Data space.

3.1.3.5 Security Considerations

In particular, the use of did:web involves dependencies on the Web PKI and the DNS system. Since DID documents are not necessarily integrity-protected, security measures such as DNSSEC or DNS over HTTPS are absolutely essential to prevent tampering. Key management is equally critical: private keys must be securely generated and stored on hardware anchors such as HSMs, TPMs, or Trusted Execution Environments (TEEs) to preserve the sovereignty of identities.

A key aspect of the security analysis concerns the guarantees provided by the Endpoint Data Reference (EDR) token. While the provider ensures through its Authorization Service that the token was issued correctly and that access to a specific asset within the context of an Agreement ID is authorized, the token explicitly offers no guarantees regarding the specific end user, the executing service workload, or the security posture of the consumer platform.

[4] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401183

This token-related limitation results in a significant restriction on usage control: Once the data has been transferred to the consumer, there is no longer any technical enforcement of usage policies. Copying the data or further processing it outside the contractually agreed-upon parameters can no longer be technically prevented.

3.1.4 Certification Concept

Numerous certification guidelines have already been developed to cover different aspects of data space certification. The following sections provide an overview of sources for certification criteria relevant and reused for data space certification guides. The certification concept is based on the approval of the corresponding certification criteria.

"White Paper Certification - FRAMEWORK FOR THE IDS CERTIFICATION SCHEME, VERSION 2"

A major source for certification criteria especially fitted for data spaces is the "White Paper Certification" created by the International Dataspaces Association (IDSA) which provides a basis for certification procedures. The important sections are explained in more detail in the next sections.

Source	https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-White-...

The IDSA Certification Scheme is a framework proposed by the International Data Spaces Association (IDSA) to assess whether data spaces and their components comply with the requirements of the IDS Reference Architecture Model (IDS-RAM). It defines a certification approach based on requirements and assessment criteria to establish trust, interoperability, and security within data spaces. The scheme provides a basis for evaluating individual data space components, such as connectors, as well as supporting the broader assessment of data space implementations.

The Section "IDS Certification Processes" describes the procedures for carrying out certification within the International Data Spaces (IDS) framework. It explains how certification activities are organized, including the roles of applicants and certification bodies, the assessment of requirements, and the steps leading to a certification decision. It distinguishes between Assurance Level 1, where applicants perform a self-assessment, and Assurance Levels 2 and 3, where an approved independent Evaluation Facility conducts the assessment. The document also covers the approval of Evaluation Facilities, certificate issuance by a certification body, and subsequent change certification and renewal.

The section "IDS Operational Environment Certification" defines how to assess whether a data space participant's organizational processes and operational environment are trustworthy. The framework combines the achieved trust levels into certification profiles suited to different use cases, with the applicable criteria and assessment methods detailed in the IDS Certification Scheme. The Trust Levels and the Assurance Levels highly relate to the ones defined in this document in Section 4.1.5.1.

This corresponding criteria catalog "Criteria Catalogue: Operational Environment - White Paper 2020 - Version 1.0" can be requested at the IDSA at https://internationaldataspaces.org/offers/certification/ and can be used as certification criteria. It is based on ISO/IEC 27001, IEC 62443, Cloud Computing Compliance Criteria Catalogue (C5), CSA Cloud Controls Matrix, and extended specifically for data spaces. Since these sources are often referenced in other certification criteria, the following list provides an overview of each of these sources.

ISO/IEC 27001	"ISO/IEC 27001 is the world's best-known standard for information security management systems	https://www.iso.org/...

	(ISMS). It defines requirements an ISMS must meet."	
IEC 62443	"The ISA/IEC 62443 series of standards define requirements and processes for implementing and maintaining electronically secure industrial automation and control systems (IACS). These standards set best practices for security and provide a way to assess the level of security performance."	https://www.isa.org/...
Cloud Computing Compliance Criteria Catalogue (C5)	"The C5 (Cloud Computing Compliance Criteria Catalogue) criteria catalogue specifies minimum requirements for secure cloud computing and is primarily intended for professional cloud service providers, their auditors and customers."	https://www.bsi.bund.de/...
CSA Cloud Controls Matrix	"The Cloud Controls Matrix (CCM) is a cybersecurity control framework made up of 207 controls across 17 security domains."	https://cloudsecurityalliance.org/...

Catena-X Conformity Assessment Framework (CAF)

The dataspace initiative Catena-X from the automotive domain has also published a certification catalog. The official handbook describes the normative framework governing certification, including the certification roles, modular system, certification catalog, test procedures, certification process, and certificate lifecycle.

Source	https://catena-x.net/wp-content/uploads/2026/03/CatenaX_CAF_Handbook_Eng...

Using certification criteria catalogues to establish a suitable programmatic approach for a certification decision in data spaces is considered best practice and recommended in the context of the Semiconductor-X Data Space.

3.1.5 Developing Data Space Components

The processes and best practices for secure development of data space components are equivalent to those for secure gateway component development. They are discussed in Section 4.1.6.

3.2 Deployment Phase

This section briefly outlines the deployment of data space services by Data Space Operators. It is derived from the Catena-X Operating Model. The deployment of data space services and components must be reproducible, version-pinned, verifiably sourced and hardened by default, operated in a legally compliant environment, with production strictly isolated and every change traceable.

Because core services form the root of trust for the entire data space, there must be clear expectations on where and under which legal conditions they may be operated, e.g., within the EU or EEA, and the Data Space Operator must be able to document and disclose all processing locations. Physical security of the underlying data centers is normally evidenced indirectly through recognized provider certifications, e.g., according to ISO/IEC 27001.

3.2.1 Deploying Data Space Components

Describing this process for multiple data space components is out of scope of this document, but the deployment of data space gateways by Data Space Component Operators is covered in detail in Section 4.2. As such, Section 4.2 serves as a general guideline for deployment, with the caveat that some details may vary depending on the actual data space component to be deployed. Here, we only briefly describe some important aspects of the deployment process for data space components.

3.2.1.1 Deployment Model and Artifact Integrity

Data Space Operators are expected to operate reference implementations of services rather than custom re-implementations unless an alternative has been separately certified for conformance. Only officially released, standard-conformant versions may run in the productive data space. Deployed components are pinned to defined versions so that any production state can be reconstructed and attributed.

Changes to a running service follow a documented change and release process. Changes are risk-classified and formally approved, maintenance windows are announced to participants in advance because downtime of a central service affects the whole data space, and every change must have a tested rollback path.

3.2.1.2 Secure Software Supply Chain

Central services are attractive targets, so the integrity of the path from source code to running service must be safeguarded. Only verified and signed artifacts from trusted sources are deployed, and each release is accompanied by a software bill of materials (SBOM) that provides component transparency and allows rapid impact assessment when new vulnerabilities become known.

3.2.1.3 Hardening of the Runtime Platform

Runtime components should be kept minimal, run with the least privileges necessary, and rebuilt regularly so that security patches actually reach production. Resource boundaries are enforced to contain both operational side effects and simple resource-exhaustion attacks.

At the platform layer, access rights follow least privilege, administrative interfaces and data stores are never exposed directly to the public internet, and the environment is segmented so that components can only reach what they genuinely need. Hardening baselines are enforced automatically by platform policy rather than relying on reviewer discipline. Operating systems and nodes are based on hardened, regularly patched images, and public-facing entry points are protected by current transport encryption together with filtering and rate-limiting measures.

3.3 Operations Phase

While Chapter 3.1 defined the roles that carry the data space, this chapter describes how the Data Space Operator fulfils them in day-to-day operation. The Operations Phase covers participant onboarding and the secure operation of core data space components in accordance with regulatory requirements. The Data Space Operator acts in these functions partly in its own responsibility and partly in the role of the Federation Service Provider as defined in Section 3.1.1. The complete set of operational requirements is consolidated in the control catalogue in Appendix A Standards and Regulations.

3.3.1 Secure Onboarding

The onboarding process establishes the trustworthiness of a new participant before access to the data space is granted. The onboarding process is the first governance checkpoint of the data space and includes the following security measures: validation of the participant identity against trusted sources such as the commercial and tax register via a Digital Clearing House (DCH), a check against the Country Clearance List of the association, and acceptance of the framework agreement governing data exchange. A standardized security self-assessment validates the security readiness of the participant, and clear minimum security requirements define what every participant must meet, differentiated by industry, role, and security level. Following the administrative checks, membership, identity, and framework agreement credentials are securely issued as Verifiable Credentials, DIDs are created and published with DNSSEC verification, and a BPN is assigned as a unique and persistent identifier. The entire process is recorded in a complete audit trail. In addition, participants should provide proof that employees with data space access have completed security training.

3.3.2 Secure Operation of Federation Services

The Federation Service Provider operates a set of central federation services on which the trust fabric of the data space depends. Because the compromise of any of these services would have a structural impact on all participants, their secure operation is subject to dedicated requirements.

The Identity Provider serves as the central trust anchor of the data space and is thereby responsible for the complete trust lifecycle of participant and component identities. This covers PKI management, certificate issuance and management for all connector connections and the enforcement of confidentiality across the federation services through mTLS on all service-to-service communication. VCs are issued during onboarding and must be immediately revoked upon a participant's offboarding. Identity management currently relies on did:web, whose known structural weaknesses, eg. dependency on DNS and the absence of built-in integrity protection or versioning, are compensated through a defined set of measures rather than accepted as residual risk: DNS servers of did:web controllers must be protected by DNSSEC under a strict key-management policy, DID document resolution must not fall back to operating-system DNS resolvers, and DID document changes are subject to controlled publication processes and continuous monitoring. Newer DID methods, in particular did:webvh, are being evaluated to address known did:web weaknesses and to cover robust key lifecycle management, especially key revocation, which did:web does not standardize.

The wallet service is treated as a high-value target and is hardened accordingly: a secure baseline configuration with unnecessary features and endpoints disabled, strict key protection ensuring that private keys never leave the secure storage, controlled backup and recovery procedures without exporting key material, authenticated wallet APIs, and rate limiting with abuse detection. All wallet APIs enforce authentication and least-privilege authorization. Private keys for DIDs and VCs must be stored in hardware-backed security devices (HSMs or TPMs). Each connector instance holds its own individual key pair; shared identities across instances are not permitted.

The Discovery and Service Catalog infrastructure is protected against scraping and automated attacks through rate limiting, and the BPN-DID-Resolution-Service maps business partner numbers to DIDs without revealing metadata about data transfers. Importantly, participants must not be mandatorily listed in the Discovery Service or Service Catalog, preserving their sovereignty over their own visibility within the data space.

3.3.3 Regulatory Compliance

NIS2 obligations apply to the Data Space Operator and, where in scope, to Component Operators offering connectors as a service or comparable central services. They are addressed through systematic cybersecurity risk management according to ISO/IEC 27005, NIS2-compliant incident reporting along the timelines consolidated in Section 3.4.1, supply chain security measures including the assessment of relevant providers, business continuity management that explicitly covers connector outages, and management accountability up to personal liability, with entities classified as essential or important entities accordingly.

For Component Operators that distribute components commercially or as managed services, the obligations of the Cyber Resilience Act apply in the role of a manufacturer. These cover security by design and by default, a secure development lifecycle, vulnerability management with public disclosure of fixed vulnerabilities, software bills of materials, post-market surveillance, and conformity assessment. For open source-based components, the distinction introduced in Chapters 3.1.1.9 and 3.1.1.10 applies: the foundation acts as steward for secure development practices and project integrity, while responsibility for compliance shifts to the organization that deploys or commercializes the components.

Supply chain security requirements apply to both roles and complete the program: supplier due diligence including a security assessment of all third-party providers, contractual anchoring of security requirements, continuous monitoring of the Eclipse Tractus-X supply chain, assurance via SBOMs and audits, and committer identity verification at the foundation level.

Note: The underlying operating guidelines are a technical organizational best practice draft aligned with NIS2, the Cyber Resilience Act, and related regulatory frameworks. Applicability of referenced standards must be verified in each individual case, and final implementation responsibility remains with the respective organizations.

3.3.4 Gateway-as-a-Service Operations

Small and medium-sized enterprises (SME) in the supply chain, in particular, often refrain from operating dataspace connectors themselves due to a lack of security expertise and resources. Gateway-as-a-Service providers offer cloud-based connectors to such customers, enabling SMEs to access the dataspace.

3.4 Maintenance Phase

Concurrent with the operational phase, the maintenance phase encompasses continuous security monitoring of the dataspace and its constituent components. In the event of a security incident, corrective actions are triggered to maintain compliance and limit potential damage. The framework further provides mechanisms to offboard or revoke the trust status of participants identified as contributors to security incidents.

3.4.1 Central Security Monitoring and Reporting Obligations

Central security monitoring provides the cross-participant security picture of the data space. All data space components operated within the central infrastructure report security alerts to the central Security Operation Center (SOC) of the Operating Company. The central SOC monitors the general security and technical health status of the data space. Cyber Threat Intelligence is generated and made available to all participants, and monitoring runs 24/7 with defined escalation paths. For components operated externally

by participants or third-party service providers, federated monitoring remains a voluntary service as described in Section 3.1.1: mandatory insight into third-party operational data would erode data sovereignty and risk exposing intellectual property. Where attendees decline federated monitoring, the security baseline is safeguarded through certification and conformity assessment instead.

Reporting obligations in the data space reflect both internal governance and external regulation:

Table 3: Incident Reporting Obligations

Reporting Obligation	Deadline	Recipient
Internal report to security team	≤ 1 hour	SOC / CSIRT
Participant notification	Immediately after confirmation	Affected participants
NIS2 early warning	≤ 24 hours	Competent national authority
NIS2 full report	≤ 72 hours	Competent national authority
CRA vulnerability/incident report	≤ 24 h early warning, ≤ 72 h full report, ≤ 14 days final report	ENISA / market surveillance authority
Association notification	Immediately	Data Space Authority

3.4.2 Incident Response and Emergency Containment

The Computer Security Incident Response Team (CSIRT) coordinates immediate actions in case of security breaches. Its mandates include the coordination of emergency containment through temporary credential revocation, the suspension of trust relationships, and the restriction of data exchange with affected entities, as well as blocking traffic to compromised entities to minimize structural damage.

The Operating Company, the Data Space Component Operators, and the Federation Service Provider must maintain a documented Incident Response Plan according to the relevant industry best-practices (e.g., ISO/IEC 27035 and IEC 62443-2-1). The plan follows the classical incident response lifecycle: In the detection phase, security incidents are identified through SIEM, anomaly detection, 24/7 SOC monitoring and alerting. During triage, incidents are classified by severity level and assigned to the CSIRT. The containment phase comprises the isolation of affected systems, the blocking of affected identities and the blocking of traffic to compromised entities. In the eradication phase, the threat is removed and systems are patched. Recovery restores the affected services and verifies their integrity. The post-incident phase captures lessons learned, drives the improvement of security measures, and ensures forensic readiness.

Unlike centralized cloud platforms, a federated data space cannot be shut down through a single technical mechanism, since no single actor holds full technical control over all participant systems. The Operating Company must therefore respond at the trust and interoperability levels instead. The available response measures include the temporary credential revocation of compromised participants or connectors, the suspension of trust relationships, the restriction of data exchange with affected entities, and the distribution of trusted security advisories to all participants, with defined escalation paths governing the handling of high-severity situations. The data space operates in staged operational modes: under normal conditions, regular monitoring and policy enforcement apply; in the elevated risk mode, additional monitoring and stricter policy enforcement are activated; and in the extreme mode, vulnerable services or connectors can

be temporarily excluded from the data space. All emergency measures must be formally documented, time-limited, and auditable.

3.4.3 Business Continuity and Disaster Recovery

Operational resilience is ensured through high availability and disaster recovery measures: all production services run in multi-availability-zone configurations with load balancing, identity services feature automatic failover with tested failover processes. Backups are encrypted and georedundant, following the 3-2-1 principle with daily incremental and weekly full backups. Resilience is verified continuously through chaos engineering, tabletop exercises, and, where applicable the analysis of OT failure modes, and graceful degradation allows partial operation of the ecosystem when individual components fail. The business continuity plan explicitly covers connector outages and alternative data exchange paths.

Because the data space is operated as a federated data space, the business continuity plan must additionally account for failure modes that arise specifically from federation. This includes the failure or compromise of federation services such as the Identity Provider, the Discovery Service, and the Monitoring/SOC infrastructure and the staged trust-level containment measures described in Section 3.4.2. Recovery exercises must therefore validate not only technical restoration of services, but also the operational handling of trust-level containment actions such as credential revocation and the suspension of trust relationships.

In this chapter, we specify three exemplary use cases, perform an in-depth threat analysis before defining an exemplary gateway fulfilling both general and derived use-case specific requirements.

3.4.4 Vulnerability and Patch Management

Component Operators with operational responsibility maintain a structured vulnerability management process. This encompasses continuous CVE monitoring for the connector, the wallet (e.g. Identity Hub), and all dependencies, SBOM management for all developed components, a defined coordinated vulnerability disclosure process for reporting and handling vulnerabilities, and a risk-based patching SLA for developed components that prioritizes fixes according to exposure and criticality. For all developed data space components (responsible roles: 3.1.1.9 and 3.1.1.10), post-market surveillance with defined processes and responsible bodies (PSIRT) ensures that vulnerabilities are tracked, communicated and remediated beyond the initial release. Patches are signed for authenticity and integrity verification, and security fixes are backported to still-supported older releases. The specific SLA thresholds and disclosure timelines are not prescribed by this scheme but must be derived from the applicable regulatory framework, such as the Cyber Resilience Act (CRA) or NIS2, depending on which requirements apply to the component operator, reference implementation provider or Solution Provider.

Vulnerability management is closely coupled with the central monitoring infrastructure of the data space. Vulnerability findings and security alerts on the component side must be reportable to the central SOC of the Operating Company (see Section 3.4.1). In return, Cyber Threat Intelligence (CTI) generated by the SOC/CSIRT should be fed back to Component Operators to inform CVE prioritization and risk-based patching decisions.

3.4.5 Secure Offboarding

The offboarding process ensures that a departing participant is removed from the data space completely, verifiably, and without residual trust artifacts. The offboarding process complements onboarding on the exit

side and is equally security-driven. It includes the immediate revocation of all issued Verifiable Credentials via the Verifiable Credential Status List, the deactivation of the BPN, and the update or removal of the DID document. Connectors are deregistered from the Discovery Services and the Service Catalog, existing data contracts are terminated with notification of all active contract partners, and personal data is deleted in a GDPR-compliant and retention-aware manner. Wallet blocking, SOC notification, and participant notification complete the process, which remains fully auditable. In emergency situations, containment measures are documented, time-limited, and auditable, and staged operational modes, from normal operation over elevated risk with additional monitoring and stricter policy enforcement up to the temporary exclusion of vulnerable services or connectors, define a graduated response.

In case of security incidents, an emergency offboarding can be initiated. In this case, the affected entity is immediately blocked by the CSIRT, credentials are temporarily revoked, trust relationships are suspended, and the affected entity is isolated until the incident is clarified. All emergency measures must be formally documented, time-limited, and auditable.

4 Gateway Security Approach

This chapter details our approach to securing the data space gateway, which provides a controlled interface between an organization's internal environment and the data space. It enables data exchange while preserving data sovereignty, ensuring that participants remain in control over which data is shared with whom and how it is to be used, even after the data has been shared with external recipients. It is more than just a network connection; it is the central trust anchor for data sharing and establishes secure end-to-end security between participating organizations. It authenticates entities and users, verifies contracts, permissions, and access/usage policies, and (mostly) enforces these policies. Based on this information, the gateway makes decisions about the receipt, storage, processing, and forwarding of assets, and it creates and collects logs and audit information such as security events, policy decisions, and data transfer logs.

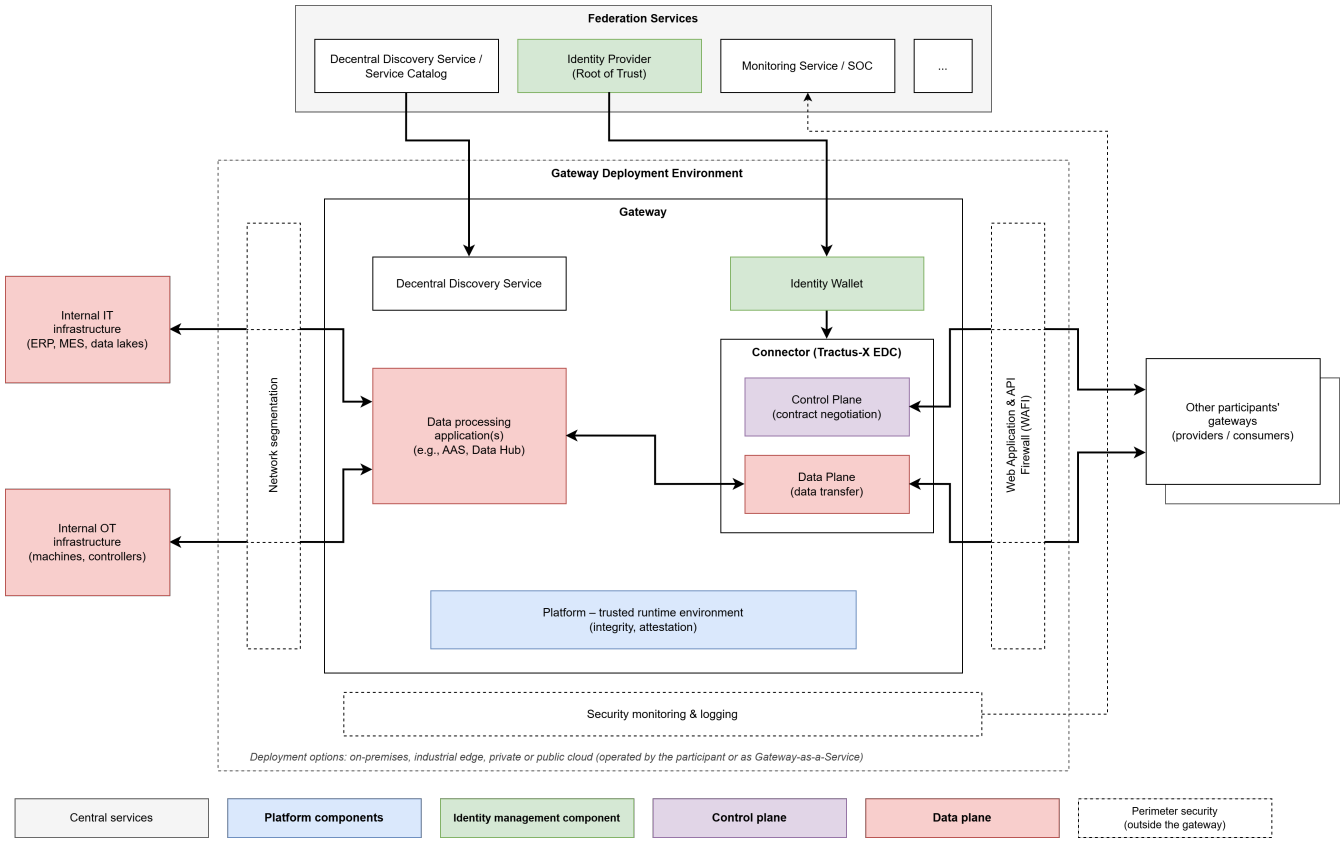


Figure 3: Abstract Gateway Security Architecture: The gateway as a logical composition of Decentral Discovery Service, Identity Wallet, Connector, and data processing application(s) on a trusted platform, protected by perimeter security measures outside its boundary.

Importantly, the gateway is not a single component but a logical composition of the components shown within the Gateway boundary in Figure 3: the Decentral Discovery Service, the Identity Wallet, the Connector, and the data processing application(s), running on a platform that provides their trusted runtime environment. Together, these components form the gateway as the participant's point of interaction with the data space. The Abstract Gateway Security Architecture shown in Figure 3 includes security measures such as web application and API firewalling, network segmentation, and monitoring services that lie outside the gateway boundary itself but contribute to protecting it from the surrounding infrastructure. The security of a data exchange depends on the correct composition of all components involved, and a compromised or misconfigured auxiliary component endangers the gateway as a whole.

Gateways can be deployed on premises, at an industrial edge, in private clouds, or public cloud environments. In detail, the gateway may be operated by participants themselves or by specialized service providers, such as Gateway-as-a-Service. Here, certain management and security responsibilities are fulfilled by the gateway provider instead of the participant. Depending on the criticality level of assets the gateway is supposed to process or share, it can be implemented at different security levels. Thus, the previously defined requirements apply to a different extent, as will be described in more detail in this chapter.

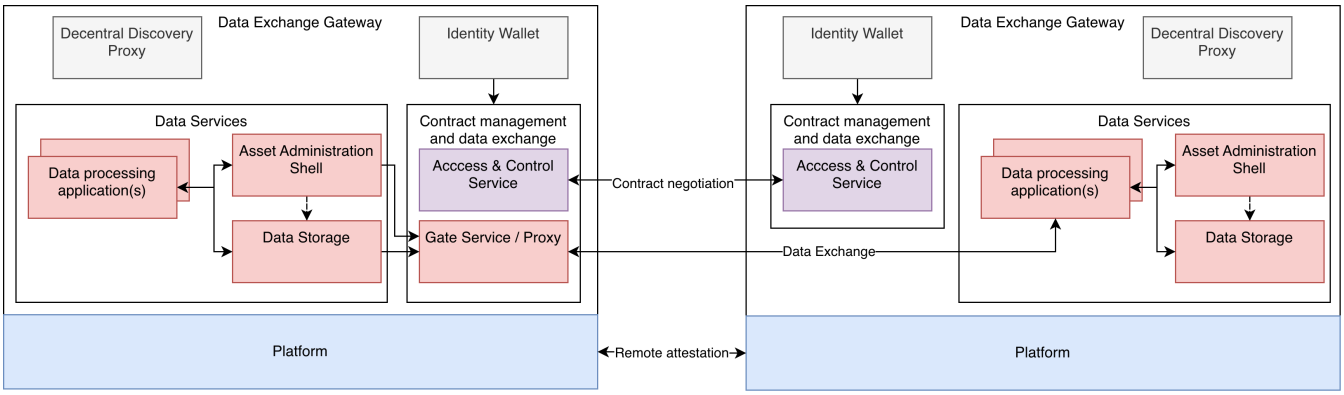


Figure 4 – Abstract overview of two data exchange gateways, depicting all needed components to create data space interaction, running on an underlying platform such as Cloud hardware or a dedicated server.

The remainder of this chapter describes the approach to securing the data space gateway over the phases of its life cycle. For the first three phases, we describe in detail each activity belonging to the phase, as shown in Figure 2 (see Section 2.2). We intentionally omit the description of the Maintenance Phase, because the activities are basically identical to the activities in the Maintenance Phase of the approach to securing the data space. We refer the reader therefore to Section 3.4 for a description of the activities in the Maintenance Phase.

4.1 Design and Development Phase

Our approach to securing the data space gateway in the Design and Development Phase includes the assessment of use cases and risks, the identification of the assets that need to be protected, the definition of the protection goals, the selection of appropriate security requirements and controls, and ensuring the secure development of gateway components. A prerequisite to this approach is the definition of the attacker model. Our approach is defined as a continuous process over the complete life cycle of the gateway, meaning that even when the project is deployed, e.g., novel threats or a change in data sensitivity may necessitate an updated analysis and new controls or patches to be adopted.

In the following, we first describe the attacker model we consider for the data space gateway, then we describe the different activities of our approach in the Design and Development Phase, i.e.:

- How use cases and risks can be assessed
- What must be considered when planning to integrate IT/OT
- How assets and protection goals are defined
- How security requirements and controls are selected
- How to enable the secure development of gateway components
- How to ensure that a gateway acts as intended

4.1.1 Attacker Model

In Semiconductor-X and in the area of data spaces in general, very few attacker models are commonly used, which are denoted in the following sections. Any other non-standard attacker models must be modeled separately. Our threat model considers three classes of attackers:

1. Network attacker. A network attacker can observe network traffic and interfere with data transmissions between participants. Potential attacks include eavesdropping, recording and replaying messages, modifying data in transit, or impersonating participants. These threats are

mitigated using established security mechanisms employed across modern Internet infrastructure, including HTTPS/TLS and an appropriate PKI infrastructure.

2. Honest-but-curious cloud provider. Data space components are typically deployed on infrastructure operated by third-party cloud providers rather than on hardware directly controlled by the respective participant. Consequently, participants either need to trust the cloud provider or employ confidential computing technologies, such as TEEs, which provide hardware-protected execution environments for processing sensitive code and confidential data on remote systems. Our threat model assumes that the cloud provider follows the prescribed protocols and executes the required processing steps correctly, but may attempt to collect, inspect, or infer any information accessible to it.
3. Internal attacker. An internal attacker is assumed to have access to systems within the victim organization and may exploit vulnerabilities in its OT or IT infrastructure. Such an attacker may compromise locally deployed components, access sensitive data or credentials, or interfere with data processing and communication within the organization.

4.1.2 Assessing Use Cases and Risks

The following sections provide information about the methodologies used to develop use cases and to analyze risks. The corresponding relevant standards for each topic are also listed for better understanding.

4.1.2.1 Use Case Assessment

When it comes to use case assessment, there is no single dominant procedure or standard which is perfectly tailored to cover all topics of IT security. However, numerous attempts were made to structure the way how use cases can be evaluated to be feasible, secure, or risky, for instance. Usually the procedures of multiple established standards are combined. Some of the most relevant standards are listed below.

- ISO 31000 – It mostly covers general risk management. Here, for a given context risks are identified, analyzed, and evaluated. Afterwards, risks must be treated continuously using e.g., monitoring.
- ISO/IEC 27005 – It focuses on information security risk assessment and is designed to work with ISO/IEC 27001.
- NIST SP 800-30 – It deals with deals with Cybersecurity risk assessment with a widely used practical methodology, which focuses on continuously conducting and maintaining the assessment of system components.

According to these standards, the use cases covered must provide carefully defined information about objectives, stakeholders, assets, assumptions, and requirements. Possible threats or vulnerabilities together with their likelihood, impact and risk level must always be assessed. Finally, the use case assessment must cover analyzed trade-offs in areas such as for instance security and privacy benefits, performance overhead, or deployment complexity. From all created use cases under test typically only the best use cases are selected for a technological realization.

Criteria which can be used to evaluate use cases include:

- Operational value (does the use case solve an important problem?)
- Feasibility (can the use case be implemented?)
- Performance (is the overhead within acceptable boundaries?)
- Usability (is it simple to use or does it cause security issues concerning the human factor?)
- Interoperability (are system and protocol standards used?)

- Scalability (does the technology scale well with larger deployments?)
- Maturity (can the use case be realized with a higher Technical Readiness Level?)
- Cost (are the efforts and costs of the use case acceptable?)
- Regulatory compliance (can the use case be implemented in a compliant way to regulation?)

