EDGE-Skills
Training Curriculum

Project 101123471 – EDGE-Skills

WP4: D4.6 EDGE SKILLS Training program

Empowering professionals to create interoperable, human-centric data ecosystems

A practical, role-based learning journey across strategy, governance, data, technology, and operations.

Version 10.5 | July 2026

Co-funded by the European Union

Summary

Data spaces are becoming a cornerstone of Europe's digital transformation. They enable organisations to share and reuse data securely, ethically, and interoperably, while preserving sovereignty and building trust. In the skills and education domain, data spaces unlock new possibilities for personalized learning, talent development, workforce mobility, and AI-driven services.
The EDGE-Skills project, built on the Prometheus-X open source stack, provides the technical, semantic, and governance foundation for such a federated ecosystem. However, successful data spaces require more than technology: they demand aligned value creation, clear governance, semantic interoperability, robust architecture, and sustainable operations.
This curriculum equips professionals with the end-to-end knowledge and skills to design, build, and operate real data space services. It follows the natural lifecycle of a data space project: from strategic vision and use case design, through governance, data and semantics, technical architecture, implementation, and finally operational scaling.
The program is role-based and highly practical. Project managers, data experts, and software engineers each develop deep expertise in their domain while learning to collaborate across roles. Through real use cases, templates, service chains, and hands-on integration, participants gain the confidence to move from concept to production.
This curriculum empowers organisations and ecosystems to collaborate and engage within trustworthy, interoperable, and human-centric data spaces that create lasting value.

Chapter Overview

Table of Contents

About this Training Curriculum

Data spaces are becoming a cornerstone of Europe's digital transformation. They enable organisations to share and reuse data securely, ethically, and interoperably, while preserving sovereignty and building trust. In the skills and education domain, data spaces unlock new possibilities for personalized learning, talent development, workforce mobility, and AI-driven services.

The EDGE-Skills project, built on the Prometheus-X open source stack, provides the technical, semantic, and governance foundation for such a federated ecosystem. However, successful data spaces require more than technology: they demand aligned value creation, clear governance, semantic interoperability, robust architecture, and sustainable operations.

This curriculum equips professionals with the end-to-end knowledge and skills to design, build, and operate real data space services. It follows the natural lifecycle of a data space project: from strategic vision and use case design, through governance, data and semantics, technical architecture, implementation, and finally operational scaling.

The program is role-based and highly practical. Project managers, data scientists, and software engineers each develop deep expertise in their domain while learning to collaborate across roles. Through real use cases (centered on the Skill Gap Analytics scenario), templates, service chains, and hands-on integration, participants gain the confidence to move from concept to production.

Target Groups and Learning Objectives

This training curriculum targets three professional groups:

1. Project Managers / Business & Ecosystem Leaders

Who they are: Project managers, product owners, business developers, ecosystem facilitators, decision-makers, public sector stakeholders, innovation leaders.

Their role in a data space: They are responsible for defining the strategic vision, prioritising use cases, orchestrating stakeholders, and shaping governance and sustainability decisions, while coordinating closely with data and software roles for detailed design and implementation.

Learning objectives:

- Explain what a data space is and why it matters (strategic and societal perspective).
- Identify high-value use cases and translate them into actionable projects.
- Design sustainable business and governance models.
- Manage stakeholders, incentives, roles, and trust frameworks.
- Navigate legal, regulatory, and operational implications (GDPR, AI Act, Data Act).
- Lead interdisciplinary teams through the full lifecycle of implementation.
- Ensure long-term adoption, scalability, and ecosystem growth.

Goal: Be capable of designing, leading, and scaling a real data space initiative from concept to operation.

2. Data Scientists / Data Engineers / Semantic Specialists

Who they are: Data scientists, data engineers, data architects, data stewards, semantic modelers, ontology specialists, ML engineers, interoperability experts.

Their role in a data space: They design and implement the data pipeline: from ingestion through semantic alignment, transformation, quality assurance, analytics, and re-offering of results. They are responsible for making data interoperable, trustworthy, and AI-ready across organizational boundaries.

Learning objectives:

- Map the complete data journey within a data space (ingestion → semantics → transformation → quality → analytics → re-offering).
- Apply semantic standards (xAPI, JSON-LD, RDF, ESCO, ROME) for interoperability.
- Design transformation pipelines, ontology mappings, and normalization rules.
- Implement data quality assurance with provenance and lineage tracking.
- Build ML models and analytics pipelines that produce re-offerable data assets.
- Apply the data journey model (Data → Transformation → Analytics → Application) to a concrete pipeline.
- Configure Dataspace Connectors (PDC), catalog offers, and governance policies from a data perspective.

Goal: Be capable of designing and implementing the complete semantic, data, and analytics backbone of a data space.

3. Software Experts / Engineers / Integrators / DevOps

Who they are: Software developers, solution architects, system integrators, DevOps engineers, cloud/edge specialists, API designers, security engineers.

Their role in a data space: They build the technical infrastructure: Connectors (PDC), Data Planes, Identity Hubs, policy engines, service chains, and deployments that make the data space operational. They work with the Prometheus-X open source stack, implementing DSP for interoperability and extending it with the PDI and Consent Manager for personal data sharing.

Learning objectives:

- Describe the Prometheus-X open source stack: building blocks, core components, and the consent-driven PDI approach.
- Deploy and configure Connectors (Control Plane) with DSP catalog, negotiation, and transfer capabilities.
- Build and register custom Data Planes (Pull, Push, Stream patterns) using SDKs (Go, Java, Rust, .NET).
- Implement Identity Hub with DIDs (did:web), Verifiable Credentials, and DCP proof composition.
- Define and evaluate ODRL policies at runtime using the policy engine.
- Deploy multi-tenant environments using the PDC Tenant Manager with Activity Agents.
- Validate end-to-end flows: credential presentation → catalog discovery → contract negotiation → data transfer.
- Apply DevOps practices: Kubernetes, NATS messaging, PostgreSQL, Prometheus metrics, CI/CD.

Goal: Be capable of technically building, integrating, and operating a compliant and scalable data space infrastructure using protocol-driven architecture.

4. Cross-Role Shared Outcome

Each group learns their part of the puzzle, but also how to collaborate across roles to deliver real value:

- PMs understand enough technology and data to lead effectively.
- Data scientists understand business needs, governance, and technical constraints.
- Developers understand why data semantics and governance matter.
- All share a common vocabulary, mental model, and reference cases.

Proposed Structure and Didactic Model

Seven Modules

- Module 1: Foundations and Mindset (shared)
- Module 2: Use Case and Value Design (PM-led)
- Module 3: Governance and Trust (PM-deep, Data/Dev practical)
- Module 4: Data, Semantics, and Analytics (Data-deep, PM aware, Dev technical)
- Module 5: Technical Architecture and Building Blocks (Dev-deep, Data supporting, PM overview)
- Module 6: Implementation and Integration Journey (hands-on, cross-functional)
- Module 7: Operations and Scaling (PM/Business deep, Dev technical, Data analytical)

Didactical Model

- **Learning Paths:** Distinct learning journeys for each role (PM / Data Scientist / Developer), aligned with typical responsibilities and required depth.
- **Shared Core Modules:** Common foundation modules for everyone, building a unified vocabulary, mindset, and reference framework.
- **Interdisciplinary Project Modules:** Cross-functional, practice-oriented sessions where all roles collaborate on the EDGE-Skills Skill Gap Analytics use case.
- **Reference Case Methodology:** Each training uses a concrete reference case traced end-to-end through the technical stack, anchoring abstract concepts in practical scenarios.

Design Principles

- **Aligned to the Data Space Community:** The curriculum mirrors how successful data space projects are built and operated.
- **Pedagogical progression:** Each module builds cognition from “why” → “what” → “how” → “operate.”
- **Role-balanced:** Each role gets depth at the right time, but also shared understanding across disciplines.
- **Technically grounded:** Content reflects actual released training materials and real implementations, not theoretical constructs.
- **Future-proof:** Extensible into certifications (Practitioner → Expert → Architect), online modules, and domain-specific accelerators.