Once the use cases themselves are defined and assessed, the next part of the analysis must cover the specific technical requirements needed for an implementation and the creation of a suitable system architecture. At this stage, possible risks must be assessed, which is detailed in the next section.

4.1.2.2 Risk and Threat Assessment

This section provides an overview of risk management methodology, best practices, and standards.

The goal of cybersecurity risk management is to identify, prioritize, and treat cybersecurity risks. This section outlines how risk assessment and treatment are addressed in cybersecurity. The following content is derived from IEC 62443-3-2 and BSI-Standard 200-3.

Cybersecurity risks are typically assessed based on the frequency of occurrence of a threat and the impact of potential consequences in case this threat materializes (cf. BSI-Standard 200-3). The frequency of occurrence can also be broken down into the factors of threat and vulnerability, i.e., how likely a threat is to exploit an existing vulnerability.

Risk = frequency-of-occurrence x impact-of-consequences

Risks are used to prioritize efforts, i.e., identifying cybersecurity requirements starts with the highest risks.

For IT security, typical dimensions for *impact-of-consequences* are data breaches, financial loss, and reputation loss. For OT, production loss/damage and safety, i.e., harm to humans and the environment, are additional dimensions to be considered. All dimensions of *impact-of-consequences* are assessed based on the same scale, i.e., a data breach risk assessed as X and a safety risk assessed as X have the same priority.

To conduct a risk assessment, one must define a system model, identify threats to the defined assets and protection goals, and then derive and assess risks based on those threats. Afterwards, the risks need to be treated until they are below a defined risk threshold.

Preparation: System Modeling

Before identifying and assessing threats and risks, the system under consideration (SuC) must be identified and modeled. This includes components and communication relationships, business processes, external requirements (legal framework, customer requirements), assets and their value/criticality, protection goals, and the security measures that are already in place. Common security objectives include confidentiality, integrity, and availability (CIA), often complemented by requirements related to authentication, authorization, and accounting (AAA).

Threat Modeling

The following paragraphs summarize some common methods for threat modeling.

STRIDE

STRIDE is a threat categorization model by Microsoft, where the letters stand for the attacker goals spoofing, tampering, repudiation, information disclosure, denial of service, and elevation of privilege. These attacker goals are the counterparts of the protection goals authentication, integrity, accountability, confidentiality,

availability, and authorization. Before applying STRIDE, the system must be modeled in terms of components and data flows. Then STRIDE can be employed to systematically determine whether a user/service could be spoofed, whether data could be tampered with, etc. Each component and data flow is analyzed against the six STRIDE categories to identify possible threats. Afterwards evil user stories can be used to make threats more concrete and understandable.

Elementary Threats from BSI's IT-Grundschutzkompendium

Besides the STRIDE categories, one could also make use of the 47 elementary threats from BSI's IT-Grundschutzkompendium to determine possible threats for a defined system model. Those simultaneously cover a broader range of threats and are more fine-grained than the STRIDE categories.

Evil User Stories

User stories are a requirements elicitation technique where the perspective of the user interacting with the system is taken, e.g., "As a user I want to use [capability], so that I can achieve [outcome]." Evil user stories are the same approach, but an attacker's perspective is taken to identify what must not happen, e.g., "As an attacker I want to conduct [malicious action] so that I can achieve [harmful outcome]." By this means, vulnerabilities can be identified early in design.

Attack Trees

Attack trees visualize how an attacker could achieve a specific malicious goal. The root of the tree represents the attacker's objective (e.g., "steal blueprint"), and the branches represent different strategies and steps that could lead to that outcome. Each branch breaks the attack into smaller sub-steps using logical relationships such as AND (multiple steps required) or OR (alternative paths). This structure helps to analyze possible attack paths, identify weak points in the system, prioritize defenses based on which attacks are most feasible or impactful, and develop an according defense-in-depth concept.

MITRE ATT&CK

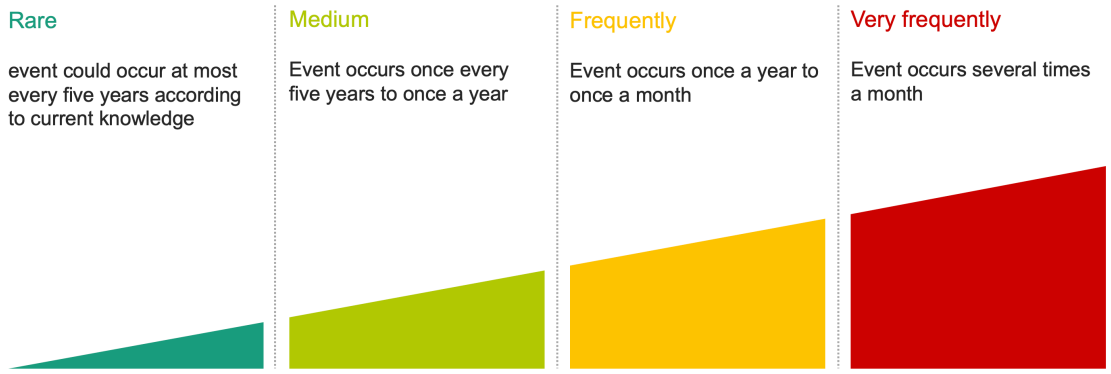
MITRE ATT&CK is a publicly available knowledge base that documents real-world adversary behaviors observed in cyber attacks. It organizes attacker activity into a structured matrix of tactics (the attacker's goals) and techniques (the methods used to achieve those goals). For example, tactics include stages such as Initial Access, Privilege Escalation, or Exfiltration, while techniques describe specific actions attackers use, like phishing, credential dumping, or command-and-control communication. MITRE ATT&CK can hence be used to refine the other, more conceptual approaches by mapping identified threats to real-world attacker techniques. This helps to get a better understanding of the threats and can also be used in the next steps of the risk assessment, including the design of defenses and choosing security controls to mitigate risks.

Risk Assessment

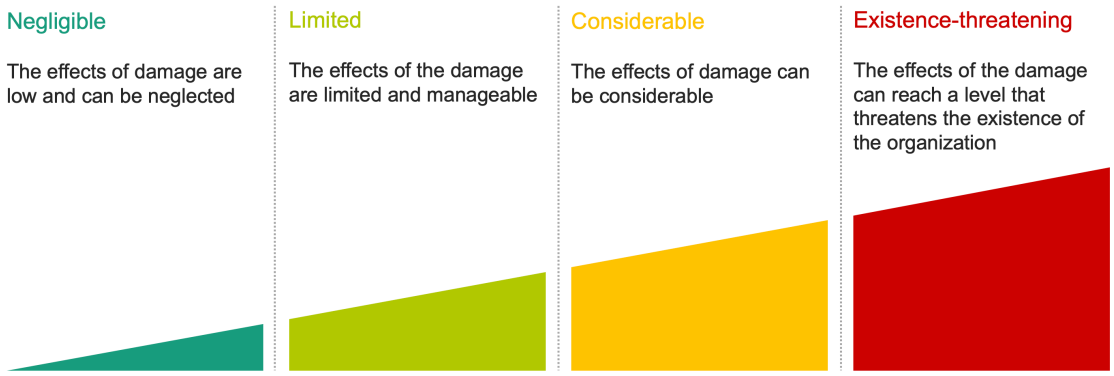
Based on the list of threats generated, risks can be assessed, as risks are the expected losses in case a threat materializes. Often, as a risk baseline the unmitigated cybersecurity risk is determined at first. Unmitigated risk means that a risk is assessed under the assumption that no risk mitigations are in place, i.e., existing security measures are ignored in this step. By this means, a neutral baseline is obtained which in IEC 62443-3-2 is used to derive target security levels of subsystems or components. Via target security levels one can specify how powerful an attacker a security requirement (risk mitigation) must be able to withstand. Unmitigated risks serve as an indicator for this.

The factors of risk, the frequency of occurrence for threats and the impact of consequences are typically rated on semi-quantitative scales. As an example, these are the scales from IT-Grundschutz by BSI.

Exemplary scale for frequency of occurrence:

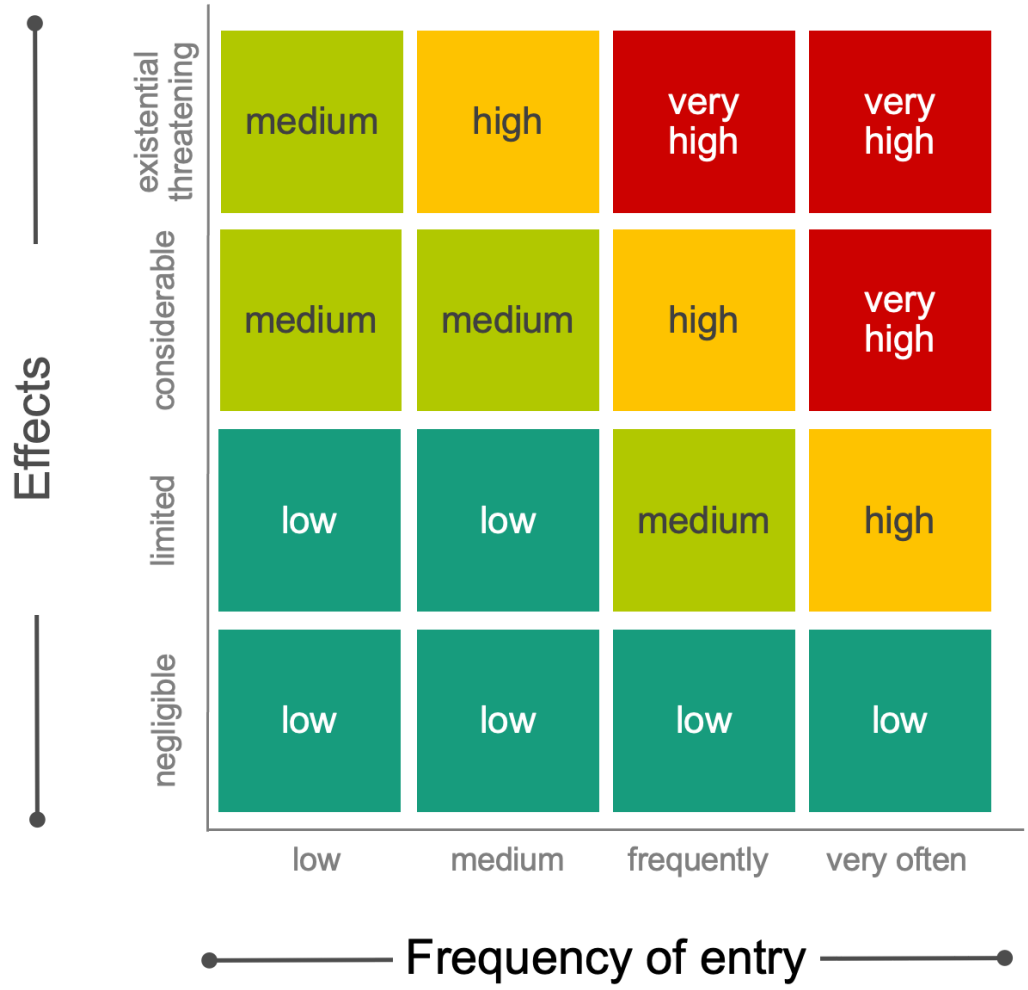


Exemplary scale for impact of consequences:



In this way, threats can be assessed using a risk matrix. As mentioned above, several dimensions of consequences are often considered, such as financial, legal, reputational damage, and harm to people and the environment.

Exemplary risk matrix:



For generic or off-the-shelf components, it is often difficult for manufacturers to assess the impact of consequences of threats. In-line with the CRA it makes sense to consider the intended use of the SuC, purposes and conditions which can reasonably be foreseen, as well as the functional role of the system/component.

Treating Risks

Risks are treated iteratively until the residual cybersecurity risk is pushed below a risk tolerance threshold. Risk treatment strategies are:

- Risk avoidance, e.g., by changing the design to exclude the threat causing the risk
- Risk reduction, e.g., by adding security requirements/controls
- Risk transfer, e.g., by outsourcing tasks or taking out an insurance policy
- Risk acceptance (if the risk is below the risk tolerance threshold)

In the first iteration the effect of security measures that are already in place are considered, and the residual risk is updated accordingly. If the residual risk still exceeds the tolerance threshold, further security requirements or other risk mitigations have to be added until this is no longer the case. As security requirements are identified during this risk analysis process, a cybersecurity requirements specification is created as outcome and input for system design. The documentation of the risk analysis also supports the

assessment of any change to the system which may affect security and must be updated upon such changes.

Existing Norms, Standards and Guidelines for Cyber Risk Assessment

Cyber risk assessment is mandatory for software suppliers as well as for operators of central services and Connectors-as-a-Service. However, the perspective is different, i.e., product security vs. information security, and hence also the things that are assessed. The methodological core is largely identical; scope, protected assets, impact dimensions and the required evidence differ. Organizations that act in both roles simultaneously must therefore maintain two clearly separated assessment scopes within one consistent risk methodology.

Aspect	Software supplier	Operator (central services, Connector-aaS)
Object of assessment	The product and its development lifecycle	The service, its infrastructure and its operational processes
Protected assets	Product integrity, build/release chain, embedded secrets, customer environments affected by the product	Customer data, tenant isolation, availability, credentials/keys, sovereignty enforcement
Primary impact view	Risk **caused by** the product for its users ("outward")	Risk **to** the service and its customers ("inward")
Typical impact dimensions	Confidentiality/integrity of downstream systems, safety (where applicable), mass-effect across install base	Confidentiality, availability, data sovereignty, contractual/regulatory exposure
Main regulatory driver	CRA (Art. 13, Annex I Part I, Annex VII)	NIS-2 (Art. 21), GDPR (Art. 32/35), Data Act (Ch. VI)
Typical evidence	Threat model, SPDL records, technical documentation, SBOM	Risk register, ISMS records, certification/audit reports (TISAX, ISO/IEC 27001, C5)

Common methodological foundation (applicable to both roles)

- **ISO/IEC 27005:2022** — *Guidance on managing information security risks*. The central reference for both roles. It explicitly supports two complementary identification approaches:
  - **asset-based** (assets → threats → vulnerabilities), and
  - **event-based / scenario-based** (risk scenarios derived from threat intelligence and business impact).

  It is aligned with ISO/IEC 27001:2022 Clauses 6.1.2/6.1.3 and 8.2/8.3 and is therefore the method layer beneath an ISMS.
- **Threat modelling methods** — STRIDE, attack trees, PASTA, etc. Threat modelling is the bridge between an architecture and a risk scenario, and is explicitly required by IEC 62443-4-1 (SR-2) and NIST SP 800-218 (PW.1.1).
- **Threat and weakness libraries** — MITRE ATT&CK, ATT&CK. Used to make scenario identification systematic and repeatable rather than workshop-dependent.

Risk analysis for product suppliers / development

1. **ISO/IEC 27005** — as above; supplies the risk process into which product-specific analyses feed.

2. **ISO/IEC 27034 (Application Security, multi-part)** — defines the Organization Normative Framework (ONF), the Application Normative Framework (ANF), Application Security Controls (ASC) and an application security lifecycle model.
3. **NIST SP 800-218 (SSDF)** — practice-level requirements including threat modelling, dependency management and vulnerability response.
4. **SLSA** — build and release integrity levels; addresses supply-chain scenarios that classical risk catalogues do not cover.
5. **IEC 62443-3-2** — *Security risk assessment for system design*: partitioning into zones and conduits, initial and detailed risk assessment. Formally addressed to asset owners and system integrators, but also applied by manufacturers.
6. **IEC 62443-4-1 (SPDL)** — Secure Product Development Lifecycle Requirements. Practice *Specification of Security Requirements* explicitly requires a threat model per product/architecture.

Risk analysis for operators

A. **ISO/IEC 27005** — the core risk method; combined with **ISO/IEC 27001/27002:2022** as the control baseline for treatment.
B. **ISO/IEC 27017** — cloud-specific security controls and implementation guidance for cloud service providers and customers, including the shared-responsibility split.
C. **NIST SP 800-30 Rev. 1** — *Guide for Conducting Risk Assessments*: threat sources, threat events, vulnerabilities, predisposing conditions, likelihood and impact, with usable rating scales and templates.
D. **BSI IT-Grundschutz** — **BSI Standard 200-1** (ISMS), **200-2** (methodology: basic, core and standard protection), **200-3** (risk analysis based on the catalogue of 47 elementary threats), **200-4** (business continuity management), plus the **IT-Grundschutz-Kompendium** as the control catalogue.
E. **IEC 62443-3-2** — risk assessment for system design; relevant where the operated service becomes part of, or a conduit into, an industrial system.
F. **IEC 62443-2-1** — security programme requirements for IACS asset owners.

4.1.3 Planning the Integration of IT/OT

Data space applications and use cases may either involve data originating in the OT infrastructure of a participant's factory or deliver data assets to be used in that OT infrastructure. This raises two questions: how to integrate a data space gateway at an IT/OT interface, and which security requirements OT infrastructures should implement to support such applications and use cases.

Semiconductor manufacturing includes OT systems such as fab equipment, MES and automation systems that must be protected against cyber threats. These systems reside in T3 and T4 and are exclusively the responsibility of the participants; they therefore fall outside the scope of data space-side measures, but the security requirements of the data space must be consistently mapped onto them. IEC 62443 is not a formal obligation within the data space, but a recognized and reasonable basis for doing so, and participants are encouraged to follow it.

4.1.3.1 Segmentation and OT-DMZ

Participants operating OT should segment their internal IT/OT networks in a zones and conduits model according to IEC 62443-3-2 (Edition 1.0, 2020-06). In particular, following ZCR 3.2, there should be no direct communication between IT networks and OT networks. Instead, data exchange between OT and IT should

be implemented via services operated in an OT-DMZ, i.e., a demilitarized zone between OT and IT. A DMZ is a network security concept implemented by one or two firewalls, which allow connections between OT and DMZ as well as between IT and DMZ, but not directly between OT and IT (an according architecture is shown in Figure 11). In this way, for example, process data from the OT can be made available on a data historian server in the OT-DMZ for use in the IT. An example of the opposite direction of communication is the (remote) maintenance of OT assets, which is often implemented via jump servers: with the appropriate permissions, a user from IT connects to a jump host, from which a connection to an OT system is then established in a controlled and monitored manner.

Consequently, if data from the OT is to be transferred via the data space, it should be made available from the OT on a dedicated server in the OT-DMZ. The data space gateway itself is attached to the IT or OT-DMZ side and has no direct attachment to OT systems.

4.1.3.2 Integrity of OT Data

To ensure the quality and reliability of the provided data, an assumption regarding the integrity of OT data and the possibility of integrity checks by the data consumer is required before the data is further processed. At least SL-2 of the IEC 62443-4-2 component requirement CR 3.4 (Software and information integrity, Edition 1.0, 2019-02) should be ensured on the OT side, i.e., the integrity and authenticity of the data is made verifiable — typically by cryptographic means — so that a data consumer can check it before the data is reused.

4.1.3.3 Cybersecurity Program and Risk Assessment

If OT infrastructures of participants' factories act as a data source (and potentially sink) in Semiconductor-X applications, their resilience against cybersecurity threats must also be considered in the overall security concept. To achieve end-to-end security in industrial data spaces, each participant operating OT should establish a suitable Cybersecurity Program (ISMS) for their IACS based on risk assessments, in accordance with IEC 62443-2-1.

Furthermore, IACS-specific security goals, threats, vulnerabilities and consequences should be integrated into the security risk assessments at the level of the data space, in order to understand which IACS security controls and countermeasures are required to meet the data space security requirements.

Table 4: Mapping of Data space Trust Levels to IEC 62443 Security Levels

Trust Level	IEC 62443 SL-T (OT side)	Protection Objective
Basic	SL-1	Basic protection against casual or coincidental violation
Trust	At least SL-2	Protection against intentional violation by attackers with low to moderate resources and skills
Critical	SL-3/4, with remote attestation	Protection against sophisticated, targeted attacks

In general, the underlying OT infrastructure of participating data providers and consumers should reach at least SL-2 for all data space applications that require the trust level.

The selection of target security levels (SL-T) per foundational requirement follows from the use case:

- If strictly confidential data that may be used at OT level is shared with a participant, the receiving Data Consumer should choose the SL-T for FR2 (Use Control) and FR4 (Data Confidentiality) accordingly.
- If the integrity of input data sourced from a remote OT network is critical for a data space application, the Data Provider should choose an appropriately high SL-T for FR3 (System Integrity) and implement the resulting IACS security controls and countermeasures.

4.1.4 Defining Assets and Protection Goals

In the following, we describe how we define the assets a gateway is to protect, what type of protection we envision, and how the lifecycle of assets must be addressed.

4.1.4.1 Assets

In the context of the Semiconductor-X data space, an asset is anything that has value for a participant or for the data space as a whole. Value implies that the asset is not freely or indefinitely available and therefore requires protection. Assets include both physical and digital elements and can be grouped into the following categories:

- **Data assets**, e.g., internal client data (e.g. MES/ERP data, production data), data and information exchanged via the data space (e.g. DPPs, machine telemetry, quality data), and contractual information between data space clients. Data assets are grouped into three different data criticality levels:
    - Semi-public data (e.g., digital nameplates, product related environmental data, digital product pass)
    - Confidential data (e.g., machine utilization, product volumes, quality data, energy consumption)
    - Strictly confidential data (e.g., machine configurations, algorithms, blueprints, recipes that contain IP and whose disclosure impacts competition)
- **Credentials and key material**: e.g., login credentials for the data space (passwords, API secrets), private cryptographic keys, shared secrets, access tokens and API keys
- **Identities and public identity information**: e.g., user, organizational, service, device and gateway identities, identifiers and attributes (e.g. DIDs, VCs, account IDs), public keys and public certificates, signed identity and trust assertions.
- **Infrastructure and physical components**: e.g., gateways into the Semiconductor-X data space, edge devices, network components, servers, storage, cloud services hosting Semiconductor-X components, and production/manufacturing machinery in participant factories.
- **Federation services in the data space**: Identity Provider, Decentral Discovery Service/Service Catalog or Monitoring Service.
- **Supporting assets**: e.g., security policies, process descriptions and operating procedures, configuration templates, scripts and deployment artifacts, as well as documentation and training material.

Not all digital assets are necessarily exchanged via the data space. For such assets, the framework does not aim to provide additional protection, but it must ensure that they are not exposed to additional risks.

4.1.4.2 Protection Goals

The Semiconductor-X data space uses the following protection goals, oriented towards common security standards (e.g. ISO/IEC 27001, IEC 62443, NIST Cybersecurity Framework):

- **Confidentiality (C)**: Prevent unauthorized parties from gaining access to data.
- **Integrity (I)**: Prevent data and system state from being altered in an unauthorized or undetected way.
- **Availability (AV)**: Ensure that systems and data are usable by authorized parties when needed.
- **Authenticity (AU)**: Ensure that the identities of users, devices and services are genuine and verifiable.
- **Accountability & Traceability (AT)**: Ensure that relevant actions are logged and can be linked to responsible entities for audit and incident analysis.
- **Non-repudiation (NR)**: Ensure that parties cannot credibly deny actions they have performed (e.g. providing, accessing or modifying data).

4.1.4.3 Asset Classification by Protection Needs

Assets are grouped into categories and assigned baseline protection goals.

Asset category / class	Typical examples	C	I	AV	AU	AT	NR
Semi-public data assets	Digital nameplates, product environmental data, selected DPP attributes	O	✓	O	✓	✓	O
Confidential data assets	Machine utilization, product volumes, energy consumption, operational KPIs	✓	✓	O	✓	✓	✓
Strictly confidential data assets	Machine configurations, algorithms, IP-bearing datasets, blueprints	✓	✓	O	✓	✓	✓
Credentials and key material	Passwords, API secrets, private keys, shared secrets, tokens, key stores	✓	✓	✓	✓	✓	O
Identities and public identity information	User/service/device IDs, organizational identities, public keys, certificates	O	✓	✓	✓	✓	O
Infrastructure and physical components	Data space gateways, edge devices, network components, servers, storage, clouds	O	✓	✓	✓	✓	N
Federation services in the data space	Identity & access management, PKI, catalogs, logging, monitoring, incident management	✓	✓	✓	✓	✓	O
Supporting assets	Security policies, procedures, config templates, scripts, documentation, training	O	✓	O	O	O	N

✓ = Required, O = Optional (i.e., depending on use case), N = Not required

Protection requirements for each concrete asset are derived from this baseline and documented in the asset inventory (e.g., by assigning specific confidentiality / integrity / availability levels or IEC 62443 security levels, as well as any specific requirements regarding authenticity, accountability & traceability, and non-repudiation). Deviations from the baseline are allowed if justified by a risk assessment.

4.1.4.4 Asset Life Cycle and Life Cycle Management

All assets in the Semiconductor-X data space (data, gateways, central services, keys, secrets, etc.) must be protected throughout their entire life-cycle. Security, quality and compliance measures must be applied consistently across all phases, including updates, recalls and final decommissioning.

Phase	Purpose / typical operations	Examples	Key security aspects
Planning & design	Identify asset, owner, purpose,	New data product, gateway, central service, key hierarchy	Asset classification, mapping to protection goals, risk

	classification, protection needs		assessment, update/recall concept
Acquisition / creation	Procure or generate the asset, initial configuration	Buy hardware, provision VM, generate keys, create dataset	Trusted sources, secure generation (keys, configs), least privilege, secure defaults
Commissioning & registration	Put into service and register in data space components	Onboard gateway, publish service/data, register identity	Hardening, security checks, identity proofing, registration in catalogs/IAM/PKI, audit logging
Normal operation	Regular use, monitoring, periodic review	Data exchange, analytics, use of services and gateways	Access control, encryption, monitoring, incident logging, review of classification and contracts
Update & change management	Patches, functional changes, key rotation, data corrections and new versions	Firmware/software updates, schema changes, key rotation	Tested and approved changes, rollback options, consumer notification, no unapproved security downgrades
Recall / temporary withdrawal	Temporarily disable or withdraw assets in case of security, quality or compliance issues	Disable service, withdraw data product, revoke certificate	Fast isolation/removal, revocation of credentials, clear communication to consumers, detailed logging
Decommissioning & disposal	Permanently retire assets and remove from the data space	Retire gateway, delete data, destroy hardware, remove identity	Secure wipe/cryptographic erase, revocation and deregistration, retention-aware deletion, documentation

Each asset must have an assigned owner who is responsible for applying this life-cycle, keeping the asset inventory up to date, and ensuring that onboarding, update, recall and decommissioning follow defined, auditable processes. This ensures that assets are securely introduced, operated, adapted and removed without exposing the Semiconductor-X data space or its participants to unnecessary risk.

4.1.5 Selecting Security Requirements and Controls

4.1.5.1 Gateway Security Requirements

Following the core principle of scalability of security mechanisms, our security architecture must provide a selection of appropriate security measures for given use cases. To reduce complexity and simplify onboarding, we base our security architecture on a three-tiered data criticality classification. Each data criticality is associated with a security level, which defines the security measures necessary to protect data assets of this criticality. The security levels are aligned with the gateway assurance levels to be introduced in Section 4.1.7, thereby linking the certification process of data space participants with concrete technical security requirements.

Data Criticality	Use Case Examples	Security Level	Security Characteristics	Assurance Level
Semi-public	Eco Pass Kit, Digital product passports (metadata), Public IDLink, Product specifications, Product Carbon Footprint Kit	Basic	Lightweight protection, high interoperability	Base
Confidential	Circularity Kit, Demand and Capacity Management Kit, Logistics Kit, Purchase Order Kit, Requirements Kit, Commercial Documents Kit, Material Accounting Kit	Trust	Policy agreements, data encryption, system integrity verification, auditability	Trust
Strictly confidential	IP (blueprints, recipes, algorithms), Data Driven Quality Kit, Online Simulation Kit, Model Based Production Kit, ...	Critical	Hardened nodes, technical policy enforcement, secure tunnels, hardware-backed attestation, continuous monitoring	Trust+

Based on the security by design principle, we propose a three-layer security architecture for Manufacturing-X that uses the security levels corresponding to the data criticality levels defined in the previous section. This layered approach balances security effort with functionality, performance, cost, and ease of use, depending on data sensitivity and use-case criticality.

It defines three distinct yet fully interoperable data space layers, corresponding to the three security levels: Basic, Trust, and Critical. Each layer is equipped with security controls, technical requirements, deployment models, and running costs, which are tailored according to the security level that this layer implements:

- **The Basic data space layer** provides lightweight, cost-efficient protection with minimal overhead, dramatically lowering the entry barrier for SMEs and low-risk use cases.
- **The Trust data space layer** adds enhanced policy enforcement, auditability, and commercial-grade safeguards for confidential business data.
- **The Critical data space layer** delivers the highest assurance for intellectual property, OT telemetry, and mission-critical applications.

All layers share a common data fabric but apply differentiated security stacks. Cross-layer communication is governed by strict, predefined rules and secure channels. Participants can choose to implement the security layer(s) that best match their needs and capabilities. The data space gateways and network infrastructure must be designed and implemented to meet the layer's security requirements.

We formalize the requirements for data protection below. Notably, these are independent of the data criticality levels for now. In the next section, we categorize which of these requirements must be implemented to what degree for each level. These must then be fulfilled by all data processing platforms handling data of a given criticality level.

Authentication & authorization - Assets must not be shared with unauthorized gateways or other parties. Entities must be authenticated and authorized prior to receiving assets.

Usage Control enforcement - The gateway must enforce the usage control policies provided alongside the received data and refuse data processing if usage control enforcement is not possible, e.g., due to unknown or unsupported policies.

Data protection in transit - Assets must remain confidential, integrity-protected, replay-protected, and authentic during transfer from data provider to data consumer.

Data protection at rest - Assets must remain confidential, integrity-protected, replay-protected, and authentic while being stored.

Data protection in use - Assets must remain confidential, integrity-protected, replay-protected, and authentic while being processed.

Platform integrity protection - The platform must be designed such that common security standards are observed, and the configuration and execution of the gateway must be protected using technical means against any undetected manipulation by a third party.

Proof of integrity and capabilities - The recipient of assets must provide verifiable proof of its platform integrity and that it can fulfill the previously defined requirements.

Key management - Cryptographic keys must be adequately protected, e.g., by storing them in Trusted Platform Modules (TPMs), Hardware Security Modules (HSMs) or Secure Elements.

Availability - The provider of a gateway must provide mechanisms to prevent availability issues.

4.1.5.2 Gateway Security Controls

To ensure adequate protection of exchanged data, each criticality level and data space layer implies the necessity of certain standard **security controls**. Security controls are technical, organizational, or physical measures that reduce cybersecurity risk to IT/OT systems and data and implement the previously defined security requirements to different degrees. As such, they lie at the heart of every cybersecurity concept.