Teaching Formats

- Case-based workshops built on real EDGE-Skills use cases and reference implementations
- Templates and toolkits for immediate application (Offering Definition Card, Policy Templates)
- Practical labs (Connector configuration, Data Plane SDK integration, policy design with the policy engine)
- Role-based collaboration exercises across PM, Data, and Dev teams
- Optional certification pathways (“EDGE-Skills Practitioner / Expert / Architect”)

Learning Matrix

Module	PM Focus	Data Scientist Focus	Developer Focus
Foundations	Strategic context, ecosystem vision	Data journey overview, standards landscape	Architecture overview, protocol paradigm
Use Case & Value	Deep: ecosystem design, incentives	Data requirements for value cases	Service chain positioning
Governance & Trust	Deep: roles, contracts, AI Act	Practical: ODRL policies, consent	Deep: DCP, VC, policy enforcement
Data & Semantics	Business implications of data quality	Deep: 4-layer flow, semantics, ML	API schemas, validation logic
Tech Architecture	Overview and alignment	Interface requirements, data access	Deep: PDC, DSP, PDI, Data Planes
Implementation	Coordination, testing	Data flows, validation, analytics	Deployment, DevOps, E2E validation
Operations	Licensing, KPIs, growth	Quality monitoring, stewardship	PDC Tenant Manager, observability, scaling

Role-Based Guidance for Engaging with the Curriculum

This curriculum follows a role-based structure that mirrors real responsibilities in data space initiatives. Each module is relevant to all participants but addresses different roles with different depth and expectations.

Project Managers / Business and Ecosystem Leaders

As a project manager, the curriculum primarily supports strategic design, coordination, governance, and long-term operation of a data space.

- Module 1 (Foundations) establishes common understanding of purpose, vision, roles, and the paradigm shift from platforms to data spaces. It includes the definition of a data space based on the DSSC Blueprint and ISO/IEC 20151 and seven essential elements that frame all subsequent work.
- Module 2 (Use Case and Value Creation) is central to your role. It directly reflects your responsibility for ecosystem design, value propositions, incentive structures, and sustainable service chains.
- Module 3 (Governance and Trust) strongly relates to governance models, roles, contracts, compliance frameworks, and the Decentralized Claims Protocol (DCP).
- Module 4 (Data and Semantics) provides awareness of how data quality, semantics, and the data journey model affect feasibility, trust, and scalability.
- Module 5 (Technical Architecture) provides understanding of architectural implications, the Prometheus-X building blocks, and the consent-driven PDI approach for realistic planning.
- Module 6 (Implementation) supports coordination, sequencing, and risk management during integration using reference cases.

- Module 7 (Operations and Scaling) directly reflects licensing schemes, KPIs, ecosystem growth, and VisionsTrust Platform operational tooling.

Data Scientists / Data Engineers / Semantic Specialists

As a data scientist, the curriculum supports your responsibility for the complete data journey: from ingestion to re-offering of analytics results.

- Module 1 (Foundations) provides the conceptual reference for understanding how your data and semantic work fits into the broader ecosystem.
- Module 2 (Use Case and Value Creation) clarifies which data assets, quality levels, and semantics are required to support concrete value creation.
- Module 3 (Governance and Trust) directly affects your work through ODRL policies, consent, contracts, and compliance requirements shaping data handling.
- Module 4 (Data and Semantics) is your core module. It covers the complete data journey across four layers (Data → Transformation → Analytics → Application), semantic standards, transformation pipelines, quality assurance, and privacy-preserving processing.
- Module 5 (Technical Architecture) clarifies how semantic and data requirements are implemented through Dataspace Connectors, Data Planes, and service chains.
- Module 6 (Implementation) relates to implementing semantic models, transformation pipelines, and validation within integrated data flows.
- Module 7 (Operations and Scaling) supports long-term data stewardship, quality monitoring, and contribution to KPIs.

Software Experts / Engineers / DevOps

As a software expert, the curriculum supports your responsibility for building, integrating, and operating the technical infrastructure using protocol-driven architecture.

- Module 1 (Foundations) establishes the paradigm shift from monolithic applications to protocol-driven ecosystems and the seven essential elements of a data space.
- Module 2 (Use Case and Value Creation) provides value context that guides service design and service chain orchestration.
- Module 3 (Governance and Trust) directly affects your role through identity, consent, Verifiable Credentials, ODRL policy evaluation, and compliance-by-design.
- Module 4 (Data and Semantics) relates to implementing schemas, semantic APIs, JSON-LD payloads, and validation logic.
- Module 5 (Technical Architecture) is your core module. It covers the PDC Connector, Data Planes with SDKs, Identity Hub with credential management, PDI for consent-driven data sharing, PDC Tenant Manager multi-tenant deployment, DSP-based negotiation, and end-to-end flow validation.
- Module 6 (Implementation) reflects integration, federated data flows, policy execution, and DevOps practices.
- Module 7 (Operations and Scaling) supports PDC Tenant Manager operations, monitoring with Prometheus/NATS, SLAs, reliability, and scalable production.

Module 1: Foundations (shared)

> **Didactic Goal:** *Build a unified mental model before diving into complexity. Eliminate the biggest early risk: misalignment of vocabulary, expectations, and mental models across roles.*

> 📚 **Training Material:** PM Training slides 1–8 | Developer Training slides 1–13 | Data Scientist Training slides 1–13 (shared foundation across all three trainings; the PM deck covers the same core concepts in fewer slides because it focuses on strategic understanding, while the Developer and Data Scientist decks include additional technical detail on architecture and building blocks)

A data space only works when everyone involved shares the same understanding of what a data space actually is. This module creates a common foundation across all roles by establishing a precise, standards-based definition and exploring the seven essential elements that define any data space.

The module introduces the paradigm shift from traditional centralized platforms to protocol-driven, federated ecosystems. Participants learn that a data space is an environment (not a platform) where trust is established through Verifiable Credentials, communication happens peer-to-peer via Connectors, and governance is maintained by a Governance Authority using machine-readable ODRL policies.

The Prometheus-X and EDGE-Skills architecture is introduced at a high level, showing how the Prometheus-X open source stack's core components (PDC Connector, Data Planes, Identity Hub, Catalogue, Contract Service) and the consent-driven PDI (Personal Data Intermediary) interact to enable real data spaces. Prometheus-X implements DSP for interoperability and extends it with the PDI and multi-party contracts.

Outcome: All participants leave with a shared language, a clear understanding of the data space vision and architecture, and the right mindset to collaborate effectively.

1.1 Purpose and Vision of Data Spaces

- Strategic goals behind data spaces: economic, societal, and innovation potential
- Data spaces as enablers of trusted collaboration across organizational boundaries
- The role of data spaces in European digital policy and sectoral strategy

1.2 Human-Centric Principles and EU Data Strategy

- Data sovereignty and individual control
- Transparency, consent, and ethical data use
- EU Data Strategy and sectoral data spaces
- Human-centric design as a guiding principle for all architecture decisions

1.3 Overview of Prometheus-X and EDGE-Skills Architecture

- High-level architecture: three layers (Legislative, Economic, Technical)
- Core components: Connector (PDC), Data Planes, Identity Hub, Catalogue, Contract Service
- Interoperability: DSP for discovery, negotiation, and transfer; DCP for credential exchange; PDI (Personal Data Intermediary) for consent-driven personal data sharing (Prometheus-X's key differentiator)

1.4 Roles and Collaboration in Data Spaces

- Core ecosystem roles: provider, consumer, intermediary, operator, governance body
- Project roles: PM, Data Scientist, Software Expert, and their intersections
- Role interactions and dependencies in real data space projects
- Four elements every role must understand: Catalog, Connector, Contracts and Policies, Consent