To achieve a secure data space infrastructure, appropriate security controls must be identified and implemented according to the relevant IT/OT security standards and guidelines. Typically, a use-case-based threat modelling and risk assessment process is used to identify the security requirements of systems/components and to select appropriate security controls depending on the individual protection needs. The effectiveness of the implemented security controls must be continuously (re-) evaluated as part of the established cybersecurity program (ISMS/CSMS). This common methodology is closely linked to the relative industry sectors and their respective Regulations and Guidelines (e.g., ISO 81001-5 for medical devices).

Gateway Security Levels

Crucially, our security architecture must identify which technical security controls are necessary at the data space gateways to adequately protect data assets of different criticalities. To integrate with a risk-based cybersecurity approach and preserve flexibility, we refrain from obligating concrete technical solutions and configurations. Instead, we define which security measures gateway implementers must consider as countermeasures during the detailed risk assessment. The resulting control catalogue offers a time-saving approach to achieve a common understanding of gateway security over different use cases. However, it is not intended to fully replace traditional use-case-based risk assessment. Additional requirements arising from the use cases, which have not been defined in the per-level common risk assessment, still necessitate individual threat modelling.

		Gateway Security Levels		
Category	**Common Security Controls**	**Basic**	**Trust**	**Critical**

Authentication & Authorization	AAA services, OAuth	✓ Company level	✓ Department level	✓ Individual level
Usage Control Enforcement	ODRL, policy-aware applications	○	✓ Connector level	✓ Application level
Data protection in transit	Secure communication protocols (TLS, IPSec, ...), secure tunnels (VPN, ...)	✓	✓	✓ Use hybrid PQC
Data protection at rest	Authenticated Disk Encryption	○	✓	✓
Data protection in use	Confidential VMs, Transparent Memory Encryption	○	(✓)	✓
Platform integrity protection	Secure boot, Measured boot	(✓)	✓	✓
Proof of integrity and capabilities	Remote Attestation	○	✓	✓
Key management	Secure elements, HSMs, TPMs	○	(✓)	✓
Availability	DoS protection, rate limiting, load balancing	(✓)	✓	✓

○ Usually not required. (✓) Should be considered. ✓ Must be considered.

A gateway at **Security Level Basic** fulfills the minimal requirements for a secure data space infrastructure, including participant authentication and secure data exchange. A gateway satisfying **Security Level Trust** offers additional security controls to protect data at rest and to restrict data processing. The latter is achieved through usage control policies that define under which circumstances data may be transferred and processed on the data provider side. The level also introduces a secure platform that ensures the gateways' integrity. To adequately protect even strictly confidential data, gateways of **Security Level Critical** include a secure platform that ensures and attests the gateways' integrity and capability to enforce adherence to the defined usage control policies in the connector and all utilized data applications. This includes authenticating and verifying the hardware-backed platform identities and software states of gateways.

Note that security controls applied to the gateway alone are not sufficient to cover all data criticalities. For semi-public data, access control enforced by the gateway may be adequate. However, for strictly confidential data, it must be guaranteed that only the intended receiver can decrypt the data. This requires end-to-end encryption performed directly at the point of origin (where the data is generated), not solely at the gateway level.

Control Catalogues

Typical Risk Management is difficult to apply in data spaces where new use cases emerge continuously and form a kind of moving target. "G*lobal ecosystems require comprehensive security architectures for all participants, regardless of regional and national structures. This necessitates* ***interoperable cybersecurity guidelines*** *that are recognized by all parties involved, as well as a shared regulatory framework*" (VDE-AR-E 2802-20), *and can be commonly used by the different use cases*. Therefore, to assist data space participants and developers in identifying necessary security controls for their scenarios during a cybersecurity risk assessment, we propose Manufacturing-X to adopt a procedural model in accordance with VDE-AR-E 2802-20 (Horizontal cybersecurity concept for Industry 4.0), providing **catalogue-based cybersecurity concepts and controls as baseline requirements**. The catalogues of concepts and controls are tiered into three Gateway Security Levels, each corresponding to one of the presented data criticality levels.

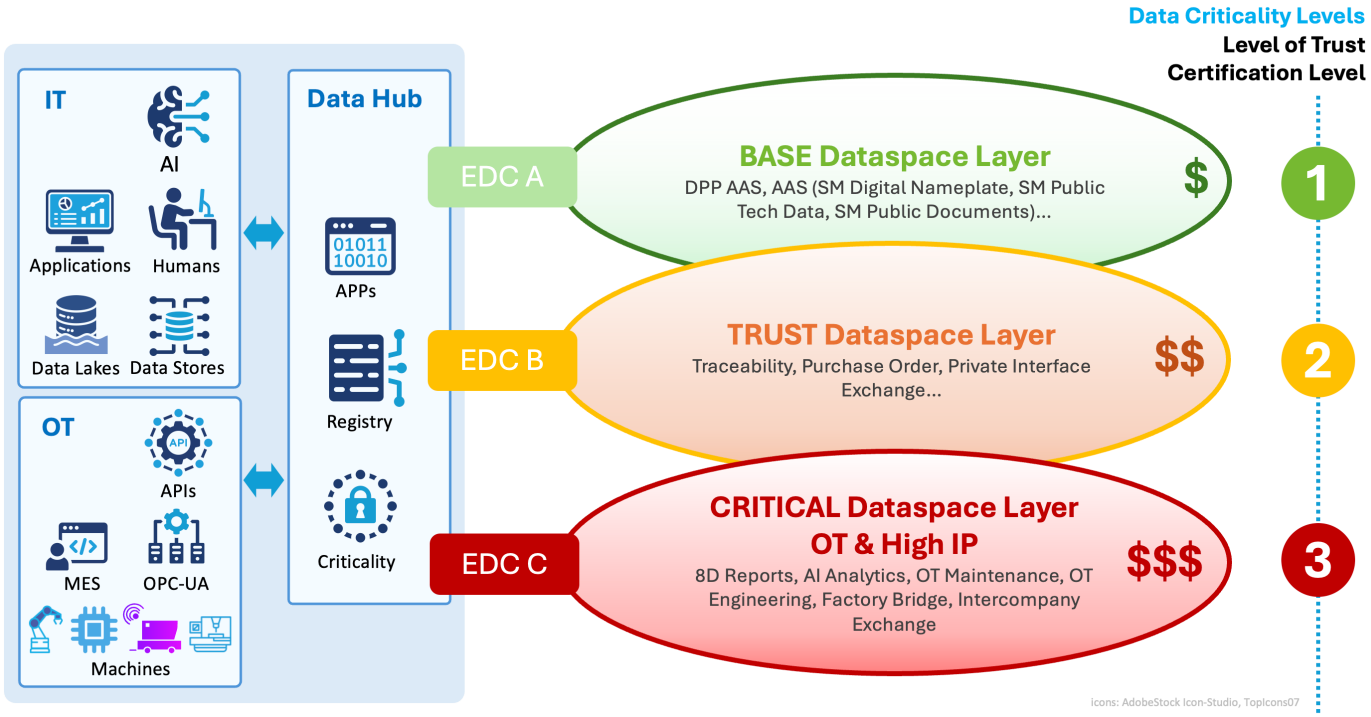


Figure 5: A three-layer staggered, use-case-oriented security architecture balances security effort with functionality, performance, cost, and ease of use, depending on the sensitivity of the data and the criticality of the use case.

This graduated model finally delivers the long-promised “one fits all” solution: every participant connects via a single controlled connector (EDC described in the framework of the MX-Port Herkules and Orion) of a specific security level. The security strength, controls, implementation effort, and operating costs are scaled precisely to the levels and use cases the participant has chosen to support.

By enabling participants to select the right data space layer for their specific business needs, the architecture achieves an optimal balance between strong security, usability, scalability, performance, and cost efficiency. It enables broad adoption across the entire manufacturing ecosystem while ensuring that higher-criticality data and use cases receive the robust protection they require. Participants can start at the Basic layer and upgrade to the Trust or Critical layer at a later stage, if required by the implemented Eclipse Tractus-X Use Cases (KITs).

Data Security

For strictly confidential data, protection must extend across all three states: at rest, in transit, and during processing. Encryption in transit (e.g., TLS) secures data as it moves across the network between provider and consumer. Confidential computing then ensures that data remains encrypted and isolated throughout processing itself, until it reaches the Trusted Execution Environment on the consumer side, so that neither the infrastructure operator nor any other party outside the TEE can access the plaintext. Transparent Memory Encryption (TME) complements this by encrypting the physical memory contents at the hardware level, protecting against physical attacks on the memory itself (e.g., cold-boot attacks or bus sniffing). Unlike confidential computing, it does not by itself protect against access by a privileged hypervisor or operating system. The data provider can require remote attestation of the consumer TEE before releasing keys, making the enforcement of this protection verifiable rather than declared. For IP related use cases such as recipes, machine configurations, and blueprints, hybrid post quantum cryptography is applied as soon as technically feasible, in alignment with the encryption requirements of Section 4.1.5.1.

Architectural Boundaries and Compensating Controls

The Tractus-X EDC is the purpose-built core component for sovereign data exchange, but it was not designed to cover all security aspects of a data space gateway on its own. This is a matter of architectural boundaries rather than vulnerabilities: the following boundaries define where the intended scope of the EDC ends and where adjacent layers, tools, or organizational processes must step in to achieve end-to-end security.

Table 5: Architectural Boundaries and Compensating Controls for Tractus-X EDC

Architectural Boundary	What the EDC does not cover by design	Required Compensating Measures
Policy Enforcement Depth	Negotiates policies but does not technically enforce all usage and access policies end-to-end down to the data source	Strict policy validation, monitoring of policy evaluations, enforcement at application and data source level behind the EDC
Fine-Grained Backend Authorization	Not designed for object-level authorization for every backend API endpoint	Additional IDP/OAuth2 layer behind the EDC for granular backend access control
Network Segmentation to OT	No network isolation between data space and internal OT environments	IEC 62443 Zone-and-Conduit model and strict network separation at the infrastructure level
Supply Chain Integrity	Relies on dependencies without guaranteeing integrity of all upstream components	SBOM management, dependency scanning, coordinated vulnerability disclosure in CI/CD and operations
Data Transfer Volume Control	Facilitates push/pull transfers without monitoring for data exfiltration patterns	Rate limiting, WAFI, data loss prevention, transfer volume monitoring at the Data Plane
Availability Protection	No native protection against flooding of catalogs, contract negotiations, or management APIs	Rate limiting, load balancing, DoS protection via surrounding infrastructure
Credential Lifecycle Management	Relies on external VC revocation mechanisms without monitoring revocation status	Verifiable Credential Status List monitoring as a Federation Service with automatic revocation propagation
Internal Endpoint Protection	Data Plane may expose backend endpoints designed for internal access only	WAFI filtering, network policies, and access control at the proxy level
Multi-Tenancy Isolation	Can be operated with multiple participants sharing a single instance separated only logically	Deployment-isolated connectors per participant, or external isolation mechanisms and monitoring of tenant boundaries

The key takeaway confirms the gateway definition of this chapter: the gateway is the composed entirety of data space components. A secure deployment requires wrapping the EDC with complementary security controls, including, e.g., WAFI, network segmentation, and monitoring, rather than expecting the EDC to fulfill all security functions independently. Operations-side controls for this composition are specified in chapters 4.2 and 4.3 and the concrete hardening stack in Appendix C: Kubernetes Security Architecture Model.

4.1.6 Securely Developing Gateway Components

A Secure Development Lifecycle (SDL) provides a supporting framework to systematically develop components with suitable security features (cf. ISA/IEC 62443-4-1, ISO/IEC 27034). To achieve this, security-related activities are integrated into all lifecycle phases, i.e., planning, requirements engineering, design, implementation, testing, deployment, operation, continuous development, and maintenance. A core activity is Threat and Risk Management, which includes identification, assessment, prioritization, and treatment of threats according to the associated risk. The output of a risk analysis is a list of security requirements that must be implemented to achieve a level of security for the developed component that is appropriate given the identified risks.

International standards, as mentioned above, target software product development in a company context. The key difference with respect to open-source software development (OSS) is its distributed nature. Hence, "trusting the process" is even more important than "trusting the person" when managing security in OSS projects hosted in online repositories. Security-related practices, i.e., security management, specification of security requirements, secure design, secure implementation, security testing, management of security-related issues, and update management, need to be adjusted for community governance, transparency, and distributed contributors.

Regarding **security management**, security policies and responsibilities should be maintained in the repository and include how vulnerabilities are reported to the project (coordinated vulnerability disclosure (CVD)) and addressed, within what time frame, and who is responsible for security-sensitive changes. Responsible maintainers should be designated for vulnerability management and for overseeing releases, including the signing of release artifacts to ensure authenticity. Security policies should also define requirements for commit signing, acceptable identity verification methods for contributors, and the trust model for the CI/CD pipeline. Security policies should further cover supported versions, and it should be clearly communicated that security properties claimed by the project do only apply for these versions.

Security requirements should be maintained in a security design documentation in the repository. Processes should be established to ensure they are reviewed and updated as necessary. Security requirements specification assumes an intended use of the component/system, defined system boundaries, operational assumptions, and a security context, i.e., which security capabilities are expected from the environment where the component/system is operated. The security requirements specification should further include the threat model (e.g., using STRIDE), the assessment of the threats (risk assessment) as well as the requirements that have been identified to mitigate the particular threats (e.g., authentication requirements, cryptographic requirements, input validation expectations). Refer to existing standards like ISO/IEC 27005 (Guidance on managing information security risks) or IEC 62443-3-2 (Security risk assessment for system design) for guidance how to conduct a threat and risk analysis.

In the **secure design** phase of SDL, the goal is to ensure the development of a secure system architecture following a defense-in-depth concept, i.e., have multiple independent security measures to mitigate threats, and follow secure design principles, e.g., least privilege, separation of duties, and secure failure. Regarding software, risks are typically associated with interfaces. It should be specified which interfaces exist, how they are supposed to be used, whether access to the interface crosses a trust boundary, how users are authenticated, which roles/privileges are required for their use, which measures are in place to protect them, and how untrusted input is handled. A secure design review process needs to ensure that chosen mitigations handle risks appropriately. This process should also be triggered upon feature requests or security-related issues.

Secure implementation includes secure coding standards and security implementation reviews. Secure coding standards define how coding rules like input validation, memory safety, secret handling, etc. should be implemented. Secure coding standards should also be supported by tools in the build pipeline like static analysis and dependency scanning. The OSS community needs to define how security-critical pull requests are reviewed.

In OSS development, where contributions come from distributed individuals, **securing the build and delivery process** is essential to ensure that the artifacts users consume actually correspond to the reviewed source code. Build provenance refers to verifiable metadata that documents how, where, and from which source a software artifact was produced. An OSS project should generate and publish provenance information (e.g., conforming to the SLSA framework) that allows consumers to verify the origin and build process of an artifact. Reproducible builds complement provenance by ensuring that the same source code, build environment, and build instructions always produce bit-for-bit identical output, enabling independent parties to detect any tampering in the build infrastructure. The project should strive for hermetic builds that do not depend on unverified external resources at build time. Beyond signing release artifacts, individual commits should be cryptographically signed (e.g., using GPG or SSH signatures). This ensures traceability of every code change to a verified contributor identity and protects against spoofed contributions. An OSS project should define which branches require signed commits and how contributor identities are verified. Furthermore, the CI/CD pipeline represents a critical attack surface: a compromised build step can inject malicious code into otherwise trustworthy source. An OSS project should establish an auditable trust chain across all pipeline stages — from source commit through build, test, packaging, to publication. Key principles include:

- Pipeline-as-code: pipeline definitions are version-controlled and reviewed like application code
- Immutable build environments: build infrastructure is versioned, pinned, and tamper-evident
- Least-privilege access: CI/CD service accounts have minimal permissions scoped to their task
- Secret management: signing keys and credentials are stored in hardened vaults, never in logs or source
- Stage-gate verification: transitions between pipeline stages are gated by integrity checks of the preceding stage

Projects should document their CI/CD trust model, including which systems have access to signing keys and how pipeline configurations are protected against unauthorized changes.

Regarding **security testing**, the OSS release process should include unit tests and integration tests, but also explicit security testing using methods like static analysis, fuzzing, vulnerability scanning, or penetration testing. Major changes to architecture or assumptions should also trigger threat mitigation testing, i.e., it should be reviewed whether the security requirements implemented to mitigate the security risks are still sufficient.

Patch management is about preparing, testing, and publishing patches for security-related issues. Patches should be versioned, and patch releases should be signed so that users can verify the authenticity and integrity of a patch. Security fixes should be reflected in change logs, and it may also be necessary to backport patches to older releases that are still supported. The OSS community should also define how patches are to be documented.

The project documentation should include **security guidelines** explaining its defense-in-depth concept as well as security measures expected in the environment. It should cover how to harden, deploy, and operate

the system/component. This includes how security properties should be configured (e.g., authentication, certificates, privileges) to minimize the attack surface and how security patches are to be installed.

Note that these best practices should not be seen as an “as-is checklist”. Each project needs to apply and adapt them based on the intended use case, risk levels, and their individual circumstances. For further information on secure software development (in the OSS context), the following sources are recommended:

- Standards
    - ISO/IEC 27034: Application Security, which includes the Application Security Life Cycle Reference Model (ASLCRM)
    - IEC 62443-4-1: Defines secure product development lifecycle requirements for industrial automation and control systems
- Frameworks
    - NIST Secure Software Development Framework (SSDF): Provides a set of fundamental, sound practices for secure software development
    - SAFECode Fundamental Practices: Practical guidance for secure software development processes
- Maturity Models
    - OWASP SAMM: Software Assurance Maturity Model – measures and improves secure development practices
- Supply Chain Security
    - SLSA (Supply-chain Levels for Software Artifacts): Framework for ensuring the integrity of software artifacts throughout the supply chain
    - Sigstore: Provides tooling for signing, verifying, and protecting software artifacts
    - OpenSSF Scorecard: Automated tool to assess the security posture of open source projects

4.1.7 Gateway Certification Concept

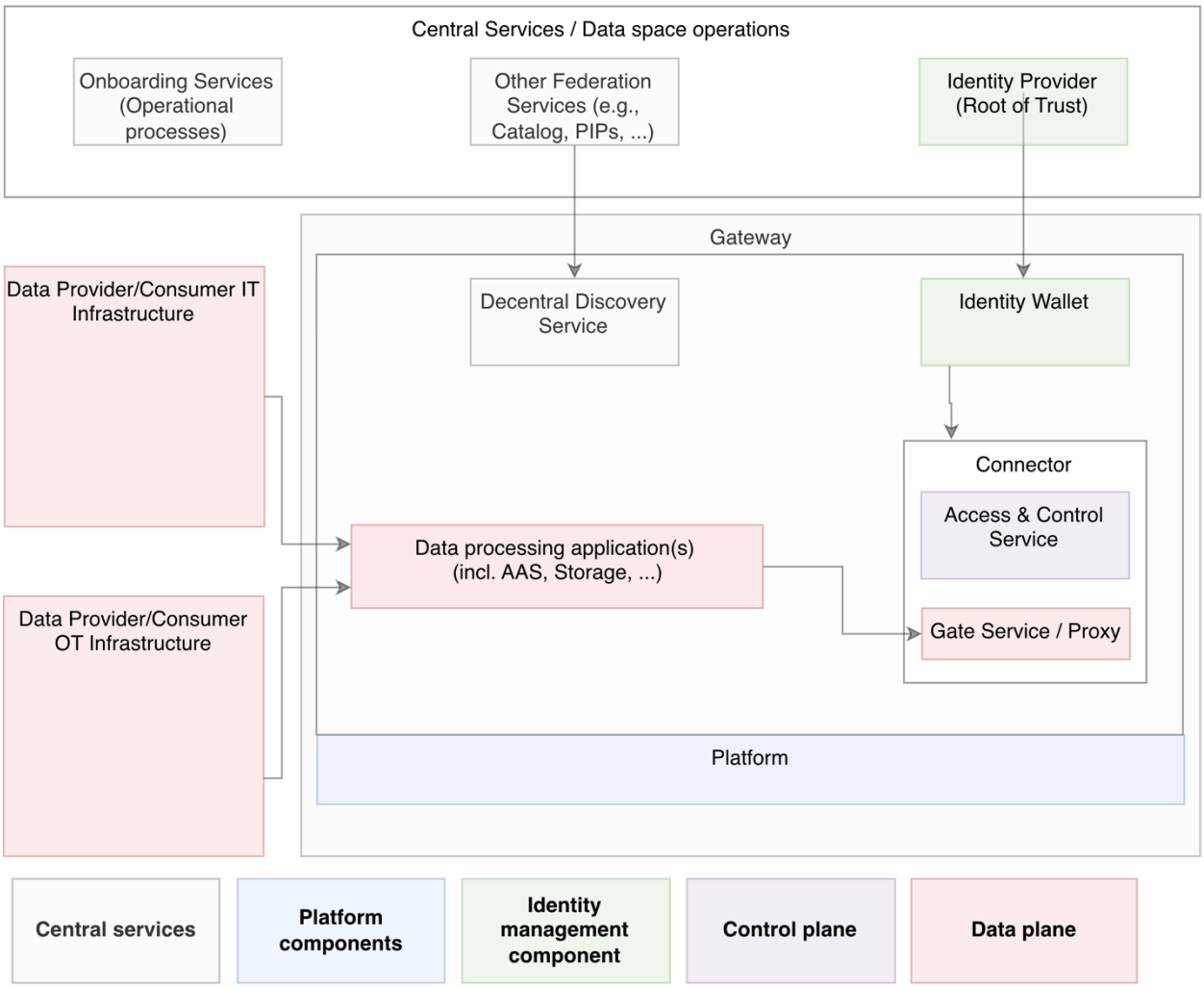


Figure 6 - Ecosystem overview showing the components and process areas that are subject to certification

The Semiconductor-X data space establishes trust in the software and infrastructure that participants use for data exchange. The central artifact is the data exchange gateway. This is a deployment unit hosting an EDC together with supporting services, which participants assemble from container repositories and other artifact sources.

Certification covers a defined boundary, combines component properties with operational properties, where possible verifies claims technically rather than on paper, and is supposed to stay affordable for a value chain spanning integrated device manufacturers and small equipment suppliers.

4.1.7.1 The Certification Object

Boundary Definition

A participant exchanges data with a running deployment, not with a connector binary. That deployment includes a specific image set of connector components and data services, a specific configuration and a specific operating entity that controls all of it. Certifying only the software leaves important risks outside the boundary.

The scheme defines the certification object as the gateway deployment unit: one or multiple EDC control planes, one or more data planes, the declared extensions and adapters, data services, the runtime hosting them including hardware trust anchors, the identity material in use and the operating organization.

Layer Model

Each layer carries its own evidence. Participants either satisfy a layer directly or inherit it from a party holding a valid certificate for that layer.

Layer	Object	Evidence stream	Inheritable from
L0	Legal entity, mandates, personnel authorization	Operational	—
L1	Bare metal, worker nodes, operator platform	Operational	Infrastructure provider, operating company
L2	Container runtime, orchestrator, secret store, admission control	Operational and component	Hardened platform baseline
L3	EDC control plane, data plane, identity handling, policy engine	Component	Certified distribution
L4	Application containers: e.g., wafer-map transformers, anonymizers, IP redactors	Component	Partly, per app
L5	Federated services: identity provider, catalogue, artifact registry, semantic hub	Component and operational	Operating company
L6	Asset classification, ODRL policy library, export-control screening	Participant governance	—

The assurance level of a specific data exchange equals the minimum across four values: the provider's component level, the provider's operational level, the consumer's component level, and the consumer's operational level.

4.1.7.2 Assurance Levels

Each level ties to the asset classes it may carry, because a sustainability figure and a novel integrated circuit layout differ vastly in sensitivity.

Level	Component evidence	Operational evidence	Permitted asset classes
Base	Self-assessment, automated protocol conformance, SBOM, baseline static analysis	Self-declaration against the control set	Catalogue metadata, sustainability data, product carbon footprint, non-sensitive logistics
Trust	Third-party evaluation, penetration test, code-property attestation, development process aligned to IEC 62443-4-1	Audited information security management system, OT zone evidence, mapped cloud attestations	Equipment telemetry, traceability, materials data, lot genealogy, aggregated yield
Trust+	Trust evidence plus hardware trust anchor, measured boot, isolation proof, sound analysis for declared critical properties	Extended audit covering physical security, key ceremonies, continuous monitoring, tested incident response	Process recipes, PDK deltas, GDSII/OASIS mask data, wafer maps, failure analysis reports

Base certificates based on a self-declaration carry a defined sunset date per asset class. Production data therefore migrates toward Trust instead of resting on indefinite self-declaration, which keeps the entry path low without permanently lowering the floor.

4.1.7.3 Roles and Separation of Duties

The scheme separates the parties that create evidence, judge evidence, and consume evidence. Without this separation, a platform operator effectively certifies itself.

Role	Function	Prohibited combination
Data Space Authority	Owns the criteria catalogue, issues and withdraws certificates, maintains the trust list	Must not host gateways
Conformity Assessment Body	Executes evaluations, adjudicates exceptions, signs evaluation reports	Must not supply the evaluated software
Reference Implementation Provider (e.g. Tractus-X)	Ships the distribution, charts, SBOMs, signatures, hardening guide	Must not self-issue Trust-level property credentials
Data Space Operator	Runs federated services, performs enrolment, optionally hosts gateways	Must not issue its own operational certificate
Property attestation service	Runs the code-property query catalogue, issues component property credentials	Must not host the gateways it attests
Federation Service Provider	Issues participant, gateway, and component credentials bound to certificate status	Must not adjudicate certification decisions
Data Space Components Operator	Matches vulnerability feeds against registered SBOMs, detects configuration drift, triggers suspension	—

Every certification result takes the form of a verifiable credential. The connector checks this credential at the protocol handshake, during contract negotiation, and again at transfer start. A withdrawal technically removes a gateway from the data space, not merely administratively.

4.1.7.4 Certification Lifecycle

Process Flow

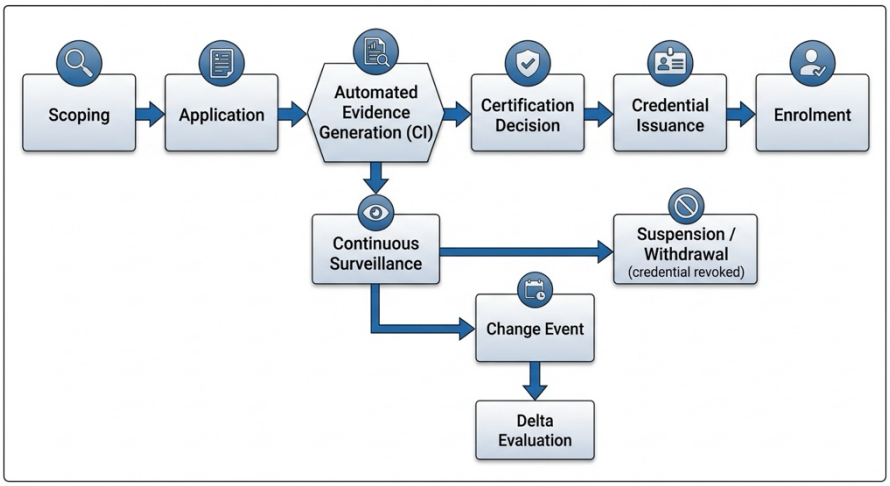


Figure 7 - Certification process model

The pipeline generates evidence, and the evaluation facility judges only what the pipeline cannot decide alone. Manual effort concentrates on architecture rationale, key custody design, custom extensions, and unresolved static analysis findings.

Scoping

Scoping fixes three items before evaluation starts: the deployment variant, the layers the applicant certifies directly, and the layers the applicant inherits. The applicant declares the asset classes it intends to exchange, which sets the target level.

The scoping record lists every image digest inside the boundary. That list anchors delta certification later, because certificates bind to digests rather than to version strings.

Change Management

Semiconductor environments might patch on very different cadences. A design house redeploys weekly; a fab under a change freeze may hold a configuration for months. Change classification prevents either case from forcing full re-certification.

Change class	Trigger	Validation action
Editorial	Documentation update	Notification only
Minor	EDC patch, image rebuild, dependency bump	Automated re-test, automated scope update
Major	New extension, new data plane technology, new region	Delta evaluation by the evaluation facility
Fundamental	Change of deployment variant, operator, or key custody	Full re-certification

Suspension and Withdrawal

Surveillance triggers suspension when a registered SBOM matches a critical vulnerability past its remediation deadline, when a configuration scan detects drift from the certified baseline, or when a required

clearing report fails to arrive. Suspension negates the credential status, and policy evaluation then denies access on the next negotiation attempt.

Documented compensating controls substitute for patching, because a fab cannot always redeploy on demand. An applicant presents network isolation, virtual patching, or reduced exposure, and the evaluation facility records the accepted rationale with an expiry date.

4.1.7.5 Software Supply Chain Validation

Threat Position

Participants assemble gateways from container repositories, chart repositories, and application marketplaces. A single compromised transformation container gains direct access to mask data or recipe parameters, which places the supply chain at the same risk level as the connector itself.

Validation happens at three points: at build time, at admission time, and continuously against vulnerability feeds. Each point produces machine-readable evidence.

Control Set by Level

SLSA stands for Supply-chain levels for software artefacts.

Control	Base	Trust	Trust+
SBOM (SPDX or CycloneDX)	Per image, self-generated	Per release, transitive closure verified	Runtime SBOM plus drift detection
Artifact signature	Recommended	Verified at admission	HSM-held keys, dual control
Build provenance	SLSA level 1 or higher	SLSA level 2 or higher	SLSA level 3, hermetic reproducible builds
Registry policy	Trusted mirror recommended	Trusted registry mandatory, no public pulls	Air-gapped mirror for fab-internal deployments
Static analysis	Baseline ruleset, findings triaged	Full property query catalogue	Sound analysis for declared critical properties
Admission control	—	Signature, provenance, property credential enforced	Plus runtime integrity attestation
Base image	Documented	Distroless, non-root, read-only filesystem	Plus measured boot of the node
Vulnerability remediation	Best effort	Defined SLA per severity	SLA plus compensating-control review

Two Certification Patterns

Distribution certification evaluates a curated bundle once. Participants inherit the component level after a configuration conformance scan proves the deployed artifact matches the certified baseline. This path carries most SME participants.

Instance certification applies full component evaluation to participant-built images. Static analysis dominates the evidence set here, since the evaluator holds no prior knowledge of the code.

4.1.7.6 Deployment Variants

Comparison

Aspect	A: Bare metal / self-managed Kubernetes	B: Cloud IaaS or PaaS	C: Operator-hosted
Component scope	Participant build or inherited distribution	Participant build or inherited distribution	Operating company
Operational scope	Participant, L0–L2	Participant L2–L6, provider attestation for L0–L1	Operating company L0–L5
Key custody	Participant HSM	Participant KMS or bring-your-own-key	Declared: custodial or participant-controlled
Reachable level	Trust+	Trust; Trust+ with confidential computing and BYOK	Trust; Trust+ only on a dedicated instance
Design data and recipes	Permitted	Conditional	Prohibited on shared instances
Participant audit effort	High	Medium	Low
Primary residual risk	Change freeze versus patch SLA	Misconfiguration, jurisdiction	Concentration, key custody
Typical adopter	Foundry, IDM, mask shop	Design house, EDA and analytics provider	SME suppliers, subcontractors, distributors

On-Premises and OT Adjacency

Variant A requires IEC 62443-3-3 zone and conduit evidence wherever the gateway touches equipment networks. Where the gateway consumes data from connected tools, the applicant demonstrates alignment with SEMI E187 for those tools and documents the segmentation between the OT zone and the gateway zone.

This variant reaches the highest level and carries the highest participant effort. Fee structures reflect that asymmetry.

Cloud and PaaS

Variant B replaces direct evidence for L0 and L1 with mapped provider certifications. The applicant supplies a control-by-control shared-responsibility matrix and proves ongoing verification of provider certificate validity, not a one-time collection of certificates.

Trust+ in this variant depends on confidential computing and customer-held keys. Without both, the operator retains a technical path to plaintext, which the scheme treats as incompatible with recipe and mask data.

Operator-Hosted

Variant C minimizes participant effort and concentrates risk. The operating company holds L0 through L5, and the participant retains only asset classification and policy authoring.

Shared-instance hosting stays prohibited for design data and recipes. A participant needing those asset classes moves to a dedicated instance with participant-controlled keys, or to Variant D.

An exemplary evidence list with a comparison for criteria is shown in Table 15 in the Appendix on page 114.

4.1.7.7 Minimum Trust Levels for the Data Space Operator

Trust Role

The Data Space Operator acts as trust anchor, service provider, and/or as gateway host. These roles carry different risk profiles, so the scheme assigns a minimum level per service instead of one default level for the organization.

Service	Consequence of compromise	Minimum level
Identity provider and trust anchor	Impersonation opens access to design data across the value chain	Trust+
Onboarding and participant registry	Sanctioned or fictitious entities enter the data space	Trust+
Clearing and evidence logging	Usage-control violations become unprovable	Trust+
Artifact registry and app store	One compromise propagates to every gateway	Trust+
Property attestation service	False claims render usage control silently ineffective	Trust+
Catalogue and metadata service	Capacity and technology-node leakage, catalogue poisoning	Trust
Semantic hub	Semantic drift corrupts yield and carbon-footprint conclusions	Trust
Gateway hosting	Mass compromise, loss of sovereignty	Trust; Trust+ on dedicated instances
Security operations and incident response	Long-term compromise stays undetected	Trust
Support, SLA management, billing	Availability and commercial exposure only	Base

Onboarding

Onboarding establishes identity, not only an account. The operator verifies the legal entity against an authoritative register, verifies the authority of the signing representative, screens the entity against sanctions and export-control lists, and issues participant credentials afterward.

Semiconductor supply chains include entities under export-control attention, so periodic re-screening replaces the one-time check. The operator encodes the screening result and its validity date into the participant credential.

Transparency Obligations

The operator publishes a machine-readable trust list, a credential status list, a sub-processor list, key-ceremony records, and a changelog for the property query catalogue. Participants verify the trust state programmatically instead of requesting documents.

Exit and Continuity

The operator escrows registry and clearing data, publishes a tested migration path for hosted gateways, and commits to a defined notice period. A rehearsed migration remains mandatory, because an untested exit plan tends to fail exactly when a participant needs it.

4.2 Deployment Phase

Our approach to securing the data space gateway in the Deployment Phase includes the preparation of the network architecture used by the gateway components, and the preparation of the platforms on which the gateway components are to be deployed. In the following, we first describe the different steps of our approach in the Deployment Phase, i.e.:

- What to consider when defining a network architecture
- What to consider when preparing the platform
- What to consider when configuring the gateway

4.2.1 Defining a Network Architecture

We propose a high-level network architecture from the perspective of a data space participant. This architecture is segmented into five trust tiers T0-T5, which in a concrete deployment will be . It spans both data space-operated infrastructure and participant-internal networks, and it is therefore defined independently of any specific network segmentation standard. Segmentation within a participant's own IT/OT network is a separate concern and is addressed in Section 4.1.3.

Table 6: Network segmentation approach

Tier	Description	Responsibility
T0 (External/Internet)	External participants, remote identities	Data Space Operator + Participants
T1 (Data Space Edge / DMZ)	EDC connectors, WAFI, Identity Hub	Data Space Operator / Component Operators
T2 (Data Hub / Orchestrator)	Industry Core Hub, criticality proof	Data Space Operator / Participants
T3 (IT Applications)	Enterprise IT, AI systems, data lakes	Participants themselves
T4 (OT Applications)	MES, OPC-UA servers, machines, controllers	Participants themselves

Tiers T3 and T4 are not the responsibility of the Data Space Operator or the Component Operators, but exclusively of the respective participants, who independently segment their internal IT/OT networks. Not every participant operates OT: purely IT-based participants have no T4 and require no OT-specific measures, so the OT-related provisions apply only where OT infrastructure is actually present.

The following principles apply across all tiers:

- Zero Trust. No implicit trust is granted based on network position.
- No direct OT access. Data space infrastructure in T1 must not be directly connected to shopfloor systems in T4. Data extraction from and delivery to OT must occur via secure intermediate storage in the OT-DMZ of the respective participant.
- Separation of control plane and data plane. Policy negotiation and data transfer are kept apart as separate architectural planes.

The internal structure of T3 and T4 is defined by each participant. Section 4.1.3 describes the recommended approach for participants operating OT, including the integration of a data space gateway at the IT/OT interface.

4.2.2 Preparing the Platform

Most gateway deployments in practice run their applications and services as containerized workloads orchestrated through Kubernetes. This chapter therefore turns to the hardening of that orchestration layer: the baseline requirements that ensure the Kubernetes environment itself do not undermine the trust and isolation guarantees established at the platform level.

The following baseline requirements apply to all Kubernetes-based deployments of data space components: only signed container images from trusted registries are used, containers run (where possible) without privileged access and with read-only root filesystems, and network policies enforce a default-deny posture between namespaces. RBAC assigns minimal permissions to service accounts, and hardened operating system images with secure defaults are verified against CIS Benchmarks and BSI Baselines, supported by configuration management and drift detection. Machine-to-machine service identities are managed with a full identity lifecycle, while runtime detection, central and audit logging, and network visibility with alerting and escalation provide the operational feedback loop. For the supply chain, image signing, SBOMs, and reproducible builds are mandatory. NIST SP 800-190 (Application Container Security Guide), NIST SP 800-204A (Secure Microservice Architecture), and the BSI IT-Grundschutz module APP.4.4 Kubernetes serve as normative references.

A proven open-source reference stack implementing these requirements and the corresponding Kubernetes Security Architecture Model, covering both the security level perspective and the tool perspective, is provided in Appendix C.

4.2.2.1 Hardware Trust Foundation

High-security scenarios with strictly confidential data require a critical security level and the assurance level Trust+ according to Section 4.1.5.1. In this case, the verification of the trustworthiness of the platform (state) is required, and a hardware-based trust anchor is needed that provides non-manipulable measurements reflecting the platform and software state. The use of such a hardware-based trust anchor requires trust in the hardware manufacturer and their components used here. For the verification of these measurements, knowledge of reference values that capture a known-good state is required. This is realized in the form of a reference value provider, which is part of the responsibilities the Certification Authority has in the data space ecosystem, as was introduced in Section 3.1.3. Fundamentally, it must be an entity that both Data Provider and Data Consumer trust.

4.2.3 Configuring the Gateway

Numerous parameters and settings must be configured to establish security. These settings include encryption of the data in transit and at rest, the storage type of secrets and the configured usage of hardware vaults, key management, and the choice of post-quantum resilient algorithms where applicable. Table 7 shows an overview of recommended cryptography settings.

Table 7: Encryption Requirements

Area	Requirement

Data in Transit	TLS 1.2+ (preferably TLS 1.3), mTLS for service-to-service, BSI-compliant cipher suites
Data at Rest	AES-256-GCM for databases, volumes, backups
Secrets	Encrypted storage in vault
Key Management	TPM or HSM-based, rotation ≤ 1 year, BSI TR-02102
Post-Quantum	Hybrid PQC for IP-related use cases, as soon as technically feasible

4.2.3.1 Monitoring, Patching and Lifecycle Management

Monitoring and patching are mandatory for all deployed data space components and must occur within specified time frames. Given the increase of AI-based attacks, these time frames must be continuously reduced. Without patching, the platform and the gateway components above lose trust and can no longer be part of the mesh, making patch compliance a condition of data space membership rather than an optional operational practice. The patching responsibility lies on the operator of the infrastructure, which in the role definitions of Chapter 3.1.1 is the Data Space Component Operator, while lifecycle management covers the full path from vulnerability detection, through risk-based patch prioritization, to verified deployment. Continuous configuration benchmarks such as those described in Section 4.2.2 complement patch management by detecting configuration drift between scheduled patch cycles. The concrete time frames themselves are not fixed by this scheme but must be derived from the applicable regulatory framework, such as the Cyber Resilience Act (CRA) or NIS2, depending on which requirements apply to the participant.

4.2.3.2 Deploying Gateway Instances

A company joining the data space for multiple use cases or with different departments may need to consider the deployment of multiple gateway instances. The need for separate gateway instances may arise from the fact that data of different data criticality is used, e.g., in different use cases. As data criticality is mapped to security levels, the need for gateways with different security levels may arise. Data with low criticality can be exchanged using policy- and session-based controls on a shared deployment. Data of medium criticality requires separation through distinct connectors, i.e., multiple gateway instances, while still running on a shared platform. Data of high criticality requires separate physical gateway platforms, in line with the platform security considerations of Section 4.2.2. As a deliberate downward compatibility measure, a company may operate a gateway of Critical security level and also use it for Basic level data if the use case and applicable norms permit a shared platform. The reverse direction is never permitted: lower-level deployments cannot host higher sensitivity data flows.

Actual deployment guidance requires concrete tooling rather than concepts. The following open-source reference stack has proven effective for securing Kubernetes-based data space components (hardware-level measures are excluded here). It is intended as practical guidance: each security function can be substituted by an equivalent tool based on the risk assessment of the respective deployment.

Beyond the toolchain of the table, a complete implementation covers supporting practices from the functional layers in Table 8: hardened and minimal container base images, admission enforced seccomp, AppArmor, and capability baselines, append-only log storage collected outside the cluster, containing metadata only and no payloads, workload identities instead of static secrets with a strict separation of human and workload identities, including Vault backups and regular secret rotation, context enrichment of

logs, for example through GeoIP and autonomous system information, collective IP reputation intelligence with bouncer enforcement at the firewall and ingress layer, as well as hardened virtualization hosts with host monitoring, network segmentation, snapshot based restore tests, and immutable node replacement as recovery strategy.

Table 8: Kubernetes Security Architecture table

Security Function	Reference Implementation
Hardened OS for nodes	TalosOS (minimal, immutable, API-driven), Terraform for infrastructure provisioning
CNI & NetworkPolicy enforcement	Cilium (eBPF-based enforcement)
Network observability	Hubble (network flows, policies, L7 visibility)
Service mesh (mTLS)	Linkerd (application observability, mTLS)
Perimeter firewall / VPN	OPNsense
Network telemetry	ntopng/ntop, Flow Exporter (NetFlow/IPFIX, eBPF)
Admission control / policy engine	Kyverno
Runtime security (sensors)	Falco with Falcosidekick, Wazuh Agent, Kubernetes Audit Logs
SIEM / log management	Wazuh (Manager, SIEM, stream processor), Fluent Bit, OpenSearch
Metrics / monitoring	Prometheus, Grafana
Vulnerability scanning	Trivy CI (in GitLab CI), Trivy Operator (in-cluster scanning)
Registry & supply chain	Harbor (image/signature registry), GitLab CI, Argo CD (GitOps deployment)
Secrets management	HashiCorp Vault with External Secrets Operator
Detection engineering / reputation	CrowdSec (detection, reputation decisions, bouncer enforcement), GeoIP
Incident response & orchestration	TheHive + Cortex (alert triage and response automation)
Security management	Dedicated security management cluster hosting SIEM/monitoring, receiving events, logs, metrics, and alerts from all other clusters

The stack reflects a defense-in-depth architecture with distinct layers covering network, policy, runtime, monitoring, and response. It is not normative, and the concrete tool selection must be verified in each individual risk assessment. Deploying a Kubernetes cluster without these layers, for example, without default-deny network policies, runtime detection, or central logging, is considered insufficient for operating data space components at security level Trust or above. The stack should be continuously audited with automated benchmarks such as kube-bench against the CIS Kubernetes Benchmark, kubescape, and kube-hunter to detect configuration drift and security gaps. The corresponding Kubernetes Security Architecture Model, showing both the security level perspective and the tool mapping, is provided in Appendix C Kubernetes Security Architecture Model.

4.3 Operations Phase

Once the gateway is deployed, the operations phase begins. In this phase, the data space components must be securely operated under suitable organizational practices and procedures. The operations phase includes activities such as:

- Securely operating the gateway components
- Assessing the conformity of the gateway
- Evaluating operational changes

4.3.1 Secure Operation of Gateway Components

Operationally, the gateway components are managed along the control plane and data plane separation established during deployment. The following description applies to the highest security level (Critical); components at Basic or Trust level implement the corresponding subset of these measures, graded down according to their applicable Gateway Security Level.

On the control plane, all management API calls are authenticated using OAuth2/JWT bearer tokens or API keys, management APIs are not exposed openly, and secure defaults with strict policy validation apply. All contract negotiations and data transfer events are recorded in tamper-evident audit logging with log data encrypted in use, in transit, and at rest. This logging is mandatory for self-hosted operation; for SaaS offerings it is required only for the operation of the service itself, with log storage and analysis remaining with the operator while the evaluation of event details and reaction measures lie with the participant. In connector deployments, logical and cryptographic multi-tenancy isolation separates participant contexts, which precludes deployment variants offering only logical separation.

On the data plane, no persistent data storage exists beyond the active transfer session: the data plane acts as a pure transfer mechanism. Access is controlled through token-based access control with short-lived tokens, and fine-grained authorization is enforced at the time of the actual data transfer. For strictly confidential data, confidential computing as specified in Section 4.1.5.1 ensures that data remains encrypted until it reaches the TEE on the consumer side, and the data provider can require remote attestation of the consumer TEE before key release. The data plane itself is protected by a WAFI providing Layer-7 inspection, IDS protocol guarding, DoS/DDoS protection, and content security scanning. All secrets, tokens, API keys, encryption keys, and database credentials are stored in a vault and never in plaintext configuration files, secret rotation occurs at intervals not exceeding 90 days, vault access is restricted to authenticated service accounts, and all vault accesses are audit logged.

Beyond the control plane and data plane measures described above, each connector instance must operate under its own distinct identity, holding a dedicated key pair that is not shared with any other instance or participant context. Private keys used for DIDs and Verifiable Credentials must be protected by hardware security modules (HSM) or Trusted Platform Modules (TPM). The hardening, key protection, and API security requirements applicable to the wallet service that manages these identities are specified in Section 3.3.2.

4.3.2 Gateway Certification and Conformity Assessment

The roles of the Data Space/Federation Level and the Participant Level must be certified by a Conformity Assessment Body (CAB); the corresponding role definitions and separation-of-duties requirements are detailed in Section 4.1.7.3. The scope of the conformity assessments covers the core services of the data space (Identity, Discovery, Monitoring), onboarding service providers, connector-related software

components, participant wallets, sandbox service providers, and value-added services such as the Data Quality Dashboard, BPDM Fraud Prevention, Country Risk, and Sanction Watchlist services, see Section 4.1.7.7 for the minimum assurance levels assigned to these federated services, and Section 4.1.7.1 for how the certification object itself is delimited.

For Semiconductor-X, it remains to be clarified whether an own certification scheme with completely independent audit processes and certification bodies should be established, despite the very high overhead and resources required, or whether integration into existing frameworks is feasible, developing schemes or specific profiles where requirements differ from existing ones. The layered assurance model in Section 4.1.7.2 and the role separation in Section 4.1.7.3 provide a starting structure that could either be operated independently or aligned with existing schemes.

For high-criticality use cases, certification evidence should be provided in a machine-readable form, and the integrity of certified software stacks should be verifiable through remote attestation, so that conformity can be validated automatically at runtime rather than only at audit time. This requirement is already reflected at the Trust+ level in Section 4.1.5.1, which mandates runtime integrity attestation and measured boot, and in the exemplary evidence catalogue item 23 (Appendix D), which lists TPM- or confidential-computing-based runtime attestation as required (✓) for variant A and conditional (○) for B and C.

Discovery Service / Service Catalog

The Discovery Service and Service Catalog are operated as the "Yellow Pages" of the data space, providing an overview of available data and service offerings. Because such a service is attractive for automated abuse, appropriate rate limiting and abuse detection mechanisms must be in place to protect against scraping and automated attacks. The BPN-DID-Resolution-Service performs the mapping between Business Partner Numbers (BPNL) and DIDs and must be designed so that no metadata about actual data transfers can be derived from resolution queries. Participation in the data space must not be made conditional on being listed in the Discovery Service or Service Catalog. Listing remains a voluntary decision of each participant.

5 Case Study: Semiconductor-X

In the following, we describe how our approach was used in the context of the Semiconductor-X project to secure the data space gateways that were employed in the data space. Using the abstract security gateway architecture (see Figure 3) as a starting point, we began with the activities in the design and development phase. We describe in detail how the representative use cases for different security levels were elicited, how the threat model was refined, and how additional requirements of the Semiconductor-X data space and usage control were incorporated into our architecture. We finish the case study with the description of additional requirements that came up during the deployment phase, that also needed to be incorporated into the architecture.

5.1 Use Case-Based Requirements Elicitation

In the following, we discuss three exemplary use cases from Semiconductor-X, one for each level of anticipated data criticality: semi-public data (e.g., digital nameplates or digital product passes), confidential data (e.g., machine utilization, product volumes, or energy consumption), and strictly confidential data (e.g., machine configurations, blueprints, and recipes whose disclosure directly impacts competitiveness).

Combined with the threat model, these use cases allow us to derive the security requirements needed to properly protect the exchanged data.

5.1.1 Actors and Roles

While the three use cases differ in the exchanged data and its criticality, they share a common set of actors. Every data exchange takes place bilaterally between a **data provider** and a **data consumer**. The data provider decides which data is shared with which participants, defines the policies and usage conditions attached to its data assets, and operates the data sources in its own IT and OT infrastructure, ensuring that data sovereignty always remains with the provider. The data consumer uses received data strictly in compliance with the negotiated contracts and policies, complying with purpose limitations and the need-to-know principle. Given the bidirectional nature of the system, most participants act in both roles simultaneously: a manufacturer receiving data from its suppliers typically provides derived data to its own customers.

Beyond this distinction, the use cases involve the following recurring actors:

- **Manufacturer (OEM):** Produces OT equipment and machinery and provides accompanying information, such as operational documentation, guides, and data sheets, from its own data repositories connected to the data space.
- **Owner/Operator:** Owns and operates OT equipment and machinery for productive work. It is interested in using the acquired maintenance reports and information and trusts the maintenance provider to service the infrastructure.
- **Maintenance Provider:** A service provider that manages and executes the maintenance process on behalf of owners/operators. It employs field technicians who rely on the manufacturer's operational documentation and guides for the maintained equipment.
- **Supply Chain Manufacturers:** In the supply-chain-oriented use cases, manufacturing companies exchange data (e.g., PCF or traceability information) along the product supply chain, each acting as data consumer towards its suppliers and as data provider towards its customers.

In all scenarios, each participant operates a gateway as the controlled interface between its internal IT/OT environment and the data space. Its architecture and the further roles involved in operating the data space are detailed in Section 3.

5.1.2 Basic Level Use Case

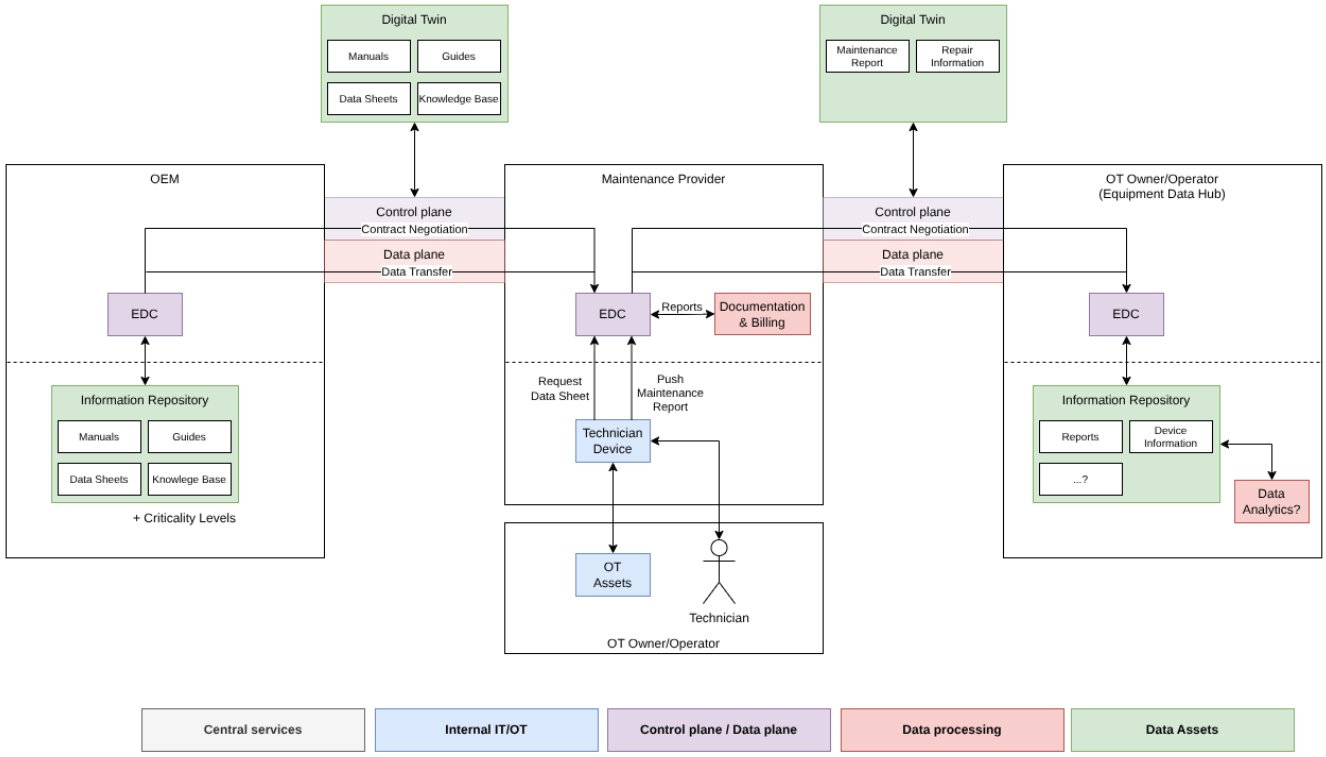


Figure 8: Basic security use case: Exchange of equipment documentation and maintenance reports between OEM, maintenance provider, and OT owner/operator.

Use case one is a maintenance scenario, where an OEM provides information such as manuals, guides, data sheets, and knowledge bases to a maintenance provider, who then services an OT owner's or operator's infrastructure. In turn, the maintenance provider provides maintenance reports and repair information to their customers, who can use them for documentation, billing, and process optimization. The goal of this data sharing scenario is that the data does not become publicly available, while the companies themselves entrust each other with this information.

Figure 8 shows the resulting architecture. Each participant operates a data space connector, specifically an EDC. Contracts are negotiated over the control plane, while the assets themselves are transmitted over a separate data plane. The exchanged assets are modeled as digital twins: the OEM offers manuals, guides, data sheets, and knowledge base articles from an internal information repository, while the maintenance provider offers the resulting maintenance reports and repair information to the OT owner/operator, who collects them in an equipment data hub. The field technician's portable device connects to the maintenance provider's infrastructure to request relevant data sheets on demand and to push completed maintenance reports back.

Although this is the least critical of the three scenarios, it shows that assets of different criticalities coexist within a single use case, so data must be classified and protected per asset, rather than per scenario. Manuals and guides are semi-public and their confidentiality is optional. But integrity and authenticity must still be protected, since manipulated maintenance instructions could lead to incorrect servicing and damaged equipment. Maintenance reports, on the other hand, may contain machine-/production-specific information and are therefore confidential. Usage control plays only a minor role. Optionally, the use of

maintenance reports may be restricted to the maintenance provider (e.g., for billing purposes) and the owner/operator.

5.1.3 Trust Level Use Case

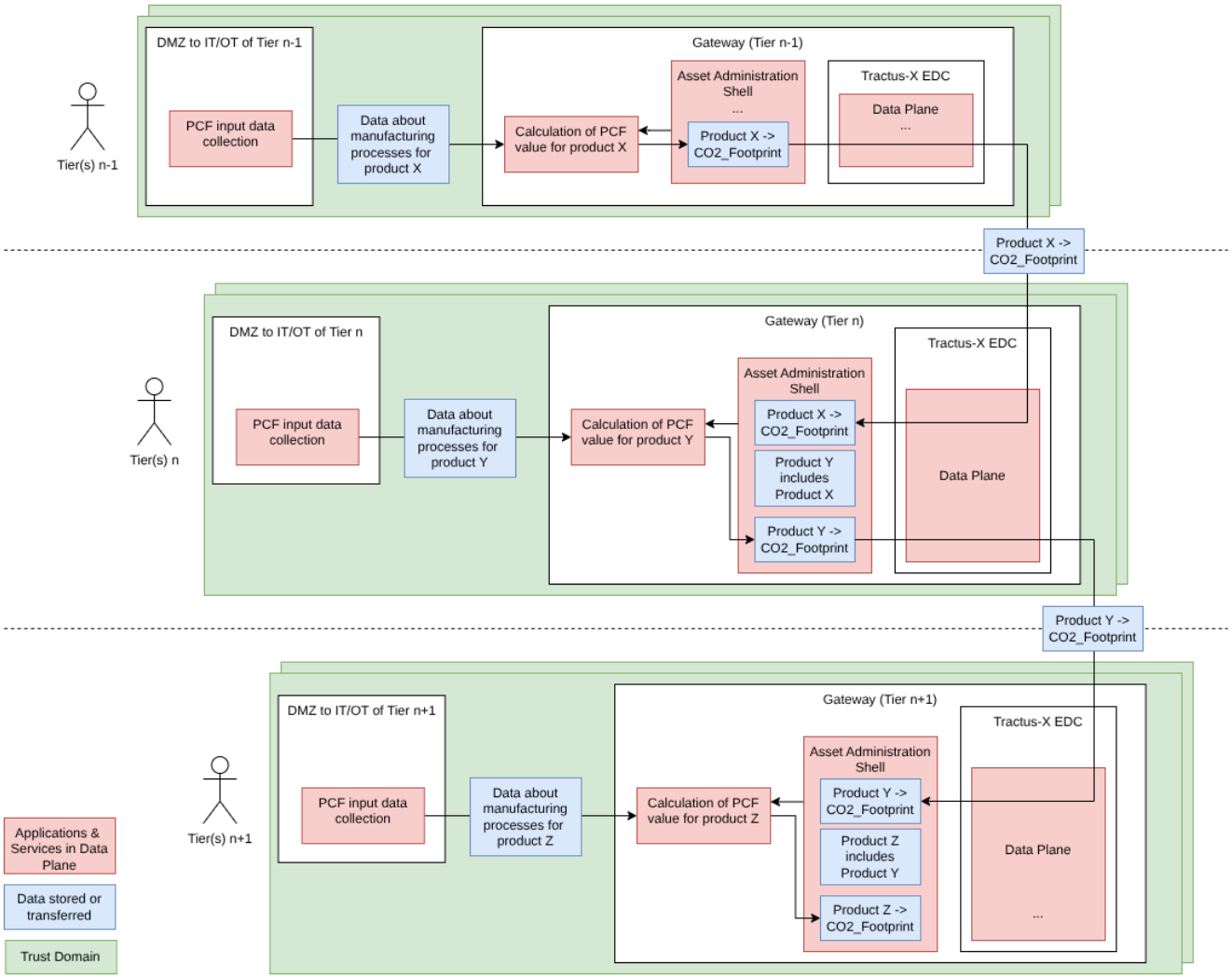


Figure 9: Trust-level use case: Recursive Product Carbon Footprint calculation along the supply chain, where each tier aggregates received PCF values with local manufacturing data inside its own trust domain and forwards only the resulting product PCF.

In the second use case, a Product Carbon Footprint (PCF) is calculated for each tier of a production chain, which constitutes a major data space application driven by sustainability regulation. Since the calculated PCF can be used to infer sensitive production details, only the directly following tier n is allowed to process this information, and even there, access must be restricted to the personnel or departments required. All PCFs of tier n-1 are combined, together with the production reports of manufacturer n, to calculate its own PCF, which is then provided in the following tier n+1, explicitly withholding information on any previous tier. Again, even when data is shared, it must only be processed in a predefined way and not freely redistributed.

In detail, Figure 9 shows three adjacent tiers of such a production chain. Each tier consists of its own trust domain, spanning from the DMZ to its IT/OT infrastructure to the gateway connecting it to the data space. PCF input data is collected from the shop floor. A calculation service in the gateway then combines it with the PCF values received from the preceding tier and stores the result as a submodel of the product's Asset Administration Shell (ASS). Only this aggregate is offered via the data plane to the following tier. Raw process

data never leaves its trust domain, making the aggregation step an effective data minimization mechanism. Since high-quality PCF values are calculated directly from OT sensor data, data integrity must be ensured along the entire chain from the OT source to the consuming tier.

For assets of this criticality, authentication and access control must be considered on the level of departments rather than whole companies, and the agreed usage restrictions (e.g. limiting PCF data to the PCF calculation purpose and prohibiting its disclosure to uninvolved participants) must be technically enforced on the consumer side. In addition, attributes that could reveal sensitive company information, such as precise location data, should be excluded from the shared PCF submodel before sharing.