1.5 From Platforms to Data Spaces

- Centralized vs. decentralized governance
- Data ownership vs. data usage rights
- API integration vs. protocol state machines

Module 2: Use Case and Value Creation (PM-led, all involved)

> **Didactic Goal:** *Shift from technology-first thinking to value-driven ecosystem design. Ensure that everything built in later modules has purpose, demand, and strategic alignment.*

> 📚 **Training Material:** PM Training slides 6–7 (service chains, critical distinction) | PM Training slides 10 and 19 (Offering Definition Card)

Technology is rarely the reason a data space succeeds or stalls; value is. This module addresses that by placing use cases, incentives, and ecosystem design at the center. Participants learn that a data space is not a product; it is an economic and collaborative system that only works when every stakeholder sees tangible benefit.

The module introduces the Offering Definition Card as a practical tool: a one-page document that defines what an offering is, who it serves, what problem it solves, and what data flows in and out. This aligns all stakeholders before any technical work begins.

Service chains are introduced as a business construct: chaining services across organizations to create composite value. Each offering is positioned as an entry point, transformation step, or end service within a broader value chain.

2.1 Ecosystem and Stakeholder Design

- Identifying relevant stakeholders and roles
- Understanding motivations, needs, and risks
- Mapping capabilities and data assets
- Aligning interests to build cooperation

2.2 Value Propositions and Incentive Structures

- Defining value for each stakeholder type (economic, operational, strategic, societal)
- Incentives to share and collaborate
- Avoiding free-rider and imbalance scenarios

2.3 Business Model Design and Validation

- Role-based business models (builder, collaborator, service provider)
- Revenue and cost structures

- Monetization vs. public value approaches
- Validating sustainability and scalability

2.4 Service Chains and Sustainability Planning

- Concept of service chains across organizations, implemented via the Data Processing Chain Protocol (DPCP)
- Positioning offerings as composable elements (entry point, transformation, end service)
- Ensuring long-term viability and operational costs
- Planning phased rollout and ecosystem growth

Module 3: Governance and Trust (PM-deep, Data/Dev practical)

> **Didactic Goal:** *Enable participants to design the rules, roles, and trust mechanisms that make data spaces actually work. Transform "trust" from a buzzword into an operational reality.*

> 📚 **Training Material:** Developer Training slides 14–16 (Identity Hub, VCs, ODRL) | PM Training slides 12–13 (governance, consent)

Technology does not create trust; governance does. This module teaches participants how to build the institutional, legal, and policy foundations of a trustworthy data space. A central focus is the PDI (Personal Data Intermediary) and the Consent Manager (Prometheus-X's most distinctive components), enabling consent-driven personal data sharing where individuals manage their consents from a single place. With the Developer Training's depth on Verifiable Credentials and ODRL policy evaluation, this module bridges the gap between legal governance and technical enforcement.

3.1 Governance Models and Roles

- Centralized, federated, and intermediary-based governance structures
- Allocation of responsibilities (business, technical, legal)
- Role of operator, facilitator, and governing body
- Decision-making and escalation mechanisms

3.2 Contracts, Consent Management, and ODRL Implementation

- Role of contracts in enabling trust and compliance
- Consent principles and lifecycle (ISO/IEC 27560)
- ODRL as a machine-readable policy language

3.3 Decentralized Claims Protocol (DCP) and Verifiable Credentials

- Why delegation is necessary in real B2B/B2C services
- DCP overview: peer-to-peer credential exchange without central authority
- Claim issuance, exchange, and verification flows

3.4 Legal Compliance (GDPR, AI Act, Data Act)

- GDPR obligations (lawful basis, consent, data minimization, rights)

- AI Act requirements (risk levels, transparency, accountability)
- Data Act implications (B2B, B2G data sharing obligations)
- Compliance-by-design approach

3.5 Lifecycle and Operational Governance

- Onboarding and certification of participants
- Managing policy changes and rule updates
- Conflict resolution and audits
- Continuous governance evolution over time

Module 4: Data, Semantics, and Analytics (Data-deep, PM aware, Dev technical)

> **Didactic Goal:** *Build the semantic and data foundation for interoperability, quality, and AI-readiness. Trace the complete data journey from ingestion to re-offering.*

> 📚 **Training Material:** Data Scientist Training slides 10–31 (complete DS curriculum) | data journey model

Without semantics, a data space is just a pile of incompatible datasets. Without quality and veracity, no one will trust or reuse the data. This module equips participants with the methods, standards, and tools to make data interoperable, meaningful, and reliable, and shows how analytics results become sovereign data assets themselves.