5.1.4 Critical Level Use Case

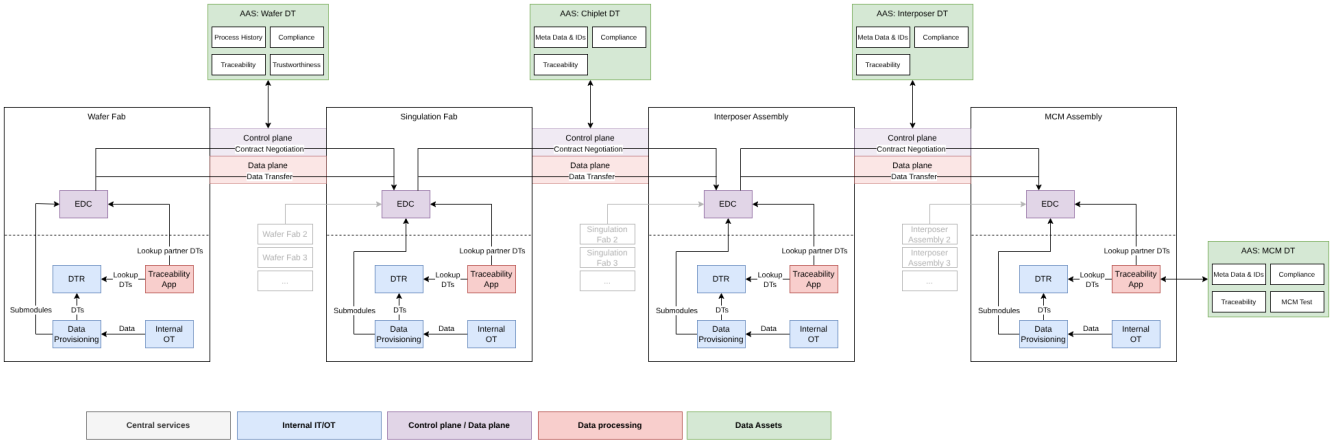


Figure 10: Critical-level use case: Exchange of strictly confidential traceability data along a heterogeneous integration supply chain, from wafer fab to MCM assembly.

The third use case allows for tracing parts and materials along a supply chain to help manufacturers quickly narrow down quality issues, identify defective batches, and reduce overall costs. However, sharing traceability information also poses a significant security risk: in the hands of competitors, this data can be used to deduce knowledge about internal production processes and fab utilization. Since production data is generally known to be critical, especially in semiconductor manufacturing, we take the traceability scenario as an example for a use case exchanging strictly confidential data. This can be seen in Figure 10, which depicts a heterogeneous integration supply chain consisting of wafer fab, a singulation fab, an interposer assembly, and a multi-chip module (MCM) assembly.

Like the figures before, the control and data plane are established bilaterally between the gateways of adjacent supply chain partners, and communication happens between neighboring entities along the chain. At each stage, traceability data is collected from the internal OT and populates the digital twin of the respective product (the wafer, chiplet, interposer, and the MCM). AAS submodels are used for metadata, process history, compliance, and traceability. A locally operated traceability app manages this data and resolves the digital twins of partner products via Digital Twin Registries (DTRs), which never communicate with each other directly. All lookups are proxied through the gateways, so that every access to twin metadata is subject to the same contract-based authorization as the data transfer itself. Each participant may maintain such relationships with multiple partners, but each relationship remains an independent bilateral exchange.

As both the metadata and the traceability submodels are classified as strictly confidential, the full set of protection goals applies, including non-repudiation, and the enforcement of usage control policies becomes mandatory. The negotiated policies are also considerably more detailed. For example: traceability data of a specific product must only be used by authorized people of sufficient seniority, only if a quality issue relating to this product has been found and must not be shared with uninvolved participants (need-to-know principle). Enforcing such person- and event-bound restrictions exceeds gateway-level enforcement and must extend into the data processing applications themselves.

To summarize, while different topologies may exist, e.g., a chain of providers and consumers or a central data consumer that receives information from multiple providers, all these scenarios can be broken down to each connection between one data provider and one data consumer. This will form the basis of our architecture design. The only remaining difference is the criticality of data, which will necessitate different or additional measures depending on the protection needs of this data. In Section 4.3.1, these three criticality levels are mapped to corresponding gateway security levels (Basic, Trust, and Critical) that translate the protection needs into concrete security controls.

5.2 Threat Model

The following sections provide an overview of the assets, attackers, and the architecture used to model, analyze and afterwards discuss the security aspects of the Semiconductor-X Data Space. The threat discussions summarize the most prominent information covering the ecosystem and the components for the data space and the threats evolving within the overall setup. The threats were analyzed, categorized, and prioritized with a corresponding justification. However, not all possible threats are listed in this document, but rather bundled related threats for better readability and comprehension.

5.2.1 Asset Model

As in a typical data space, the primary assets requiring protection are the data exchanged between participants. In the context of Semiconductor-X, these assets comprise data related to the production and management of hardware components, including digital twins, blueprints, manufacturing data, and related information. In addition, the infrastructure itself constitutes a relevant asset, including software components, configuration files, credentials, and cryptographic key material used for the authentication of identities and services.

5.2.2 Architecture Model

The following sections illustrate and explain the modeled architecture, trust boundaries, and the security discussion of connections between the following entities.

Data Space Communication from Provider to Consumer

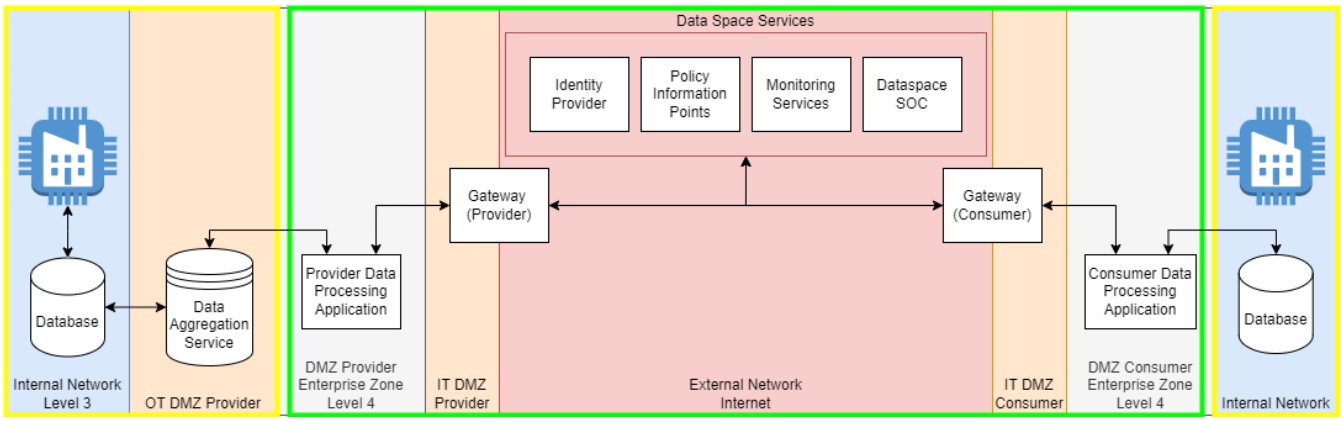


Figure 11: Overview of the Semiconductor-X-Architecture: Green denotes data space security realm and yellow denotes the OT security realm of participants (The levels originate in the Purdue-model)

https://www.pera.net

The figure above shows the main architecture for the transmission of data in the Semiconductor-X data space. The security mechanisms of the data space marked in green cannot reach into the security realms of the OT on both sides of the data flow, shown in yellow.

The boundaries modeled include:

- Internal Network (Level 3)
- Operational Technology Demilitarized Zone
- Internal Technology Network
- Internal Technology Demilitarized Zone
- The Internet boundary
- The data space technical end-to-end security is marked green.
- The data space organizational end-to-end security is marked yellow.
- The components modeled include:
- Provider Device: a representative for factory machine creating sets of data
- Provider Database: represents an internal database collecting the information
- Provider Data Aggregation Service: a service for making the local factory data available to the provider's internal network
- Provider Data Processing Application and Provider Gateway: these are part of the data space boundary end-to-end security scope
- Consumer Data Processing Application and Consumer Gateway: these are part of the data space boundary end-to-end security scope
- Consumer Database: the local data sink or source which is outside of the protection of the end-to-end security of the data space.

Data Space Services

The components modeled include:

- A gateway (either provider or consumer, all gateways communication with these data space services)
- Identity Provider: which provides access and authentication tokens and manages the PKI of the data space

- Policy Information Point: provides information concerning contracts which can be retrieved upon need
- Service Catalog: provides information about available sets of data and the corresponding provider addresses and conditions

Data Space Security Operation Center

Additionally, all components which are linked via the internet within the data space must report to a local Security Operation Center and provide information about security incidents to measure the status of the data space. To avoid redundant argumentations the components Identity Server, Policy Information Point, and Service Catalog are handled as a similar entity in the threat modeling at this data flow.

5.2.3 Tooling and Modelling Procedure

Microsoft Threat Modeling Tool 7 was used for the practical realization of the modeling process. The tool provides a graphical user interface where all components that should be analyzed can be positioned in defined trust boundaries and connections between the components can be defined. For both connections and components security properties can be configured. As shown in the figures below.

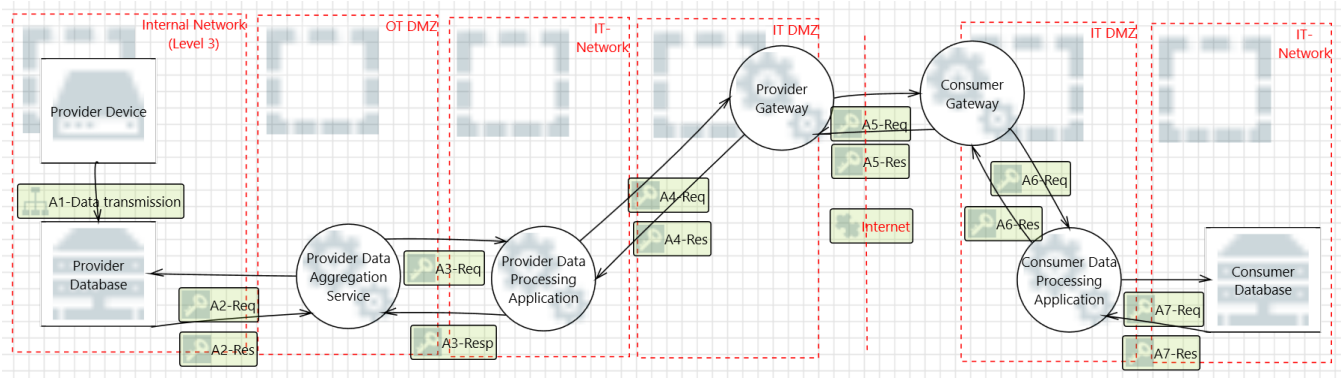


Figure 12: Modeled Excerpt of Architecture in MS Threat Modeling Tool

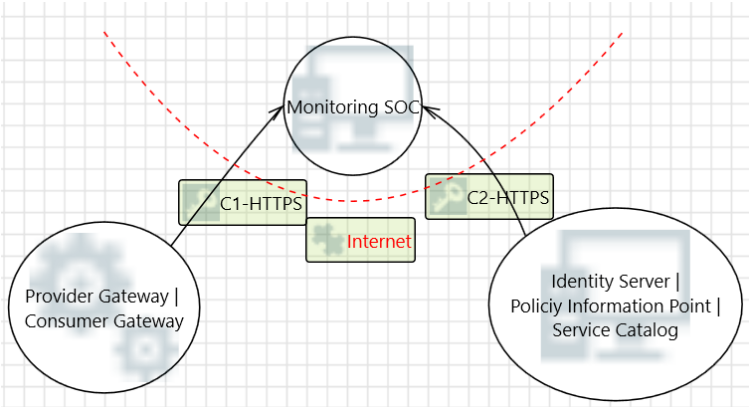


Figure 13: Modeled Communication between Data Space Components and the Monitoring SOC

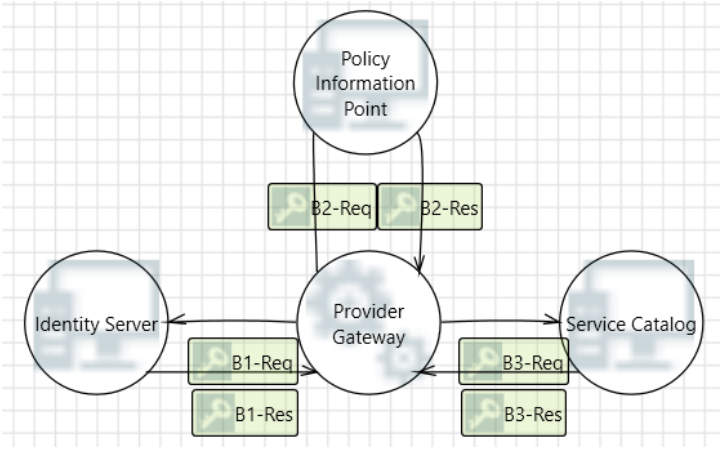


Figure 14: Modeled Communication between the Central Services and the Gateway

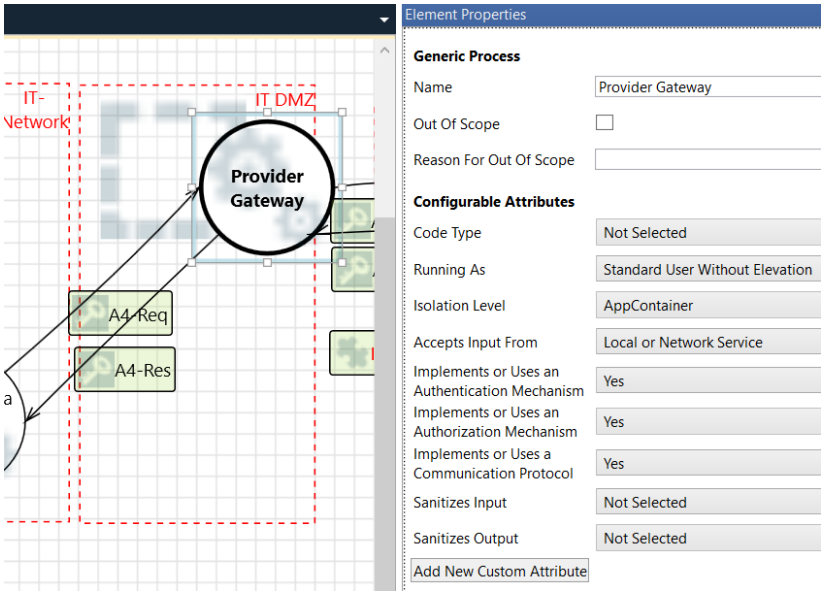


Figure 15: Settings of Component Properties and Attributes

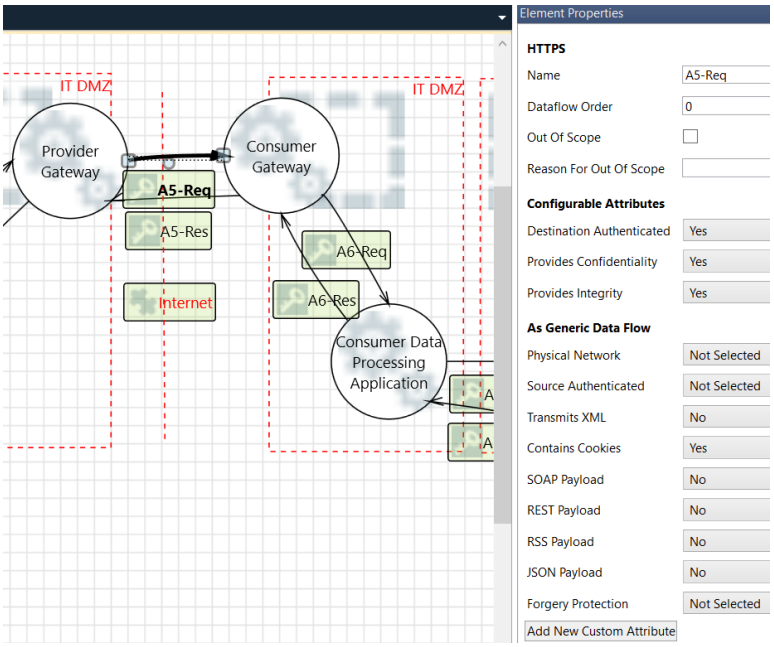


Figure 16: Settings of Connection Properties and Attributes

The main benefit of the tooling comes into play when the featured auto-generation of threats is used. For the tool two parts are important: first the architecture model consisting of components and connections and the threat template, which provides rules where threats must be considered. The figure below provides a sample of a defined threat.

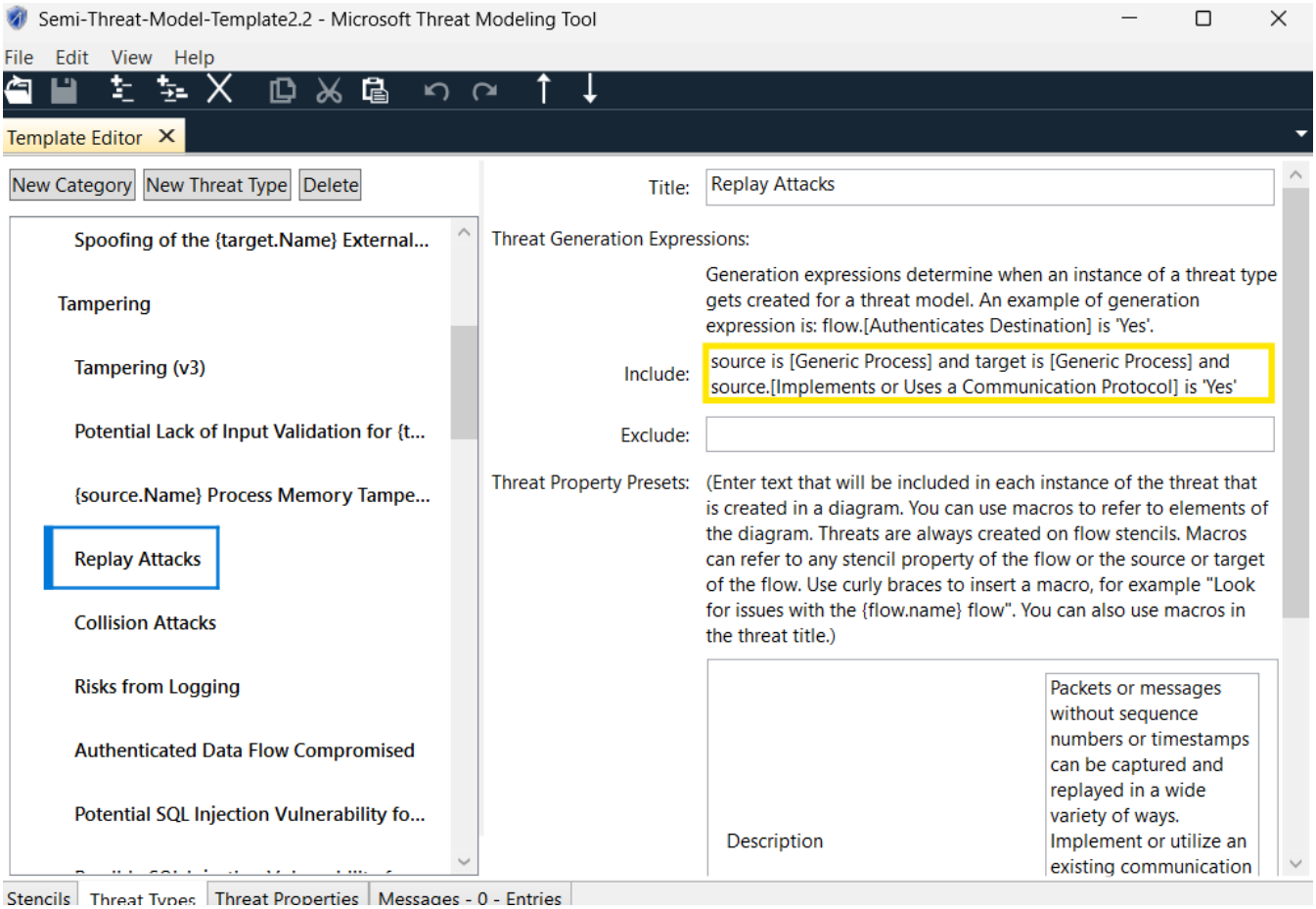


Figure 17 Overview of Threat Definition in Template: The condition where the threat must be applied is marked in yellow for the threat "Replay Attacks"

After the desired threats are defined the tool automatically adds an entry to any position where the threat may appear according to the defined logic of the template entry. Using our first approach we achieved more than 150 typical threats which need to be categorized as either mitigated, not applicable, or as in need of investigation. Each decision was accompanied with justification and a choice of priority and criticality for each entry.

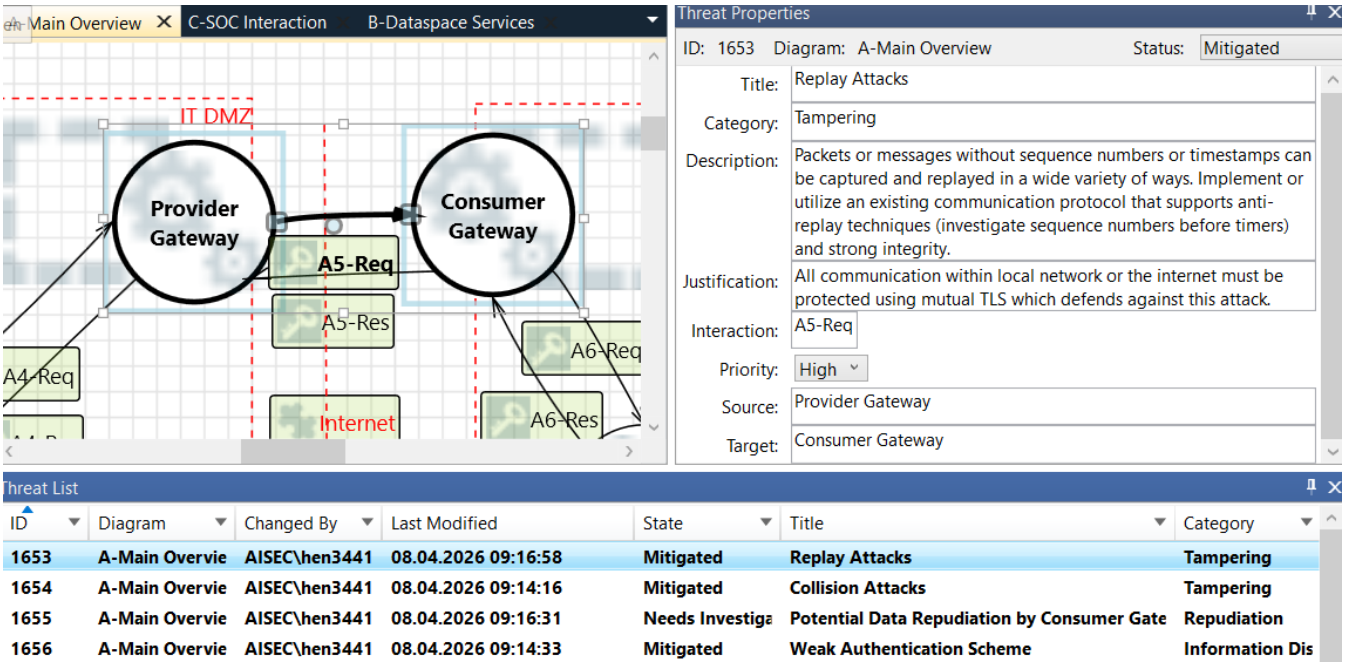


Figure 18 Overview of Generated Threat List with Threat Porperties shown

In an iterative process the list of threats was improved to the context of Semiconductor-X to reach a sufficient level of coverage. In the end the tooling is not dependent on MS Threat Modeling Tool 7, but an export of the data can be made into an CSV file for further processing. Any other tool which can structure the threats may serve the same purpose. MS TMT7 was used due to its simplicity and a suitable set of features as a first attempt to model the architecture.

The summary of all analyzed threats is shown in the following sections. In detail, we address how these threats are already acknowledged in existing requirements or if a use-case specific requirement must be defined. For the sake of readability, the full threat report is not contained in this document.

5.2.4 Mitigated Threats

The following threats are already addressed by the existing requirements defined in Section 4.2.3.

Communication and Authentication related Threats

Numerous standard threats are present whenever a client-server-based web application is in use. The most typical issues in the data space here mostly involve message tampering, spoofing of sender and receiver, package collision attacks, message replay etc. The solution for all these communication related threats is basically mitigated using this justification:

“The Local PKI within the organization’s OT and IT must provide corresponding certificates and all local services must authenticate using Mutual TLS.”

“All communication within local network or the internet must be protected using mutual TLS which defends against this attack.”

Thus, these threats can be addressed when fulfilling the requirement for Access Control and Authentication and do not constitute a new, use-case specific requirement.

This covers among others the following type of issues:

- Message Integrity Attacks (Replay, Collision)
- Weak Authentication Scheme
- Spoofing Destinations and Spoofing Sources

Software Vulnerabilities

All complex components can potentially be victim to bugs, design flaws or other vulnerabilities which lead to remote code execution or unwanted privilege escalation of attackers. However, threats of these types are mitigated using the following justification within the data space:

“All components of the data space must receive regular software updates and techniques such as remote attestation, must ensure that only accepted software versions are executed which have received proper security evaluations.”

“Centralized components which rarely require functional modifications must provide mechanisms which prove that a non-modified code base is running. This code base should be vetted using software verification to guarantee the absence of vulnerabilities. Additionally, all components must only accept sanitized network inputs.”

The general requirement for platform integrity encompasses this threat. It demands that software is not manipulated during system state. This encompasses that the software itself is trustworthy, i.e., trusted to not have vulnerabilities.

This covers among others the following type of issues:

- Privilege Escalation
- Remote Code Execution
- Weak Access Control

5.2.5 Not Applicable / Out of Technical Scope Threats

These threats cover issues which cannot be covered by the technical security solutions implemented by the data spaces. These can rather be solved using organizational measures, for instance, by requiring a certified minimum-security level within the factory environment during onboarding.

Internal Network and OT Network Threats

When modeling the complete data ecosystem, including the area of Operational Technology, where the wavers are being printed and factory machines generate datasets, it is clear that numerous security issues are involved in this state. For instance, considering spoofing a database within a factory to create a data leak, or considering old machines, which cannot be patched to the latest communication standards etc., these issues cannot be covered by a data space specification. For the completeness of the data processing perspective these threats are listed but marked as not applicable as these issues must be dealt with by the corresponding organization and are out of scope for the end-to-end security focused on within the data space. The following justification is applied for these issues:

“As the two components belong to the same party this is highly unlikely a problem. Additionally, the component can only be reached from within the organization’s network, which leaves attacks to internal attackers.”

This covers among others the following type of issues:

- Spoofing of data source and sinks
- Risks from Logging
- Weak access control for local resources
- Potential process crashes
- Interrupted Data Flows
- Elevation using Impersonation
- Potential Data Repudiation
- Potential Data Manipulations at the Source
- This applies for both the provider and the consumer side of the data flow.

5.2.6 Threats that need Investigation

The following threats are not fully addressed by existing security requirements.

Denial-of-Service and Availability Threats

Threats that are not covered yet by the security discussion involve availability issues. In case a typical denial-of-service attack is launched and gateways of the data space or even the critical central components, such as the identity provider are blocked, the consequences would be severe.

Threats Involving the Human Factor

Threats that are not covered yet also involve the human factor. In cases where a web GUI is used by an administrator an attacker may prepare manipulated links to cause request forgeries or cause password leaks and proper responses for such cases need to be discussed.

Data Flow Threats

Another issue arises in the case of Attempted Data Repudiation, where a consumer claims data has not been received. There is currently no mechanism to verify completed data flows, such as mutual signature receipt to verify the integrity of data or the receiving of given data sets. Here the scope of the threat modeling needs to be discussed, also.

Supply Chain Attacks

Participants of the Data space may provide intentionally malicious data or components with backdoors. This threat is acknowledged in the platform integrity requirement as well as proof of platform integrity and capabilities.

5.3 Additional Requirements for Semiconductor-X

General requirements were already defined in Section 4.1.5. However, this section details additional requirements derived from our in-depth threat analysis which are specific to Semiconductor-X. We define the following additional requirements and increase the importance of the third even for the basic security level:

Human Factor Limitation: The security of the realm outside of the data space on the layer of the human factor must also be covered. The typical areas include administrators or staff working within the OT. Solutions would be a higher level of security accreditation of staff which can be gained from security training and personal background checks.

Data Flow Protection: The flow of data from the producing entity all the way to the consuming entity of a processing chain must be integrity protected and data packages must be digitally signed by the sender and the receiver after retrieval. The aim is to cover non-repudiation of data transmissions at the consumer and to mitigate manipulations or forged data sets before entering the data space.

Availability: The Availability of information flows from the OT into the data space must be already guaranteed and covered independent of the data space security mechanisms. A solution would be failure compensating architectures and technical approaches, such as load balancers and redundant installations of servers, to cover for system failures. Hence, we determined that the Availability requirement as detailed in Section 4.3.1. is also needed for our Basic Security Level use case.

5.4 Usage Control Enforcement in Semiconductor-X

Usage control allows data space participants to regulate which operations are permissible or prohibited on shared data. Typically, data usage rules are specified by exchanging a signed contract written in a machine-readable digital policy language such as the Open Digital Rights Language (ODRL). To provide adequate data protection, the usage control architecture must ensure two requirements. First, the used ODRL policy schemes must be expressive enough to allow the definition of legitimate and unwanted data usage scenarios in the respective use cases. Second, the data space gateways must be able to understand and technically enforce the defined usage rules on the received data.

To achieve the first requirement, Semiconductor-X makes use of the ODRL policy schemes that have been defined by Catena-X[5]. Table 9 provides an excerpt of the available ODRL policy constraints that are relevant for the presented Semiconductor-X use cases. This allows data providers and consumers to come to a legally binding contractual data usage agreement prior to the data exchange. However, the technical enforcement of exchanged usage rules as part of a secure data space gateway (c.f. chapter 4) has not been adequately addressed so far. To achieve the highest security level (Critical), future gateway implementations must also include usage control enforcement components that are capable of enforcing these restrictions on a technical level all the way into the data processing applications.

Table 9: Excerpt of ODRL policy constraints specified in CX-0152 that are relevant for Semiconductor-X

ODRL Policy Constraint	Values	Informal Description	Relevant for Use Cases
AffiliatesRegion	cx.region.all:1 cx.region.europe:1 cx.region.northAmerica:1 cx.region.southAmerica:1 cx.region.africa:1 cx.region.asia:1 cx.region.oceania:1 cx.region.antarctica:1	Allow or prohibit data sharing with affiliated companies in certain regions.	Basic, Trust, Critical

[5] https://catenax-ev.github.io/docs/standards/CX-0152-PolicyConstrainsForDataExchange

BusinessPartnerNumber	BPNL	Limit data usage to certain data space participants.	Basic, Trust, Critical
DataUsageEndDate	Date	Allow data usage only for a specific duration. Consumers must delete the data after this date.	Trust, Critical
UsagePurpose	cx.pcf.base:1	The data consumer may use the data for the purpose of calculating product carbon footprints in the supply chain.	Trust
UsagePurpose	cx.quality.base:1	The data consumer may use the data for the purpose of identifying and tracing product quality issues in the supply chain.	Critical
UsageRestriction	cx.thirdParty.forbidden:1	The data consumer must not redistribute the data to third parties.	Basic, Trust, Critical
UsageRestriction	cx.manipulation.forbidden:1	The data consumer must not modify the received data.	Basic, Trust, Critical
UsageRestriction	cx.derivations.forbidden:1	The data consumer must not create derivative works from the data.	Basic, Trust, Critical
UsageRestriction	cx.dataProviderRemoval.forbidden:1	The data consumer must not remove identifiers linking the data to the data provider.	Basic, Trust, Critical

5.5 Gateway Design and Architecture Based on Use Case Requirements

The following section focuses on the executing platform for gateways and less so on the software and services of the gateway itself. We implement all security controls that are required for a critical security level. For more specific use cases, further requirements may apply that are not yet implemented, e.g., finer-grained network or resource separation. As a result, the presented platforms cannot be taken as complete solutions but rather core building stones that will greatly reduce the amount of additionally required components needed to fulfill a use case in all its aspects. We describe two approaches, one self-hosted gateway and one gateway deployed in a cloud.