The module is anchored by the Skill Gap Analytics reference case (Schülerkarriere → Skill Analytics Service → imc), demonstrating the complete data journey from raw data to actionable intelligence.

4.1 The Data Journey in a Data Space

- Collection does not equal value creation; each step adds value and builds trust
- Four-layer data journey model: Data Layer (collection with provenance), Transformation Layer (semantic alignment and harmonization), Analytics Layer (quality assurance and analysis), Application Layer (re-offering as new data assets)
  - Layer 1, Data Layer: Data sources with assets, Dataspace Connectors (PDC), federated discovery via catalogue, Data Offers (Metadata + Usage Conditions)
  - Layer 2, Transformation Layer: Data preparation and cleansing, transformation into a unified data model, schema harmonization, semantic alignment (e.g. ESCO mapping), feature normalization, constraint and validation modeling
  - Layer 3, Analytics Layer: Skill Analytics Service with harmonized data, skill gap analysis, trend identification, competence prediction, risk assessment, output as Skill Gap Analytics Asset re-offered to the data space
  - Layer 4, Application Layer: Skills Dashboard App, recommended learning pathways, alternative qualification routes, trade-off explanation, decision-making and validation by HR professionals and learners

4.2 Data Ingestion and Connectors

- Connecting databases, APIs, file systems, IoT streams via Dataspace Connectors (PDC)
- Each dataset becomes an Asset with metadata, DataAddress, and usage conditions
- Ingestion patterns: Pull (query on demand), Push (batch delivery), Stream (continuous flow)

4.3 Semantic Standards and Ontologies

- Purpose of semantics in cross-organizational interoperability
- Core standards: xAPI (learning traces), JSON-LD (linked data context), RDF (knowledge graphs)
- Classification frameworks: ESCO (European Skills, Competences, Qualifications), ROME (French job classification)
- Custom ontologies and domain-specific extensions
- Linking data via semantic relationships across providers

Without shared semantics, data from different organizations cannot be compared, combined, or analyzed. Semantic alignment is the foundation of every successful data space.

4.4 Data Transformation and Preparation

- Four-step pipeline: Map (schema alignment) ⟶ Transform (format conversion) ⟶ Enrich (semantic enrichment) ⟶ Validate (quality checks)
- Building reusable transformation services that other participants can chain
- Metadata tracks provenance and lineage at each transformation step

4.5 Metadata and Catalog Management

- Four types of metadata: Technical (format, schema, API endpoints), Semantic (ontology references, domain tags), Contractual (ODRL policies, license type), Operational (update frequency, data lineage)

- Catalogs as data and service registries
- Enriching metadata with semantic information
- Enabling federated discovery across multiple data spaces

Rich metadata enables discovery, which builds trust, which enables reuse: the virtuous cycle of the data space.

Note: The Data Scientist Training consolidates Operational metadata under Technical metadata for practical simplicity, resulting in three categories (Technical, Semantic, Contractual).