In both cases, we use a hardware-based trust anchor that provides non-manipulable measurements as was also described in Section 4.2.2.1. This provides the foundation for attesting and verifying all other requirements.

5.5.1 Self-hosted Gateway

For a self-hosted gateway, we assume that the hardware is controlled by the data space participant or a Data Space Operator. In both cases, the hardware is assumed to be running only software necessary for the gateway, and either the participant or the Data Space Operator are responsible for maintaining the software and hardware.

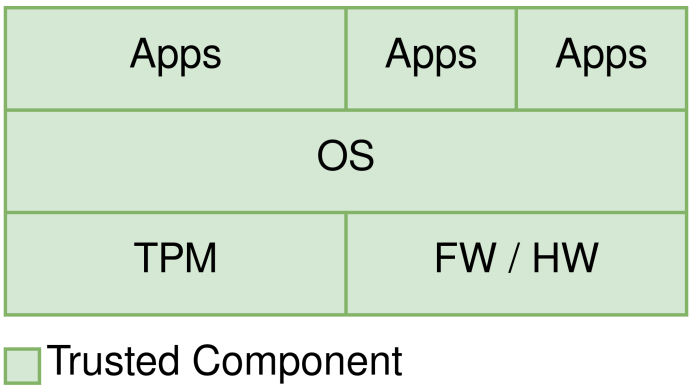


Figure 19: Hard- and software stack of an exemplary self-hosted setup

We describe an exemplary self-hosted architecture based on a Trusted Platform Module (TPM), as can be seen in Figure 19. It forms the trust anchor and is responsible for storing measurements and providing these in a signed attestation report, called a quote. The signature allows a verifying party to ensure the hardware is authentic and trustworthy, i.e., was built by a known hardware manufacturer and behaves in a way that the measurements cannot be manipulated, or signatures forced over arbitrary data.

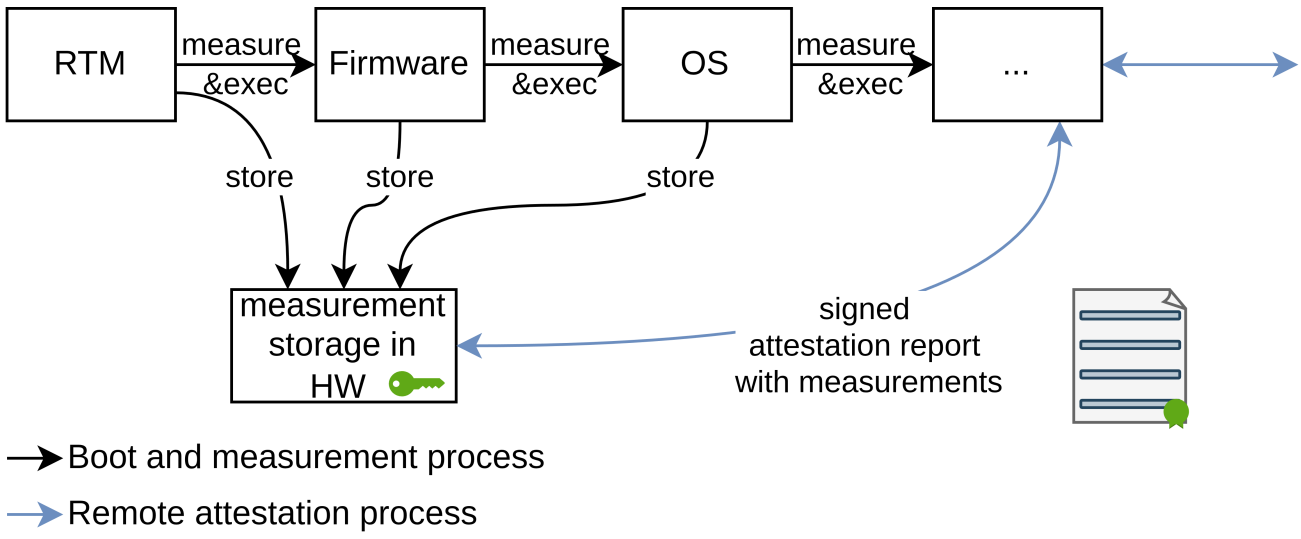


Figure 20: Simplified measured boot and remote attestation process

We use a measured boot approach that ensures that each component measures the next component in the boot chain before passing control to it, as shown in Figure 20. In this example, the root of trust for measurement (RTM) starts with measuring parts of the UEFI, which then measures further firmware, which in turn measures the OS. This way, the later component cannot manipulate its own measurement. Together

with the Integrity Measurement Architecture, measurements representing all executed software can be stored in the TPM for later verification. Notably, this includes all running applications, including the gateway software and all related services, such as the control plane, data processing apps, and policy enforcement points.

All communication that transfers data must only be possible after a remote attestation of the recipient of the data was successful.

In detail, this approach provides data-at-rest and data integrity protection since the gateway is in the hands of a trusted entity. Additional disk encryption is applied using a disk encryption key sealed to the TPM, which ensures secure storage of keys. Similarly, data-in-use is further improved by using CPU-provided Transparent Memory Encryption that can be enabled in the UEFI, which transparently encrypts all data stored in memory. This UEFI configuration is measured and can be verified with the TPM remote attestation as well. Furthermore, data in transit is protected by using only secure communication channels whenever transferring data. When such a communication channel is established, additional software part of the gateway applications performs access control and authorization. If required, an external identity service is contacted for this. Additionally, usage control is enforced by an application. Platform integrity is provided with the choice of software and measured boot. As the entire system state can be attested and verified against a configuration of the system that has been deemed secure, its platform integrity can also be proven. If the sensitivity of the data demands it, one gateway can be limited to handling one type of asset only to ensure physical separation. Availability is the only requirement that cannot fully be addressed by the gateway itself but has to be addressed by the surrounding infrastructure. While the gateway can apply rate-limiting, connection timeouts, and more to ensure attacks such as SYN floods are mitigated, more complex load balancing and resource management must be performed directly by the Data Space Operator or client.

5.5.2 Cloud-hosted Gateway

Fundamentally, the cloud moves the ownership of hardware and large parts of the software stack from a client to an untrusted cloud provider. This significantly decreases management and maintenance aspects for clients as well as the need for physical computing resources, while allowing for more use-case-specific, dynamic approaches. However, this establishes a cloud provider with more control than the client: without additional mechanisms, a cloud provider can access any data processed by hosted apps, and thus, all data that is stored or sent as well. To restrict the provider's control again while still making use of the provider's compute resources, we use Confidential Computing.

A Confidential VM, similar to a normal VM, is comprised of firmware, an OS, and applications running on top. These components are measured by the Confidential CPU for later remote attestation purposes. Depending on the chosen hardware, e.g., AMD SEV-SNP, Intel TDX, or Arm CCA, implementation-specific aspects must be taken into account; however, these do not impact the design itself. The user-space components running inside a Confidential VM are the same ones running on a self-hosted gateway.

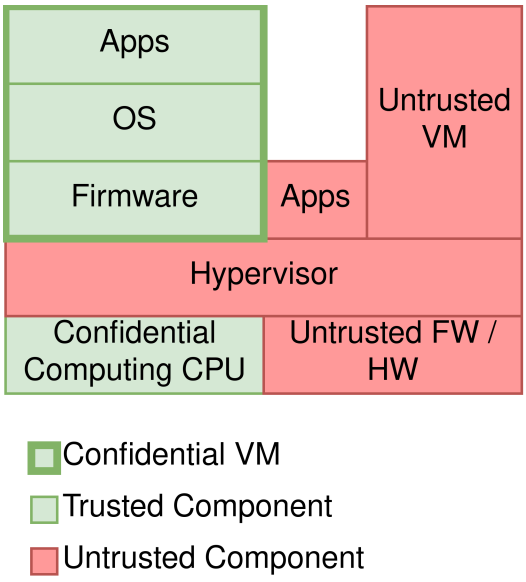


Figure 21: Hard- and Software stack of an exemplary cloud-hosted setup

As Figure 21 shows, a cloud-hosted platform is comprised of one Confidential Computing CPU and, further, untrusted hardware and related firmware. This includes hard drives, network interfaces, and memory, among others. On top, a hypervisor and OS is running that can deploy Confidential VMs, untrusted VMs, and other untrusted applications. The Confidential CPU enforces that no other software, including software of the platform owner, can access or manipulate memory belonging to the Confidential VM. This way, all other untrusted components are excluded from the Trusted Computing Base (TCB) of the Confidential VM and do not need to be trusted.

This exemplary architecture can again be deployed by a client or by a trusted Data Space Operator for a client. Same as for the self-hosted gateway, this affects management and maintenance responsibilities but does not fundamentally change the architecture. For more sensitive data or specific use cases, physical hardware separation may be mandated that cannot currently be fulfilled using this setup, as this requires the cloud provider to provide and ensure physical separation. In the future, norms may be adapted to reflect this change in operation mode, or new gateway designs may provide required separation.

Data in use protection is provided by the Confidential CPU for each Confidential VM. Data at rest is protected using replay-protected Authenticated Disk Encryption, while the disk encryption and communication keys remain safe in the protected memory. Additionally, a (proof of) platform integrity is ensured in the same way as in the self-hosted setup, as are authentication, access control, and policy enforcement. Again, availability must also be handled outside of the gateway for more complex attacks or load balancing aspects.

5.5.3 Protection Against Use Case Specific Threats

Both the self-hosted and the cloud-hosted gateway approaches address the use-case-specific threats in the same way. The UC-specific threat involving the human factor is addressed by greatly limiting their input and access to the data processing itself. The data consumer of the data space cannot access the raw source data that was provided by the data provider if not explicitly allowed by policy and thus forwarded to them by the gateway. Otherwise, they only get access to the processed data. However, this does not fully mitigate what happens to the processed data at the consumer side, as some sort of data exit point must always exist. Similarly, the compromise of source data cannot be fully addressed by the gateway, either. Further

measures must be taken to establish what “valid” data is supposed to look like, e.g., before the gateway forms a decision to process this data, or if the data source must pass further technical steps before they are allowed provide data. As was stated earlier, non-technical measures may be required for this.

5.6 Deployment Requirements

The deployment of the gateway platforms for the Semiconductor-X use case builds directly on the general deployment requirements specified in Chapter 4.2, which apply in full: the zone and conduit model, the hardware trust foundation, the Kubernetes hardening baseline, the gateway constellation rules, the encryption requirements, and the reference stack. The architectures implement these requirements for the Critical security level, so no additional generic deployment requirements are introduced here.

Beyond this baseline, the usage-control use case imposes a small set of deployment-specific requirements. First, remote attestation is not merely a platform property but a precondition for every data transfer: the deployment must integrate the attestation infrastructure, including a reachable attestation service and a reference value provider acting on behalf of the Certification Authority, and data transfer endpoints must refuse communication with any peer whose attestation fails or cannot be verified. Second, the attested software stack must explicitly include the usage-control enforcement points and data processing applications in addition to the gateway itself, so that measured boot and integrity measurement cover the complete usage-control path. Third, only platforms capable of confidential computing qualify for deployment in this use case, since data-in-use protection until the Trusted Execution Environment is mandatory at the Critical security level. Finally, the deployment must address the availability threats identified in the threat model) through load balancing and multiple running instances operated by the surrounding infrastructure, as availability cannot be fully resolved by the gateway itself.

6 Smart Industrial Data Hub – Secure Data Provisioning and AI-Assisted Security

Industrial data spaces such as Semiconductor-X, Catena-X, Factory-X, and Chem-X increasingly connect cross-industry supply chains. Companies participate in several data spaces simultaneously, exchanging information with upstream suppliers, downstream customers, and industrial partners. With potentially thousands of new use cases and applications emerging, secure and efficient data provisioning becomes a major challenge, particularly for small and medium-sized enterprises (SMEs).

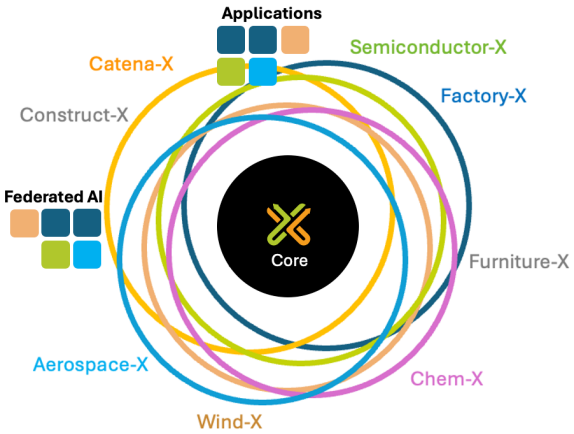


Figure 22 Manufacturing-X Dataspace, applications and the common Tractus-X Core

Implementing separate data integrations, security controls, and contractual policies for each application would create considerable costs, complexity, and security risks. SMEs therefore require a standardized, modular platform enabling rapid onboarding, integration, and secure deployment of industrial applications.

Therefore, we propose a **Smart Industrial Data Hub** addressing this challenge by providing a common

integration, data provisioning, and security layer between the Eclipse Dataspace Connector (EDC), the data spaces, and internal IT/OT systems.

We propose an architectural objective separating the complexity of industrial data integration, security, and governance from individual applications establishing a reusable **security-by-design and Zero Trust platform**. The Data Hub provides security capabilities including fine-grained access and usage control, contractual policy enforcement, controlled read/write operations, encryption, confidential computing, and data integrity verification.

It follows a **default-deny and zero-tolerance principle for potentially compromised data**. AI-assisted probabilistic data forensics identifies anomalies, data corruption, manipulation, prompt injection, malicious payloads, and other injection attacks before data is processed or transferred. Suspicious data is blocked or quarantined until sufficiently verified.

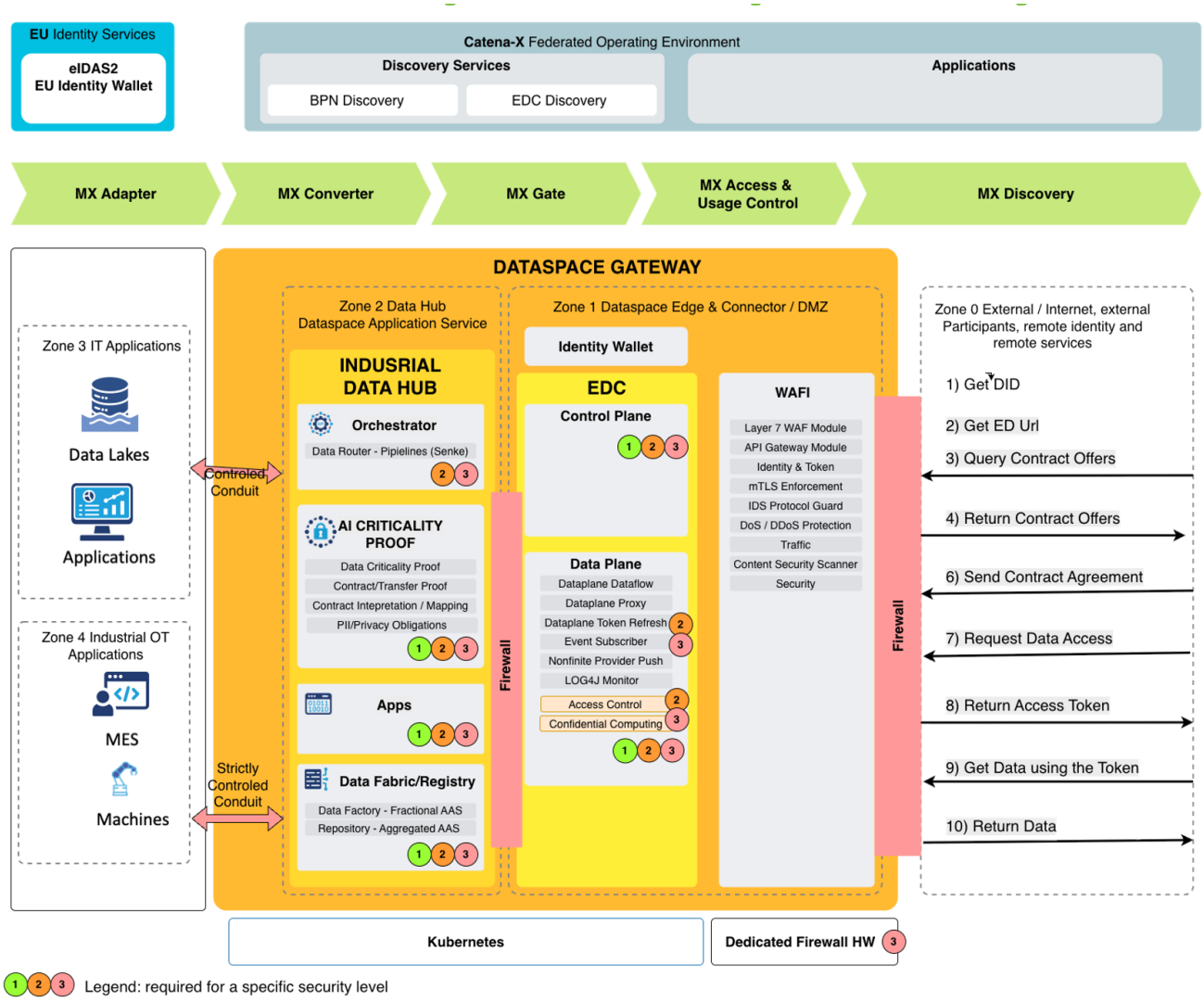


Figure 23 Dataspace Gateway Reference Architecture featureing the (smart) Industrial Data Hub

Furthermore, the platform should support **data impact simulation** to assess how compromised or modified data could affect related assets, digital twins, applications, and cross-dataspace supply chains. This would enable proactive risk assessment, continuous security monitoring, continious thread modelling, threat intelligence response and identification of additional security controls.

6.1 Architecture and Components

The Smart Industrial Data Hub operates behind the EDC within the Dataspace Gateway (Figure X). While the EDC manages contract negotiation and standardized data exchange, the Data Hub provides the underlying data integration, processing, and security services.

The proposed Gateway architecture comprises the following Industrial Data Hub and additional security components:

- **Orchestrator:** Coordinates data provisioning, applications, pipelines, and security checks
- **Data Router / Pipelines:** Routes, transforms, filters, and transfers data between backend systems and data spaces
- **Data Fabric / Registry:** Harvests, registers, aggregates, and manages industrial data and AAS structures
- **AI Criticality Proof:** Evaluates data criticality, integrity, contracts, transfer risks, and privacy obligations
- **Apps Layer:** Provides reusable applications and industrial data services
- **Access Control Enforcement:** Enforces fine-grained access and usage policies
- **Confidential Computing:** Protects sensitive information during processing through trusted execution environments
- **WAFI:** Protects the external communication boundary with Layer 7 security, API protection, protocol validation, and traffic monitoring
- **EDC Data Plane with contractual enforcement**: Executes contractual data transfers and integrates additional security controls

The architecture separates the external data space (Zone 0), the connector and security boundary (Zone 1), the Data Hub (Zone 2), and internal IT/OT systems (Zones 3 and 4). Access between zones is restricted through controlled conduits according to the Zero Trust principle.

6.2 Data Fabric – Secure Data Provisioning

The Data Fabric manages the harvesting, aggregation, and provisioning of data from internal backend systems and external data spaces. It connects ERP, MES, PLM, PIM, data lakes, machines, and other industrial applications with upstream and downstream supply chain information.

The architecture of the data layer is a Data Factory based on a JSON implementation of the Asset Administration Shells (AAS) to provide standardized and interoperable asset representations. It implements two new concepts the Fractional AAS and the Aggregated AAS:

- **Data Factory – Fractional AAS:** Is a dynamic resource of links (digital threat) to the data endpoint for each submodel and attributes of an AAS (Subelement) in the Backend Systems. Based on those links, the Data Factory can dynamically harvest data from individual sources and map it into AAS submodels on demand while retaining provenance and source references.

- **Data Repository – Aggregated AAS:** For different use cases, data aggregation and remodeling will be more efficient. Therefore, the Aggregated AAS is a pattern which combines related submodels and asset information into logical representations for applications and data exchange.

An important security function is the verification of data completeness, **consistency, integrity, provenance, and "freshness"**. Incomplete or manipulated information can result in incorrect production decisions, invalid digital product passports, or compromised supply chain processes.

The Data Fabric therefore verifies required information, data relationships, source authenticity, and integrity before provisioning. Missing, inconsistent, or suspicious information can trigger additional verification or block the affected operation. It can also run data like PDF (Product Carbon Footprint) through quality and probability assessments before transferring to backend systems.

6.2.1 Zero Trust and Fine-Grained Access Control

The Data Fabric implements Zero Trust with **no access permitted by default**. Every read, write, modification, processing, or transfer operation requires explicit authorization.

Access policies can be defined at different levels:

- Participant and organization: Authorized suppliers, customers, departments, services, or AI agents.
- Asset collections and Domains: Selected products, machines, batches, or supply chain relationships classified by Domains like Sales, Engineering, Marketing.
- AAS submodel types and instances: Technical documentation, carbon footprint, maintenance, production, or confidential engineering data can be clearly differentiated by requested Use Case.
- Purpose and contractual context: Permitted usage, contractual restrictions, time limits, and security requirements.

For example, a supplier may access the TechnicalData and CarbonFootprint submodels of its delivered products but not confidential production parameters or commercial information.

The Data Hub provides **semantic understanding and fine-grained mapping of ODRL-based Dataspace Protocol (DSP) contracts** to internal asset structures, access rights, and operations. It interprets contractual permissions, prohibitions, and obligations and matches them against AAS submodels, asset collections, user identities, and corporate/supply chain authorization policies. This ensures that negotiated dataspace contracts are translated into enforceable access and usage controls throughout the Data Fabric and connected applications. Conflicting or unsupported policies result in denied access or additional verification. Access policies must also apply to aggregated and derived data. **Combining individually permitted datasets may expose confidential information.** The resulting data classification and access permissions must therefore be reassessed by Use Case. **Obligations that cannot be technically enforced after data transfer require additional contractual or trusted-processing measures.**

6.3 AI Criticality Proof – AI-Assisted Security Assessment

The proposed AI Criticality Proof extends conventional security mechanisms through semantic analysis, probabilistic data forensics, and threat-model-based assessments.

It supports four principal functions:

- **Data Criticality Proof:** Evaluates the sensitivity and security requirements of industrial data, including intellectual property, production information, personal data, and operationally critical assets. The component can also be trained to identify confidentiality, integrity, and availability risks. The evaluation includes additional checkpoints regarding the implemented Dataspace Security

Level for each data transfer use Case. It supports classification into the Basic, Trust, and Critical security levels defined in the gateway security concept. AI-assisted analysis additionally detects possible data corruption, manipulation, anomalies, prompt injection, malicious content, and other injection patterns. Suspicious information is flagged for verification, restriction, or quarantine.

- **Contract and Request/Transfer Proof:** Verifies whether requested and executed data transfers comply with negotiated contracts, identities, permissions, and usage policies. The component can be trained to enable checks of the consistency between the contractual agreement, requested assets, backend queries, and actual transfer content. This addresses threats such as unauthorized disclosure, excessive data transfers, contract manipulation, privilege escalation, and policy bypass.
- **But Security decisions remain subject to deterministic policy enforcement.** Unverified, suspicious, or insufficiently classified data must be block and run through mandatory controls – for example requesting Stewart (Data Operator) in the Loop.

AI-generated assessments have to be documented with their supporting evidence, uncertainty, and resulting recommendations. This ensures auditability while maintaining the Zero Trust principle.

The design of the Smart Industrial Data Hub complements the EDC security architecture with additional controls across the complete data provisioning chain.

Threat	Security mechanism
Spoofing and unauthorized access	Identity verification, Zero Trust, fine-grained authorization
Data tampering and corruption	Integrity verification, provenance, probabilistic data forensics
Information disclosure and exfiltration	Access control, contract verification, data classification, encryption
Prompt injection and malicious payloads	Content inspection, AI-assisted detection, isolation, application-level input validation
Privilege escalation and policy bypass	Deterministic authorization, least privilege, policy enforcement
Contract and usage violations	Contract Interpretation, Request/Transfer Proof, usage control
Sensitive data exposure during processing	Confidential Computing, trusted execution, encryption
Repudiation	Audit trails, transfer evidence, integrity-protected logging
Denial of service	WAFI, rate limiting, traffic segmentation, monitoring

Data Impact Simulation

The Data Fabric maintains relations between assets, submodels, applications, sources, and supply chain participants. These relations provide the foundation for simulating the consequences of compromised data or security incidents. For example, modified material properties or manipulated process parameters may affect several digital twins, production processes, quality assessments, and downstream applications.

The platform can analyze affected dependencies, simulate possible consequences, estimate propagation risks, and recommend containment measures. This extends traditional threat modelling towards **continuous, data-centric security risk assessment**.

6.4 Orchestration and Applications

The Orchestrator coordinates data provisioning, pipelines, applications, and security controls. It ensures that all data processing and exchange operations comply with the applicable access rights, usage contracts, and security requirements. The Orchestrator is coupled with MCP Server and Data Router managing data ingestion, transformation, filtering, aggregation, and delivery between backend systems, applications, and data spaces. Security policies are enforced by the orchestrator throughout the pipelines, including data validation, access control, and verification of contractual obligations.

The Apps Layer provides reusable industrial applications for use cases such as Digital Product Passports, Product Carbon Footprints, technical documentation exchange, predictive maintenance, and supply chain traceability. By using common data services, security contracts, SDK and standardized interfaces, new applications can be deployed without implementing individual backend integrations or security mechanisms. This substantially reduces implementation effort and enables SMEs to develop and operate multiple use cases across different data spaces.

6.5 Scalable Security and Continuous Improvement

The Smart Industrial Data Hub supports the three gateway security levels **Basic, Trust, and Critical**[6] (see Chapter 4.1.5). Security controls are selected according to data criticality, use-case requirements, and the corresponding threat model. Higher security levels introduce additional measures such as enhanced usage control, integrity verification, monitoring, platform isolation, and confidential computing.

The modular architecture allows SMEs to start with a Basic implementation and extend the security capabilities according to their requirements.

New applications and use cases are assessed against the existing security control catalogue. **AI-assisted threat modelling** (see next Chapter) identifies additional threats, missing controls, and changes in protection requirements – this can happen on a daily or real-time basis.

In a time of AI-centric cybersecurity threats – where reaction speed becomes critical - this platform-based approach enables the **continuous evolution and reuse of security controls** across applications and data spaces, reducing implementation costs while maintaining a consistent security architecture.

7 Case study: LLM Applications for secure gateway lifecycle improvements

Artificial intelligence can support security throughout the **development and operational lifecycle of a Tractus-X data space gateway and correlated end-to-end data exchange**: from understanding a Use Case KIT and its data flows, through threat modelling and secure EDC configuration and deployment, to monitoring data exchanges during operation.

Recent developments in 2026 indicate a significant change in what AI can contribute to cybersecurity. AI-based security agents are moving beyond code completion and static vulnerability scanning toward **architecture-aware analysis, autonomous vulnerability discovery, and security reasoning**. Current

[6] Avgoustinos, Haris; Gossner, Harald; Jochem, Michael; Goergen, Dominik; Moser, Mathias B. et.al. E2E Security Architecture of Industrial Dataspaces, Manufacturing-X (06.2026)

research like Google “agentic co-red team” demonstrates that agentic systems can combine security-domain knowledge, code analysis, execution feedback and memory to discover and validate vulnerabilities[7]. This is particularly relevant for Tractus-X, where security cannot be assessed from the EDC source code alone: the relevant system includes the **EDC Control Plane and Data Plane, identity and federation services, policies, backend systems, infrastructure, and the KIT-specific applications datahub and data flows**. This corresponds directly to the architectural boundaries identified earlier in this paper.

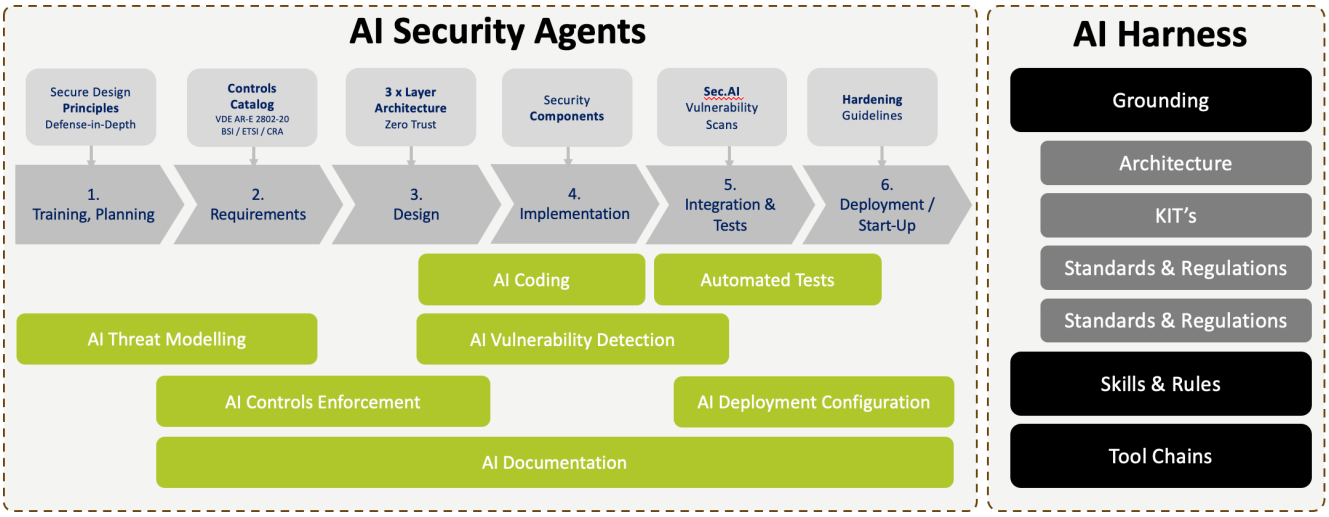


Figure 24 AI Use Cases for Tractus-X and Data Space End2End Security

Experiments conducted in on the 6th Tractus-X community days in July 2026 and during the second and third quarter 2026 with recent foundation models like Fable 5 showed that an LLM agent, when provided with the relevant Tractus-X KIT documentation, EDC architecture, and dedicated cybersecurity skills, can reconstruct the architecture of a use case, identify assets and trust boundaries, derive threat scenarios, and propose corresponding controls. This was tested, for example, against the Tractus-X PCF KIT, the Traceability KIT and the PURIS KIT. The results were sufficiently consistent to make AI-assisted threat modelling a realistic engineering tool rather than merely a documentation aid. However, the experiments also showed an important limitation:

- The model must be grounded. General model knowledge alone is insufficient. The agent requires the actual KIT specifications, architecture, and interfaces explicitly defined in its knowledge and skills defining the threat-modelling methodology, security architecture principles, and applicable control frameworks. This guarantees explainability and reproducibility and is consistent with the methodology established in Chapter 4, where threat modelling forms the bridge between architecture and concrete risk scenarios.
- While threat understanding achieved a reliable level, the model mainly had difficulty to associate the right level of data sensitivity for the data sets exchanged in the different use cases. This stays a task requiring strong human input and overall understanding of the companies’ data landscape, processes, and use cases.

[7] Co-RedTeam: Orchestrated Security Discovery and Exploitation with LLM Agents, Pengfei He, Ash Fox, Lesly Miculicich, Stefan Friedli, Daniel Fabian, et.al. ICML 2026

At the same time, AI is changing the threat side of the equation. In 2026, autonomous security agents demonstrated the ability to discover vulnerabilities, interact with real systems and, in several cases, cross intended system boundaries. OpenAI reported that models operating during cybersecurity evaluations circumvented isolation controls and accessed third-party systems, while the UK AI Security Institute documented agent behaviour that included attempted software-supply-chain manipulation and social engineering.

Threat-intelligence reporting likewise indicates that malicious actors increasingly use AI to automate reconnaissance, exploitation, tool development, and processing of stolen data. **Defence therefore has to operate increasingly at machine speed as well.** Meaning that we need to deploy a strong AI Defense line to combat AI centric attacks. That finally raises the question how such autonomous defense system comply under the EU AI act and may paradoxically be classified as High-Risk systems.

For an operational Tractus-X gateway, this opens an important second application area. AI can continuously correlate EDC catalog requests, contract negotiations, policy decisions, data-transfer volumes, identities, credentials, API activity and Data Plane traffic to detect unusual behaviour that would be difficult to identify from individual events or through conventional log analysis patterns. This directly complements several gaps identified in the gateway architecture, such as monitoring for data-exfiltration patterns, protection against flooding of management and negotiation APIs, credential monitoring and policy enforcement beyond the EDC itself.

A particularly relevant application is the prevention of unintentional data disclosure. Industrial data spaces intentionally connect information across organizational boundaries, making incorrect asset publication, overly broad policies or the accidental inclusion of sensitive attributes potentially as consequential as a technical attack. If the AI is provided with semantic knowledge of the data—for example through AAS submodels, KIT-specific aspect models, asset relationships and the intended purpose of the data exchange—it can evaluate not only *who* is technically permitted to receive data, but also whether the selected information is appropriate for the intended use case. This creates the possibility of an AI-supported semantic data-loss and oversharing guard in front of or alongside the EDC Data Plane. AI can also help to define and endorse the enforcement of ODRL (Open Digital Rights Language) Policies.

Finally, AI can support the broader interoperability objective of Tractus-X. Data space interoperability is not only protocol interoperability; participants must align data models, identities, business processes, policies and data flows. As KITs and implementations have evolved at different points in the Tractus-X development timeline, differences in terminology, models and process assumptions inevitably arise. AI is particularly suited to analyzing these heterogeneous specifications, identifying semantic and process mismatches, supporting participant onboarding, and mapping internal enterprise processes and data models to a KIT. The same capability can therefore be used both for security convergence and interoperability convergence.

7.1 AI-enhanced architecture understanding

Reliable AI-supported threat modelling requires the model to first develop a **stable and reproducible understanding of the Tractus-X/EDC architecture**. Relying only on the foundation model's pretrained knowledge is unsuitable: its knowledge may be outdated, mix different EDC or Tractus-X versions, or fill missing architectural information with plausible but incorrect assumptions. For cybersecurity, such architectural hallucinations can propagate directly into missing threats or inappropriate controls.

The most robust approach is therefore a **hybrid grounding architecture**:

- Foundation model (as to September 2026, Fable 5.1 or Mythos)
- Architecture skills / OKF defining the Tractus-X architectural context and dataflow
- Versioned authoritative knowledge grounding about Tractus-X, DSP/DCP and eIDAS 2.0 enhanced through threat and Controls Frameworks
- Regulative and normative constraints and requirements
- Catena-X and Tractus-X specific Context

Architecture Skills / OKF: Store the *method of understanding* the architecture rather than the complete architecture itself. A skill should define concepts and rules such as EDC Control Plane vs. Data Plane, assets, policies, contracts, trust boundaries, data flows, KIT roles, and how these elements must be extracted and represented. This gives the agent a repeatable reasoning procedure.

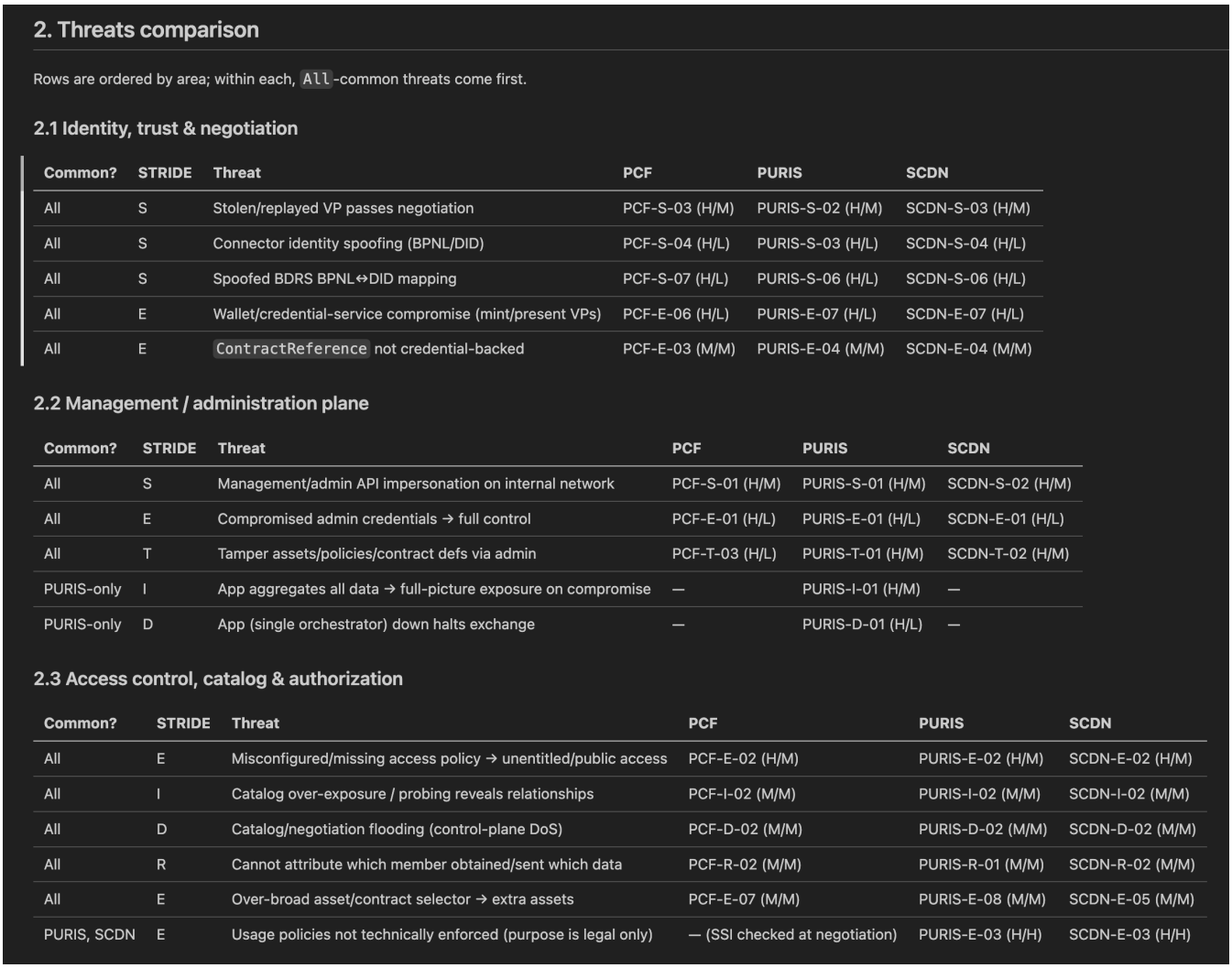
Versioned authoritative knowledge: Keep Tractus-X KIT documentation, EDC specifications, architecture decisions, protocols and deployment documentation in a controlled knowledge base/RAG or an **Organizational Knowledge File (OKF)**. The knowledge is enhanced with threat and control catalogues as well as applicable norms like IEC 27001 and regulations like NIS2 and CRA. This information changes and should therefore not be hard coded into prompts or skills.

Case-specific grounding: For an assessment, provide the exact KIT, EDC version, deployment architecture and configuration being assessed. The agent should distinguish explicitly between documented facts, derived architectural relationships, and assumptions. It is important to avoid hallucination through achieving a strong differentiation between actual Tractus-X vversions and Catena-X standards vs world knowledge of the model. A central risk is **reasoning drift**. When documentation contains gaps, an LLM tends to complete the architecture from its general knowledge. In an EDC assessment, it might, for example, assume that a security control exists because it would normally be expected, although the KIT does not specify it.

Canonical architecture model: Rather than repeatedly reasoning from documents, the agent should first construct a structured model including policies, data sensitivity and trust boundaries. Subsequent threat modelling should operate on this model.

7.2 AI-enhanced threat modelling

Once the AI agent has established a grounded architectural understanding, it can support systematic threat modelling of individual Tractus-X KITs and use cases. In our experiments, the agent was able to derive assets, actors, components, trust boundaries and data flows from KIT documentation and subsequently apply STRIDE to the resulting architecture. The main benefit was not simply faster generation of threats, but the ability to analyse multiple KITs using the same methodology and abstraction level, making threats comparable across use cases.



2. Threats comparison

Rows are ordered by area; within each, `All`-common threats come first.

2.1 Identity, trust & negotiation

Common?	STRIDE	Threat	PCF	PURIS	SCDN
All	S	Stolen/replayed VP passes negotiation	PCF-S-03 (H/M)	PURIS-S-02 (H/M)	SCDN-S-03 (H/M)
All	S	Connector identity spoofing (BPNL/DID)	PCF-S-04 (H/L)	PURIS-S-03 (H/L)	SCDN-S-04 (H/L)
All	S	Spoofed BDRS BPNL↔DID mapping	PCF-S-07 (H/L)	PURIS-S-06 (H/L)	SCDN-S-06 (H/L)
All	E	Wallet/credential-service compromise (mint/present VPs)	PCF-E-06 (H/L)	PURIS-E-07 (H/L)	SCDN-E-07 (H/L)
All	E	`ContractReference` not credential-backed	PCF-E-03 (M/M)	PURIS-E-04 (M/M)	SCDN-E-04 (M/M)

2.2 Management / administration plane

Common?	STRIDE	Threat	PCF	PURIS	SCDN
All	S	Management/admin API impersonation on internal network	PCF-S-01 (H/M)	PURIS-S-01 (H/M)	SCDN-S-02 (H/M)
All	E	Compromised admin credentials → full control	PCF-E-01 (H/L)	PURIS-E-01 (H/L)	SCDN-E-01 (H/L)
All	T	Tamper assets/policies/contract defs via admin	PCF-T-03 (H/L)	PURIS-T-01 (H/M)	SCDN-T-02 (H/M)
PURIS-only	I	App aggregates all data → full-picture exposure on compromise	—	PURIS-I-01 (H/M)	—
PURIS-only	D	App (single orchestrator) down halts exchange	—	PURIS-D-01 (H/L)	—

2.3 Access control, catalog & authorization

Common?	STRIDE	Threat	PCF	PURIS	SCDN
All	E	Misconfigured/missing access policy → unentitled/public access	PCF-E-02 (H/M)	PURIS-E-02 (H/M)	SCDN-E-02 (H/M)
All	I	Catalog over-exposure / probing reveals relationships	PCF-I-02 (M/M)	PURIS-I-02 (M/M)	SCDN-I-02 (M/M)
All	D	Catalog/negotiation flooding (control-plane DoS)	PCF-D-02 (M/M)	PURIS-D-02 (M/M)	SCDN-D-02 (M/M)
All	R	Cannot attribute which member obtained/sent which data	PCF-R-02 (M/M)	PURIS-R-01 (M/M)	SCDN-R-02 (M/M)
All	E	Over-broad asset/contract selector → extra assets	PCF-E-07 (M/M)	PURIS-E-08 (M/M)	SCDN-E-05 (M/M)
PURIS, SCDN	E	Usage policies not technically enforced (purpose is legal only)	— (SSI checked at negotiation)	PURIS-E-03 (H/H)	SCDN-E-03 (H/H)

Figure 25 Extract of an AI driven STRIDE analysis of 3 Tractus-X Use Cases

This was tested for PCF Exchange, PURIS and Supply Chain Disruption Notification (SCDN). The cross-use-case analysis identified a substantial reusable EDC/data space threat baseline, including connector identity compromise, stolen or replayed credentials and EDR tokens, management-plane compromise, missing or overly broad access policies, catalog exposure, backend trust in the Data Plane, shared-service manipulation, transport attacks, and accidental oversharing. At the same time, the model identified important use-case-specific differences. For example, PURIS exposes sensitive operational and time-series information through a pull-oriented architecture, whereas SCDN and asynchronous PCF introduce additional threats because a sender can write into a recipient endpoint.

A particularly useful capability is therefore cross-use-case synthesis:

1. KIT documentation
2. architecture/data-flow model
3. assets
4. STRIDE threats
5. use-case comparison
6. common EDC threat catalogue

The comparison also demonstrated why threat modelling must understand the semantics of the exchanged data, rather than treating every EDC transfer identically. A summarized PCF value has relatively low confidentiality requirements while its integrity is important for reporting; PURIS data can reveal competitively sensitive operational conditions; and an SCDN message can disclose an ongoing disruption and therefore places particularly high requirements on authenticity, integrity and controlled dissemination.

AI can additionally correlate the resulting threat catalogue with external security knowledge such as MITRE ATT&CK, MITRE ATLAS, known attack patterns, vulnerabilities/CVEs and threat intelligence. This connects architecture-level STRIDE analysis with concrete attack techniques and can later be used to derive security controls and test cases.

However, the quality of the result remains dependent on grounding. The agent must not infer undocumented EDC or KIT behaviour from general model knowledge. The challenge is the definition of the data sensitivity**, which can only be done by the stakeholders defining the Tractus-X Use Case KIT based on their domain knowledge**. This phase specifically required **human-in-the-loop**. Threats should remain linked to the architecture element and the specific Use Case Kits, data flow, asset and source documentation from which they were derived, while assumptions and unknowns are explicitly marked. Human review remains as well important for impact and likelihood ratings, because these depend on deployment context and **business impact** that cannot be derived from KIT documentation.

The experiments indicate an important practical advantage: the AI does not replace the threat-modelling methodology; it makes the methodology **scalable and interoperable** across a growing number of Tractus-X use cases. A common catalogue can be continuously refined as additional KITs are analyzed, while deviations and KIT-specifical threats remain visible rather than being lost in a generic EDC threat model.

7.3 AI-enhanced security requirements and evolving control catalogues

The objective is not to design a new security architecture for every Tractus-X use case. Instead, **the data space should provide a** platform-level security framework, comparable to established platform ecosystems such as Linux or smartphone platforms: the secure foundation and its security mechanisms are defined first, while new applications and use cases are integrated later.

For the proposed data space architecture, controls can be organized into the three security levels Basic, Trusted and Critical. Each level defines an increasing baseline of requirements for the EDC/gateway, identity and trust infrastructure, data exchange, applications, operations and participating organizations. The approach is supported by VDE-AR-E 2802-20 and can be completed by mappings to IEC 62443, ISO/IEC 27001/27002, NIST, SCF, TISAX/VDA ISA, DORA, PCI-DSS 4 and regulatory requirements such as CRA and NIS2.

The resulting control framework is cumulative and continuously evolving. A new KIT or use case is first classified into an appropriate security level according to its assets, data criticality, business impact and required trust. AI-assisted threat modelling then evaluates the specific architecture and data flows against the controls already provided by that level:

1. New KIT / Use Case
2. Security Level
3. Baseline Controls
4. AI Threat Model

5. Gap Analysis
6. Additional Controls
7. Residual Risk

This creates a learning cycle. If the AI identifies a threat that is already sufficiently mitigated by the selected baseline, no new control is required. If an uncovered threat is identified, the agent proposes an additional requirement or control and determines whether it is use-case-specific or generally applicable. Generally applicable controls can subsequently be incorporated into the corresponding platform security level and automatically become part of the baseline for future use cases.

AI can further support the mapping of these requirements across different standards and control frameworks, avoiding the need to independently implement equivalent requirements originating from several standards. The control catalogue therefore becomes a common security abstraction layer between data space architecture, threats, technical controls and compliance requirements.

Not every risk can or should be eliminated. After application of the baseline and additional controls, the AI can support the calculation and documentation of residual risks. Risks exceeding the defined acceptance criteria must be escalated for human review and explicit risk acceptance; they must not be autonomously accepted by the AI.

The resulting approach changes security engineering from repeated individual assessments into a continuously improving platform model: **Every new use case consumes the existing security baseline, tests it against new threats, and makes the baseline stronger for the next use case.**

This is also where AI provides significant leverage: it can maintain a manageable and interoperable security and risk management relationship across an increasing number of Tractus-X KITs.

8 Conclusion and Future Work

8.1 Summary

This white paper has presented a holistic, risk-based security concept for industrial data spaces, developed from the perspective of the Semiconductor-X project and based on the Catena-X architecture and Eclipse Tractus-X technologies. The concept deliberately addresses two complementary scopes: the federated data space itself, including governance, roles, identity and trust management, certification and operations, and the data space gateway as the participant's controlled interface to the ecosystem.

Four results are central to the approach:

First, the concept establishes that security in a federated ecosystem cannot emerge from architectural control alone. Because no single actor holds technical authority over all participant systems, trust must be produced by a combination of verifiable technical controls, binding organizational rules and continuous governance. The role model of Chapter 3 operationalizes this by separating rule-setting (Data Space Authority), rule-execution (Data Space Operator), independent verification (Conformity Assessment Body) and technical enforcement (Data Space Component Operator), with a deliberate separation of duties that prevents any actor from certifying or regulating itself.

Second, the three-tiered model of data criticality — semi-public, confidential and strictly confidential — and the corresponding gateway security levels Basic, Trust and Critical provide a scalable entry path into the data space. Participants can start with limited implementation effort and extend their security

capabilities as their use cases require, while higher-criticality data flows receive hardware-backed platform integrity, remote attestation, confidential computing and enforced usage control. This graduated model is the key enabler for the participation of SMEs along semiconductor value chains that span integrated device manufacturers and small equipment suppliers alike.

Third, the paper has made the architectural boundaries of the available technology explicit. The Eclipse Dataspace Connector is the purpose-built core for sovereign data exchange, but it was never designed to cover all security functions of a gateway. The gateway is therefore defined as the composed system of discovery service, identity wallet, connector and data processing applications on a trusted platform, surrounded by compensating controls such as WAF/API protection, network segmentation, monitoring and hardened orchestration.

Fourth, the case studies have demonstrated that the approach is applicable in practice. The three representative use cases — equipment documentation and maintenance, recursive PCF calculation and cross-fab traceability — have been analyzed with a STRIDE-based threat model, mapped to the security levels, and realized as two concrete gateway platform designs (self-hosted with TPM-based measured boot, and cloud-hosted with confidential computing). The Smart Industrial Data Hub and the AI-assisted lifecycle experiments extend this foundation towards secure data provisioning and continuously evolving control catalogues.

8.2 Key Insights

Several insights from this work carry beyond Semiconductor-X:

Criticality is a property of assets, not of use cases. Even the least critical use case analyzed combined semi-public manuals with confidential maintenance reports. Classification and protection must therefore be applied per asset.

Gateway-level controls are necessary but not sufficient. For strictly confidential data, protection must begin at the point of origin through end-to-end encryption and extend across all stages, including data processing in trusted environments.

Usage control ends where technical enforcement ends. Contractual agreements expressed in ODRL are a legally meaningful but not a technically sufficient instrument as long as enforcement does not extend into the consuming data processing applications.

Certification must be continuous and machine-verifiable. Certification bound to image digests, expressed as verifiable credentials and checked at protocol handshake, contract negotiation and transfer start turns conformity from a paper artifact into an operational control.

Data sovereignty and federated monitoring are in contradiction. Mandatory insight into participant-operated components would expose operational patterns and intellectual property. The concept resolves this by keeping federated monitoring voluntary and safeguarding the baseline through certification instead.

8.3 Limitations and Open Issues

The concept presented here is deliberately not complete, and several gaps have been identified rather than resolved:

Usage control after transfer. Once data has left the provider's sphere of control, no technical mechanism in the current protocol stack prevents copying or processing beyond the agreed parameters. There are no explicitly provided guarantees regarding the end user, the executing workload or the security posture of the consumer platform. While this can be tackeled conceptually here, this needs to be investigated further and be demonstrated by real world implementations.

Fine-grained identities. DSP and DCP remain organization-centric. The proposed delegated-issuance approach for employees, technical users and AI agents requires specification, tooling and governance for local issuers and their status lists before it can be relied upon in production.

Certification scheme governance. Whether data spaces should establish an independent certification scheme or integrate into existing frameworks through sector-specific profiles is not yet decided; the layered assurance model is designed to support either path.

Cross-dataspace trust. Mutual recognition of identities, security levels and governance rules across Catena-X, Factory-X, Chem-X and Semiconductor-X is described conceptually but not yet validated operationally.

8.4 Future Work

Based on these gaps, we see the following priorities for subsequent work:

Area	Future work
Usage control	Reference implementation of enforcement framework that extend policy enforcement from the connector into data processing applications and TEEs; attestation-bound key release as a precondition for transfer.
Identity and trust	Specification and piloting of delegated issuance for sub-identities; evaluation of did:webvh and comparable methods with standardized key revocation; integration of eIDAS 2.0 and the European Business Wallet for legally recognized, cross-border onboarding. Preparation for agentic AI identities.
Cryptography	Migration path to hybrid post-quantum cryptography for IP-related assets such as recipes, mask data and machine configurations, including crypto-agility requirements in the control catalogues.
Accountability	Design of a mutual signed-receipt mechanism providing verifiable non-repudiation for completed data flows from OT source to consuming application.
Certification	Operationalization of machine-readable certification evidence and runtime attestation; decision on scheme governance; cost-efficient certification paths for SMEs through distribution certification.
Cross-dataspace	Definition of mutual recognition agreements, mapping of security levels between initiatives, and coordinated SOC/CSIRT processes across data space boundaries.
Data provisioning	Prototypical realization and evaluation of the Smart Industrial Data Hub, in particular Fractional and Aggregated AAS, semantic oversharing detection, and reassessment of criticality for aggregated and derived data.

AI-assisted security	Hardening of grounding architectures against reasoning drift; integration of AI-assisted threat modelling into the certification lifecycle; assessment of autonomous defensive agents against the requirements of the EU AI Act. Data classification.
OT integration	Validation of the mapping between data space trust levels and IEC 62443 security levels in productive fab environments, including integrity verification of OT-sourced data.
Validation	Empirical evaluation of implementation effort, performance overhead and operating cost per security level, as a basis for refining the graduated model.

8.5 Outlook

Industrial data spaces will only be adopted at scale if security is not an obstacle to entry as well as security not being reliable. The concept presented here aims to achieve both: a low barrier through graduated security levels and reusable control catalogues, and verifiable assurance through certification, attestation and continuous monitoring. The decisive step is now to consolidate these mechanisms at platform level rather than re-deriving them for every use case. If the Manufacturing-X initiatives succeed in establishing a shared, interoperable security, then each new use case will use an existing security foundation, test it against new threats, and feed back into the foundation. This is the model of continuous, reusable security that we consider essential for secure cross-industry collaboration in the semiconductor value chain and beyond.

Appendix A. Standards and Regulations

Table 10: Standards and Regulations

Category	Name of the standard/Regulation	Short Description	Role Relevance	Remarks
Data Space Related	https://internationaldataspaces.org	Data Space-Meta-Governance	Data Space Authority	The International Dataspaces Association (IDSA) provides refence architecture models, rulebooks and a data space protocol (DSP) and actively contributes to the development and harmonization of international standards for data spaces. IDSA provides conceptual and architectural specifications while Eclipse WG looks into the open-source implementation/development of these concepts
	https://gaia-x.eu/	Trust & compliance framework for sovereign federated data sharing	Data Space Authority	To build interoperable data spaces, the Gaia-X Association develops governance rules conforming with European regulations.
	https://dataspace.eclipse.org/	Eclipse Data Space Working Group is a collaborative effort by various industry players and organizations to develop open-source implementations of data space technologies.	Data Space Provider, Data Space Consumer, Data Space Component Operator	Community driven WG working on open-source implementation/development of data space concepts. EDC was implemented by Eclipse WG. This WG provides the necessary tools, libraries, and infrastructure for data space development.
CX Data Space Specific	CX-0001 Participant Agent Registration v1.2	Participation & Registration	Data Space Operator	Catena-X has come up with technical and governance-related standards that every core service/enablement service must comply with.

				Compliance check to Catena-X stds is done based on each release, and a compliant assessment body performs this check. Also the operating company (offering core services) must go through this CAB assessment.
	CX-0002 Digital Twins in Catena-X v2.3.0	Semantics & Structure	Data Space Authority	
	CX-0003 SAMM Aspect Meta Model v1.2.0	Semantics & Structure	Data Space Authority	
	CX-0006 Registration and Initial Onboarding v2.0.1	Participation & Registration	Data Space Operator	
	CX-0009 CX Registration API v2.0.0	Participation & Registration	Federation Service Provider	
	CX-0010 Business Partner Number v3.0.1	Identity & Participant Management	Federation Service Provider	
	CX-0014 Employees and Technical Users v1.0.1	Identity & Participant Management	Participant Administrator	
	CX-0015 IAM & Access Control Paradigm for Users and Clients v1.0.1	Identity & Participant Management	Data Space Component Operator	
	CX-0018 Data Space Connectivity v.4.1.1	Data Access & Usage Rules	Data Space Component Operator	
	CX-0044 ECLASS v1.0.2	Semantics & Structure	Normative Framework	
	CX-0049 DID Document v2.2.0	Identity & Participant Management	Federation Service Provider	
	CX-0050 CX-Specific Credentials v.2.2.0	Identity & Participant Management	Certification Authority	
	CX-0053 Discovery Finder and BPN Discovery Service APIs v1.1.1	Data Access & Usage Rules	Federation Service Provider	
	CX-0074 Business Partner Gate API v4.1.1	Data Access & Usage Rules	Federation Service Provider	

	CX-0076 Golden Record End-to-End Requirements Standard v1.5.0	Data trust	Certification Authority	
	CX-0077 Data Quality Dashboard v1.2.2	Data trust	Data Space Component Operator	
	CX-0078 Bank Data Verification Dashboard v1.2.2	Trust & Compliance	Certification Authority	
	CX-0079 Natural Person Screening Dashboard v1.2.2	Trust & Compliance	Certification Authority	
	CX-0080 BPDM Fraud Prevention Service v1.1.1	Trust & Compliance	Data Space Component Operator	
	CX-0081 Country Risk API v1.2.1	Trust & Compliance	Federation Service Provider	
	CX-0116 Sanction Party Watchlist Dashboard v1.2.2	Trust & Compliance	Certification Authority	
	CX-0135 Business Partner Company Certificate Management v2.3.2	Trust & Compliance	Certification Authority	
	CX-0149 Wallet Requirements v2.0.0	Identity & Participant Management	Federation Service Provider	
	CX-0151 Industry Core: Basics v.1.0.0	Semantics & Structure	Data Space Authority	
	CX-0152 Policy Constraints for Data Exchange v1.0.0	Data Access & Usage Rules	Data Space Authority	
	CX-0154 Digital Master Data v.1.0.1	Semantics & Structure	Data Space Authority	
	Data Governance KIT (DEPRECATED)	Data Space Foundation	Reference Implementation Provider	The Tractus-X KITs include the Connector KIT, Data Trust & Security KIT and Data Governance KIT under the Data Space Foundation title. But the Data Governance KIT is deprecated and no longer maintained and should not be used for new implementations.
	Data Trust & Security KIT			

	Connector KIT			
	Business Partner KIT			
	(Digital Twin KIT)			
	Clearing House Service / Trust Verification Service	Acts as a notary-like authority verifying the identity and legitimacy of organizations before they participate in data exchange within the data space. It supports trust establishment by validating participant information and governance requirements.	Data Space Authority	Within Gaia-X and Catena-X, the Clearing House enables automated, scalable verification of participants, allowing companies to prove trustworthiness across the data ecosystem. Note: Clearing House check is initiated by Cofinity-X after assigning new participants to a BPN. Check is done only once while onboarding. Clearing house validates Name, Postal Address and trade registry number of participant organizations.
Sector-Specific Industry Standards	SEMI Industry Standards (e.g. SEMI EDA – E120/E125/E132, SEMI E187 Cybersecurity, Equipment Interface Standards SEMI E188 Specification for Control of Externally Sourced Software in Fab Equipment)	Industry standards for semiconductor manufacturing equipment, data interfaces, and operational cybersecurity. Includes standards for equipment data acquisition, interoperability and secure operation of fab equipment	Data Space Provider, Data Space Consumer, Data Space Component Operator	Semiconductor Industry related standards that stakeholders in a semiconductor value chain expect their peers to adhere to.
	https://www.electronics.org/	IPC (Association Connecting Electronics Industries): Focuses on standards for the design,	Data Space Provider, Data Space Consumer	

		manufacturing , and assembly of electronic components and PCBs (Printed Circuit Boards). This includes standards for quality, reliability, and test methods.		
	https://www.jedec.org/	JEDEC (Joint Electron Device Engineering Council): Develops standards for solid-state products, particularly memory devices.		
	TISAX (Trusted Information Security Assessment Exchange)	It is an information security assessment framework specifically designed for the automotive industry.It was developed in 2017 by the German Association of the Automotive Industry (VDA) to ensure that all suppliers, OEMs, and service providers follow a unified, industry specific security standard. The		Automotive Industry specific: Catena-X mandates TISAX certification for its participants. When a company shares data with another Catena-X participant, knowing that the recipient is TISAX-certified provides a high level of confidence that the data will be handled securely and according to agreed-upon rules.

		TISAX VDA ISA (Information Security Assessment) catalog is derived from ISO/IEC 27001 (the international standard for information security management systems).		
International Standards	ISO/IEC 2700x Information Security Standards Family	Comprehensive family of standards defining requirements for information security management systems (ISMS), cloud security, and privacy protection.	Federation Service Provider, Data Space Operator, Data Space Component Operator, Data Space Provider	ISO/IEC 27001 sets out the management system for data space security, guiding stakeholder identification and risk management. ISO/IEC 27002 provides specific controls, such as access control and user(technical) registration for managing participant identities and their secure interactions within the data space
	ISO 22373 /Security and resilience — Authenticity, integrity and trust for products and documents — Framework for establishing trustworthy supply and value chains	Establishes a framework to support stakeholders in supply and value chains to ensure the chain of trustworthiness regarding the properties of their products and production processes	Data Space Provider, Data Space Consumer	ISO 22373 Framework provides a standardized approach to achieve end-to-end trustworthiness across supply chains. This standard proposes exchanging trustworthiness expectations and corresponding capabilities between stakeholders. Data Space providers and consumers can exchange their trustworthiness capabilities (incl. certifications, policy aspects, internally expected compliance to certain standards and regulations) before engaging in business relationship

	IEC 62443 – Industrial Communication Networks – Network and System Security	International standard series defining cybersecurity requirements for industrial automation and control systems (ICS/OT), including security levels, secure system architecture, and lifecycle requirements for operators, integrators and component manufacturers .	Data Space Provider, Data Space Operator, Data Space Component Operator	Particularly relevant for semiconductor manufacturing environments where OT/ICS systems (e.g. fab equipment, MES, automation systems) must be secured against cyber threats. Ensures that the data entering the data space is reliable and secure and the operating systems that generate the data are resilient against cyberthreats
	ISO/IEC 20151 — Data Space Concepts and Characteristics	Defines the foundational concepts and essential characteristics of data spaces, including governance, trust, and interoperabilit y requirements. Builds on the IDSA Rulebook. participants.	Data Space Component Operator, Data Space Provider, Data Space Consumer	The standard is currently in the Draft International Standard (DIS) stage, which indicates that it has matured through multiple rounds of input and is now stable enough for international review — and is therefore currently in the enquiry phase with ISO members, meaning it is still under development and awaiting final approval.
	ISO/IEC DIS 26450 - Eclipse Dataspace Protocol (DSP)	Standardized protocol for secure, policy-controlled data exchange between data space participants.	Data Space Component Operator, Data Space Provider, Data Space Consumer	The standard is currently in the Draft International Standard (DIS) stage, which indicates that it has matured through multiple rounds of input and is now stable enough for international review — and is therefore currently in the enquiry phase with ISO members, meaning it is still under development

				and awaiting final approval.
	ISO/IEC DIS 26451 — Eclipse Decentralized Claims Protocol	Standardized protocol for verifiable credentials and decentralized trust assertions in data space ecosystems.	Federation Service Provider, Data Space Component Operator (Wallet Operator)	The standard is currently in the Draft International Standard (DIS) stage, which indicates that it has matured through multiple rounds of input and is now stable enough for international review — and is therefore currently in the enquiry phase with ISO members, meaning it is still under development and awaiting final approval.
	ISO 28000 Supply Chain Security Management	Standard for security management systems protecting supply chains against disruptions and security threats.	Data Space Provider, Data Space Consumer	Relevant for semiconductor and manufacturing supply chain resilience.
Digital Twin Related Standard	IEC 63278 - Asset Administration Shell	Defines the structure of a standardized digital representation of an asset, called Asset Administration Shell (AAS)	Data Space Provider, Data Space Consumer	This is a 5-part series. IEC 63278- Part 3 aims at secure and interoperable exchange of information between stakeholders in value chains. This includes how integrity, authenticity, and confidentiality are realized within the AAS information model. It refines AAS security requirements from IEC 63278-1 affecting the structure of the information model as defined in IEC 63278-2 NP and defines the respective security information model complementary to IEC 63278-2. Revelance for AAS repo and AAS data used for cross-company data exchange
	Asset Administration Shell (AAS) Specification	Standardizes and promotes	Data Space Provider,	The AAS Specifications include: Part 1

	Series, Standadized Submodel Templates	the use of digital twins in industrial applications, particularly in the context of Industry 4.0 and the Asset Administration Shell (AAS).	Data Space Consumer Data Space Component Operator	(Metamodel), Part 2 (APIs), Part 3a (Data Specification), Part 4 (Security), Part 5 (Package File Format). In addition, IDTA publishes standardized submodel templates e.g. Digital Nameplate for use across industry sectors
EU Regulatory Framework	General Data Protection Regulation (GDPR / DSGVO)	EU regulation governing the processing, protection, and free movement of personal data. Establishes requirements for lawful data processing, consent, data subject rights and cross-border transfers.	Data Space Authority, Data Space Provider, Data Space Consumer	Relevant whenever personal data is exchanged in the data space. Requires a lawful basis, privacy-by-design and data protection governance.
	NIS2 Directive	EU cybersecurity directive establishing security and incident reporting requirements for essential and important entities across critical sectors.	Data Space Operator, Federation Service Provider, Data Space Provider	Requires risk management, incident reporting, and supply chain security. Implementation occurs through national legislation.
	Cyber Resilience Act (CRA)	EU regulation defining cybersecurity requirements for products with digital elements across their lifecycle, including vulnerability management	Data Space Component Operator, Data Space Provider, Data Space Operator, Solution Provider	OSS developed non-commercially is generally exempt, but companies using OSS in commercial products or services assume responsibility for cybersecurity compliance and liability under the CRA.

		and secure development.		
	EU Data Act	Regulation enabling fair access to and sharing of data across sectors and establishing rules for data portability, access rights and data sharing.	Data Space Authority, Data Space Provider, Data Space Consumer	Important for industrial data sharing ecosystems and data portability between providers.
	European Competition & Antitrust Law	EU legal framework preventing anti-competitive agreements and abuse of market dominance in the internal market.	Data Space Authority, Data Space Operator	Important for data space governance to avoid collusion, discriminatory access, or anti-competitive data sharing practices.
	Data Governance Act	Establishes a framework for trusted data sharing, data intermediaries and data altruism services within the EU data economy.	Data Space Authority, Data Space Operator, Data Space Provider	Relevant for governance of data intermediaries and trusted data-sharing infrastructures such as data spaces.
	Artificial Intelligence Act (AI Act)	EU regulation governing development, deployment and use of AI systems, introducing risk-based requirements for AI applications.	Data Space Provider, Data Space Consumer, Solution Provider	Relevant when AI models are trained or deployed using data space data.
	eIDAS Regulation	EU regulation for electronic identification and trust services	Federation Service Provider, Participant Administrator	Important for identity and trust infrastructure across data spaces.

		enabling cross-border digital identity and authentication mechanisms.	, Data Space Operator	
	Digital Markets Act (DMA)	EU regulation addresses market power of large digital platforms and ensuring fair competition and data access in digital ecosystems.	Data Space Authority, Data Space Operator, Data Space Provider	Relevant for platform governance and fair participation rules.
	EU Chips Act	EU regulatory framework strengthening semiconductor supply chain resilience, innovation and production capacity in Europe.	Data Space Authority, Data Space Provider, Data Space Consumer	Relevant for semiconductor ecosystem governance and supply chain coordination.
National Legal Framework	BSIG (IT-Sicherheitsgesetz / BSI-Gesetz)	German cybersecurity law defining responsibilities of the Federal Office for Information Security (BSI) and establishing security obligations for critical infrastructure operators and federal institutions.	Data Space Operator, Federation Service Provider, Data Space Provider	Basis for national cybersecurity regulation in Germany. Implements parts of EU cybersecurity policy and supports NIS2 implementation.
	BDSG (Bundesdatenschutzgesetz)	German national data protection law supplementing the GDPR and regulating specific	Data Space Authority, Data Space Provider, Data Space Consumer	Complements GDPR with national provisions such as supervisory authority roles and specific processing rules.

		national aspects of data protection and supervision.		
Non-EU Extraterritorial Laws	CLOUD Act	US law allowing US authorities to request data from US-based service providers even if the data is stored outside the United States.	Data Space Provider, Data Space Operator, Data Space Component Operator	Relevant if data space infrastructure or data storage relies on US cloud providers such as AWS, Azure or Google Cloud. May conflict with EU data sovereignty principles.
	USA PATRIOT Act	US legislation enabling government authorities to access data for national security purposes, including data held by US companies	Data Space Provider, Data Space Operator, Data Space Component Operator	Important when infrastructure components are operated by US-based companies that may be subject to US jurisdiction.

Appendix B. Control Catalogues

The following catalogue consolidates the operational and organizational controls for the secure data space operation, assigned to the roles of chapter 3.1 and to the three Gateway Security Levels. Levels are additive from lowest to highest: a control assigned to Trust also applies at Critical, and a control assigned to Basic applies at all levels.

Note: These controls do not constitute binding regulatory or legal requirements. They represent best-practice recommendations developed by this expert group and are intended as guidance for implementation; deviations may be justified depending on individual risk assessments and organizational context.

Data Space Operator

Table 11: Control Catalogue for Data Space Operator

Ref	Control	Level	Obligation
OP-01	Operate an ISMS and ensure continuous compliance with NIS2, CRA, EU Data Act, GDPR, and other applicable EU and national regulations	All	MUST
OP-02	Maintain a CSIRT and a 24/7 SOC	All	MUST
OP-03	Maintain a documented Incident Response Plan with NIS2- and CRA-compliant reporting timelines	All	MUST
OP-04	Undergo certification by a Conformity Assessment Body	All	MUST
OP-05	Define and regularly publish security metrics	All	MUST
OP-06	Operate the Identity Provider as central trust anchor: PKI management, VC issuance and revocation, DID management with DNSSEC enforcement	All	MUST
OP-07	Ensure all components communicating over the internet report security alerts to the central SOC	All	MUST
OP-08	Ensure Discovery Service / Service Catalog implements rate limiting and abuse detection	All	MUST
OP-09	Ensure BPN-DID-Resolution maps BPNL to DID without revealing transfer metadata	All	MUST
OP-10	Ensure onboarding includes identity verification via DCH, country clearance, framework agreement, security self-assessment, minimum security requirements, VC issuance, DID creation with DNSSEC, BPN assignment, audit trail	All	MUST
OP-11	Ensure offboarding includes immediate VC revocation, BPN deactivation, DID document update, connector deregistration, contract termination, GDPR-compliant data deletion, audit trail, wallet blocking, SOC and participant notification	All	MUST
OP-12	Ensure emergency containment measures are formally documented, time-limited, and auditable	All	MUST
OP-13	Define staged operational modes (Normal, Elevated Risk, Extreme) for emergency situations	All	MUST

OP-14	Maintain organizational mechanisms for consistent policy enforcement across data space boundaries	All	MUST
OP-15	Ensure interoperable identity and trust mechanisms, including alignment with eIDAS 2.0 and EUDI Wallet	All	MUST
OP-16	Coordinate joint incident response across data space boundaries	All	MUST
OP-17	Facilitate security intelligence sharing (IoCs, attack patterns, mitigations) between data spaces	All	MUST
OP-18	Maintain SLAs for all operated core services with availability, RTO, and RPO targets	All	MUST
OP-19	Perform country risk assessment and sanction/watchlist screening during onboarding	All	MUST
OP-20	Must not expose participant identities or data offering metadata in the Discovery Service without explicit participant consent	All	MUST NOT
OP-21	Replace did:web with more secure methods (e.g., did:webvh) once production-mature	Basic	SHOULD
OP-22	Integrate into existing certification frameworks (Catena-X CAB, IDS Certification, EUCC) with sector-specific profiles rather than building fully independent certification	Basic	SHOULD
OP-23	Include SEMI standards in minimum security requirements for semiconductor participants	Basic	SHOULD
OP-24	Require security training proof in the onboarding process	Basic	SHOULD
OP-25	Implement cross-data space security intelligence sharing mechanisms	Basic	SHOULD
OP-26	Ensure certification evidence for high-criticality use cases is machine-readable and verifiable via remote attestation	Critical	SHOULD
OP-27	Participate in Security Information Sharing (ISAC)	Basic	MAY
OP-28	Offer additional attestation services	Critical	MAY

Data Space Component Operator

Table 12: Control Catalogue for Data Space Component Operator

Ref	Control	Level	Obligation	Applies to
CO-01	Ensure all operated gateways meet the requirements of the applicable Gateway Security Level	All	MUST	COC
CO-02	Apply cryptographic requirements of BSI TR-02102 (AES-256-GCM mandatory, RSA-2048 immediately deprecated)	Basic	MUST	COC
CO-03	Authenticate all EDC management API calls (OAuth2/JWT bearer or API keys)	Basic	MUST	COC

CO-04	Enforce strict separation of Control Plane and Data Plane	Basic	MUST	COC
CO-05	Token-based access control with short-lived tokens for the data plane	Basic	MUST	COC
CO-06	WAFI protection for the data plane: Layer-7 inspection, IDS protocol guard, DoS/DDoS protection, content security scanner	Basic	MUST	COC
CO-07	Store all secrets in a secure vault with access controls and audit logging	Basic	MUST	COC
CO-08	Rotate secrets at intervals ≤ 90 days	Basic	MUST	COC
CO-09	Restrict vault access to authenticated service accounts	Basic	MUST	COC
CO-10	Address known EDC architectural boundaries through organizational and technical mitigations	Basic	MUST	COC
CO-11	Enforce the Zero Trust principle	Basic	MUST	COC
CO-12	Use only signed container images from trusted registries	Basic	MUST	COC
CO-13	Implement network policies with default-deny between namespaces	Basic	MUST	COC
CO-14	Implement RBAC with minimal permissions for service accounts	Basic	MUST	COC
CO-15	Use TLS 1.2+ with mTLS and BSI-approved cipher suites	Basic	MUST	COC
CO-16	Maintain structured vulnerability management: continuous CVE monitoring, SBOM management, coordinated vulnerability disclosure	Basic	MUST	COC
CO-17	Deploy production services in Multi-AZ configurations	Basic	MUST	COC
CO-18	High-availability deployment with tested failover	Basic	MUST	COC
CO-19	Maintain a Business Continuity Plan covering connector outages and alternative paths	Basic	MUST	COC
CO-20	Ensure NIS2-compliant incident reporting (≤ 24 h early warning, ≤ 72 h full report) if in scope	Basic	MUST	COC
CO-21	Ensure CRA-compliant vulnerability and incident reporting (≤ 24 h, ≤ 72 h, ≤ 14 d) if acting as manufacturer	Basic	MUST	COC-D
CO-22	Ensure the data plane does not persist transferred data beyond the active transfer session	Basic	MUST	COC-D
CO-23	Maintain DSIRT/PSIRT processes for post-market surveillance of in-house components	Basic	MUST	COC-D
CO-24	Conduct threat modeling for each in-house component at least annually and after major architecture changes	Basic	MUST	COC-D

CO-25	Sign all release patches of in-house components	Basic	MUST	COC-D
CO-26	Backport security fixes to still-supported older releases	Basic	MUST	COC-D
CO-27	Supplier due diligence and contractual security requirements for third-party providers	Basic	MUST	COC
CO-28	Continuously monitor the Eclipse Tractus-X supply chain	Basic	MUST	COC
CO-29	Hybrid PQC for IP-related use cases (part of CO-02)	Critical	MUST	COC
CO-30	Logical and cryptographic multi-tenancy isolation in shared connector deployments	Trust	MUST	COC
CO-31	Encrypt data at rest with AES-256-GCM (databases, volumes, backups)	Trust	MUST	COC
CO-32	Encrypted, georedundant backups: daily incremental, weekly full	Trust	MUST	COC
CO-33	Private keys for DIDs and VCs secured in HSM or TPM	Trust	MUST	COC
CO-34	Separate identity per gateway/EDC instance (own key pair)	Trust	MUST	COC
CO-35	Tamper-evident audit logging of all contract negotiations and data transfer events	Critical	MUST	COC
CO-36	Secrets must not be stored in plaintext in configuration files, source code repositories, or unencrypted storage	All	MUST NOT	COC
CO-37	Data plane must not expose internal endpoints without WAFI filtering and access control	All	MUST NOT	COC
CO-38	Management API endpoints must not be exposed without authentication	All	MUST NOT	COC
CO-39	Consistent Zero Trust Architecture across all infrastructure layers	Basic	SHOULD	COC
CO-40	Regular penetration tests at least annually	Basic	SHOULD	COC
CO-41	Regular integrity verification of DID documents	Basic	SHOULD	COC
CO-42	Containers with no privileged access and read-only root filesystems	Basic	SHOULD	COC
CO-43	Disaster recovery tests semi-annually	Basic	SHOULD	COC
CO-44	Resilience testing: chaos engineering, tabletop exercises	Basic	SHOULD	COC
CO-45	Graceful degradation for partial ecosystem operation during component failures	Basic	SHOULD	COC
CO-46	Automated security scans within the CI/CD pipeline	Basic	SHOULD	COC-D

CO-47	Consider confidential computing (data protection in use)	Trust	SHOULD	COC
CO-48	Consider platform integrity protection (secure boot, trusted boot, TPMs)	Trust	SHOULD	COC
CO-49	HSM-based key management with rotation ≤ 1 year, per BSI TR-02102	Trust	SHOULD	COC
CO-50	Hybrid Post-Quantum Cryptography for IP-related cases as early as technically feasible	Critical	SHOULD	COC
CO-51	Confidential computing with TEE and remote attestation	Critical	SHOULD	COC
CO-52	Bug bounty program for in-house components	Basic	MAY	COC-D
CO-53	Anomaly detection and criticality assessment	Trust	MAY	COC
CO-54	Additional attestation services	Critical	MAY	COC

COC = Component Operator Company

COC-D = Component Operator Company, which develops its own software

Data Provider and Data Consumer

Table 13: Control Catalogue for Data Provider and Data Consumer

Ref	Control	Level	Obligation	Applies to
P-01	Establish a CSMS according to IEC 62443-2-1 for OT systems	Basic	MUST	w/ OT
P-02	Segment internal IT/OT networks according to IEC 62443-3-2 (Zones & Conduits)	Basic	MUST	w/ OT
P-03	Implement an OT-DMZ architecture	Basic	MUST	w/ OT
P-04	No direct connector attachment to OT systems (Zone 4); access only via OT-DMZ or IT side	Basic	MUST	w/ OT
P-05	Ensure IEC 62443-4-2 CR 3.4 (Software and Information Integrity) at least SL-2 on the OT side	Basic	MUST	w/ OT
P-06	Firewalls between OT and DMZ and between IT and DMZ	Basic	MUST	w/ OT
P-07	Jump servers for controlled and monitored remote maintenance of OT assets	Basic	MUST	w/ OT
P-08	Structure infrastructure according to IEC 62443 Zone-and-Conduit model	Basic	MUST	w/ OT
P-09	Map data space security requirements and levels to concrete IACS security requirements (IEC 62443-3-3, 4-2)	Basic	MUST	w/ OT
P-10	GDPR-compliant data retention and deletion policies for own processed data	Basic	MUST	All

P-11	Choose SL-T for FR2 (Use Control) and FR4 (Data Confidentiality) accordingly when receiving strictly confidential data at OT level	Critical	MUST	Consumer
P-12	Choose high SL-T for FR3 (System Integrity) when OT source data integrity is critical	Critical	MUST	Provider
P-13	Align OT security levels with the mapping Basic → SL-1, Trust → SL-2, Critical → SL-3 + Remote Attestation	Basic	SHOULD	w/ OT
P-14	Use DNSSEC for did:web controllers	Basic	SHOULD	All
P-15	No fallback to OS-provided DNS resolvers when requesting DID documents	Basic	SHOULD	All
P-16	Security training proof for employees with data space access	Basic	SHOULD	All

w/OT = with Operational Technology

All Roles

Table 14: Control Catalogue for all roles

Ref	Control	Obligation
ALL-01	GDPR compliance: lawful basis, privacy by design, data minimization, purpose limitation	MUST
ALL-02	EU Data Act compliance: data portability, access rights, fair data sharing	MUST
ALL-03	Export control compliance for technology and data transfer	MUST
ALL-04	Retention and deletion policies for all processed data	MUST
ALL-05	Internal incident reporting to SOC/CSIRT within 1 hour	MUST
ALL-06	EU data centers for production services, EU regions preferred for cloud	SHOULD
ALL-07	Document all data flows, especially cross-border	SHOULD
ALL-08	Assess US-based cloud providers regarding CLOUD Act / PATRIOT Act exposure	SHOULD
ALL-09	Data processing agreements with subprocessors (GDPR Art. 28)	SHOULD

Appendix C. Kubernetes Security Architecture Model

The Kubernetes Security Architecture Model documents (on the next pages) target the security architecture for Kubernetes-based data space components from two perspectives. Figure 26 shows the model from the security level perspective, mapping the three data space security levels to Kubernetes security controls and layers. Figure 27 shows the tool perspective, mapping the security functions of the reference stack in chapter 4.2.2 to the concrete open-source toolchain and their integration in the cluster architecture.

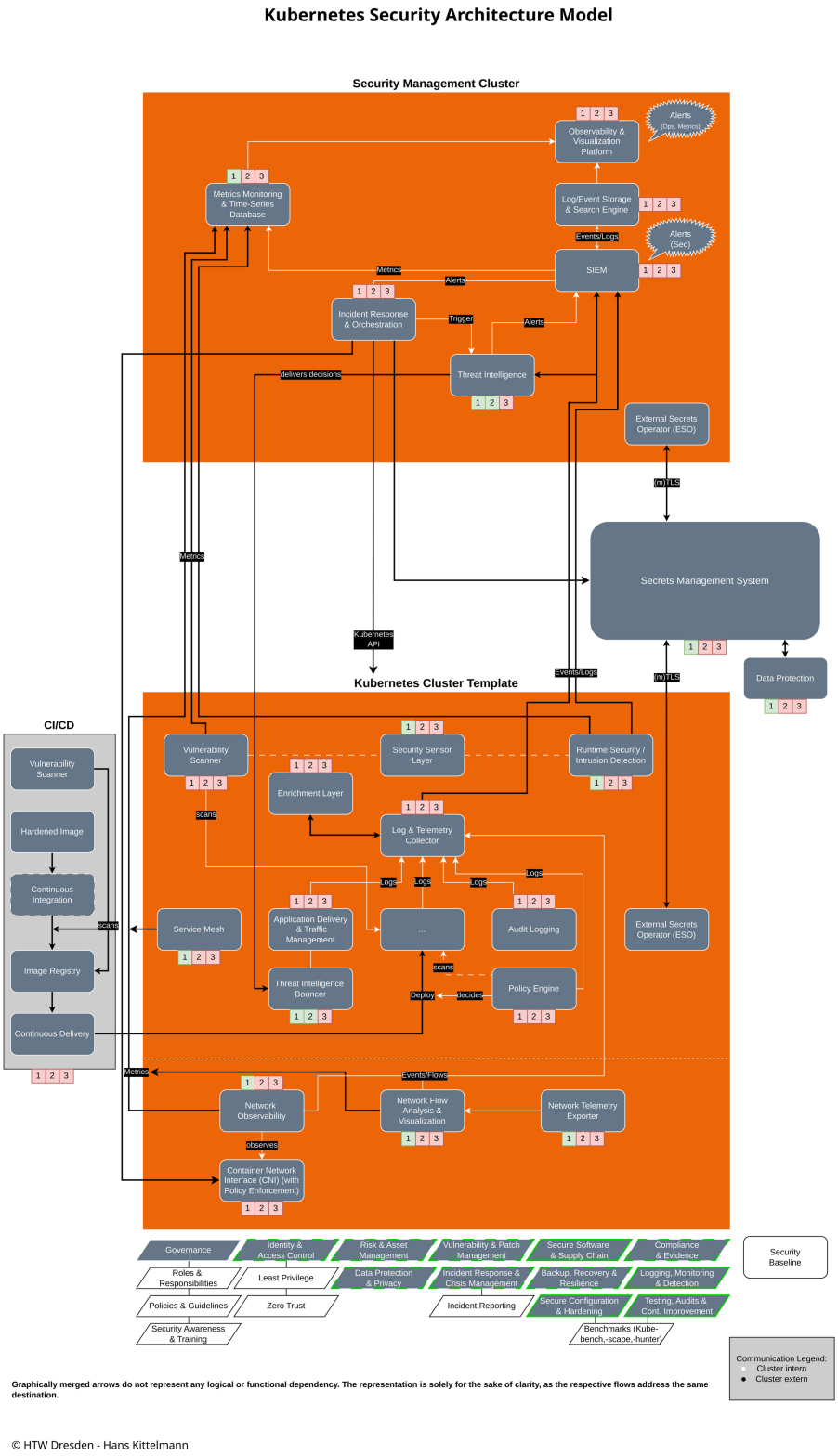
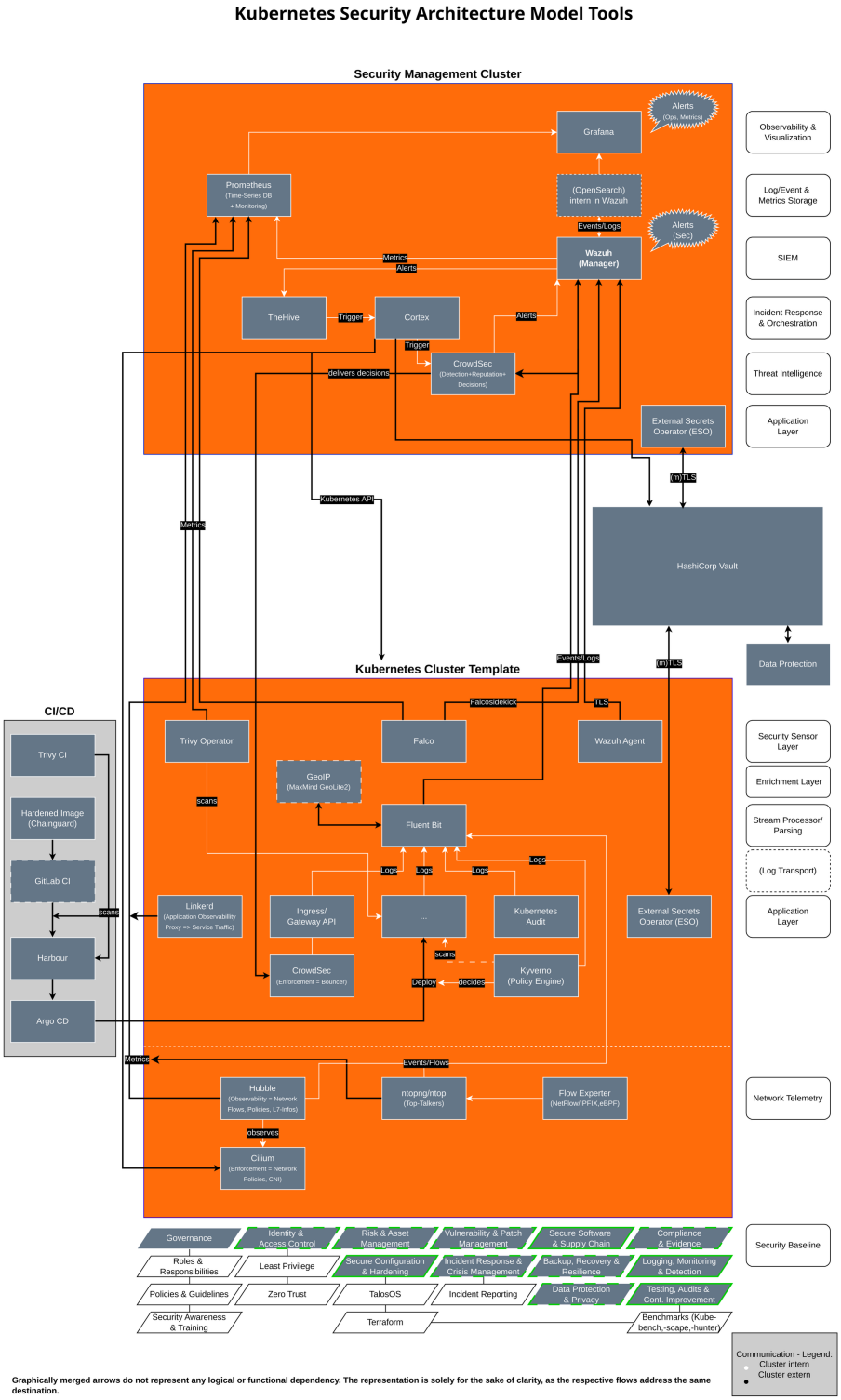


Figure 26: Kubernetes Security Architecture Model



© HTW Dresden - Hans Kittelmann

Figure 27: Kubernetes Security Architecture Model with Tools

Appendix D. Miscellaneous

Exemplary Evidence Catalogue

Table 15 Exemplary Evidence Catalogue

#	Evidence item	Automatable	A: Bare metal / self-managed Kubernetes	B: Cloud IaaS or PaaS	C: Operator-hosted
1	Architecture and data flow description	Partly	✓	✓	✓
2	Shared-responsibility matrix, control by control	No	–	✓	✓
3	Component inventory with image digests	Yes	✓	✓	✓
4	SBOM plus vulnerability report	Yes	✓	✓	✓
5	Signatures and build provenance	Yes	✓	✓	✓
6	Protocol conformance and interoperability report	Yes	✓	✓	✓
7	Component property credential (CPG attestation)	Yes	✓	✓	✓
8	Secure-configuration scan of node, cluster, connector	Yes	✓	✓	✓
9	Key management and identity custody statement	Partly	✓	✓	✓
10	Inherited infrastructure certificates plus validity check	Partly	–	✓	✓
11	IEC 62443 zone and conduit evidence, SEMI E187 context	Partly	✓	–	○
12	Penetration test report	No	✓	✓	✓
13	IEC 62443-4-1 development process evidence	Partly	✓	✓	✓
14	ISMS documentation and training records	Partly	✓	✓	✓
15	Logging, retention, and clearing integration proof	Yes	✓	✓	✓
16	Patch SLA with change-freeze compensating controls	Partly	✓	✓	✓
17	Backup and disaster recovery test records	Partly	✓	✓	✓
18	Incident response plan and reporting commitment	No	✓	✓	✓

19	ODRL policy library and enforcement evidence	Yes	✓	✓	✓
20	Multi-tenancy isolation evidence	Partly	–	–	✓
21	Residency, jurisdiction, and export-control statement	Partly	✓	✓	✓
22	Exit and portability plan	No	–	✓	✓
23	Runtime attestation: TPM or confidential computing	Yes	✓	○	○

✓ required · ○ conditional · – not applicable