4.6 Data Quality and Veracity

- Quality dimensions: accuracy, completeness, timeliness, consistency
- Provenance and lineage management with Data Value Chain Tracker
- Prometheus-X Data Veracity Assurance for integrity validation

4.7 Governance and Policies for Data

- Three governance layers from a data perspective: Access policies, Usage policies, Consent management
- ODRL machine-readable policies enforced automatically by Connectors

4.8 Publish, Re-Offer, Consume, and Integrate Data

- Register data assets with metadata, quality level, and provenance
- Set usage conditions via ODRL policies enforced automatically
- Re-offer analytics results as new data assets in the catalog

4.9 Privacy-Preserving Data Processing

- Anonymization and pseudonymization techniques
- Federated processing: compute at data source
- Differential privacy: add calibrated noise
- Secure multi-party computation and synthetic data
- Balancing utility and privacy for each use case

Module 5: Technical Architecture and Building Blocks (Dev-deep, Data supporting, PM overview)

> **Didactic Goal:** *Understand how to construct the technical foundation of a data space using modular, interoperable building blocks and protocol-driven architecture.*

> 📚 **Training Material:** Developer Training slides 14–34 (complete developer curriculum)

A data space is not a single platform; it is an orchestrated network of interoperable components communicating via standardized protocols. This module shows how the Prometheus-X open source stack provides the technological foundation, with each component serving a specific role in the trust and data transfer chain.

5.1 The Prometheus-X Architecture

- Three layers: Legislative (governance, compliance, legal), Economic (providers, consumers, marketplaces), Technical (protocols, connectors, identity, data planes)
- Core components: Connector (PDC), Data Planes, Identity Hub, Catalogue, Contract Service
  - DSP (Dataspace Protocol): Control Plane ↔ Control Plane: catalog discovery, contract negotiation, transfer coordination
  - DCP (Decentralized Claims Protocol): Identity Hub ↔ Identity Hub: credential presentation and verification
  - PDI (Personal Data Intermediary): Prometheus-X's key differentiator: consent-driven personal data sharing where individuals centrally manage their consents (Separation of Powers Principle)

5.2 The Connector (Control Plane)

- The central software agent representing a participant in the data space
  - Catalog Publication: dynamic catalog based on Consumer credentials: different Consumers see different offerings
  - Contract Negotiation: verify credentials, evaluate policies, produce binding Contract Agreements (REQUESTED → OFFERED → AGREED → VERIFIED → FINALIZED)
  - Policy Evaluation: Policy engine for runtime evaluation without redeployment
  - Transfer Coordination: selects Data Plane, issues authorization signals, never handles data directly

5.3 Data Planes

- Where actual data moves, and where developers write custom code
  - Pull: Consumer fetches from Provider (APIs, queries)
  - Push: Provider sends to Consumer (batch exports, files)
  - Stream: Continuous flow (IoT, telemetry)

5.4 Identity Hub and Decentralized Identity

- Each participant's wallet for credentials and DIDs
  - DID Management: creates did:web identifiers, manages signing keys, publishes DID Documents
  - Credential Storage: Membership, Attribute, and Role-specific Credentials
  - Proof Composition: selects relevant credentials, cryptographic presentation via DCP
  - Verification: checks signatures, validity, revocation status, trust anchors

Trust Sequence: DID Creation → Credential Acquisition → Hub Storage → Proof Presentation → Verification → Policy Evaluation

5.5 Publishing and Consuming Data via DSP

Publishing (three Control Plane objects):

  - Create Asset: Metadata + DataAddress pointing to Data Plane (Control Plane holds only metadata)

- Define Policy: ODRL policy specifying access requirements with credential-based authorization
- Contract Definition: Links Asset + Policy, makes offering discoverable and negotiable

Consuming (three-step data consumption protocol):

- Catalog Discovery: query Provider's DSP endpoint, catalog generated dynamically based on credentials
- Contract Negotiation: REQUESTED ⟶ AGREED ⟶ VERIFIED ⟶ FINALIZED (seconds, no human intervention)
- Transfer Initiation (EDC-based transfers): receive EDR, retrieve data directly from Provider's Data Plane

No portal accounts, no API keys, no custom code, no manual approvals. Standardized across all providers.

5.6 Multi-Tenant Deployment with PDC Tenant Manager

PDC Tenant Manager for scalable operations:

- PDC Tenant Manager is NOT in the trust decision path; trust decisions happen directly between Connectors peer-to-peer
- Two-step provisioning: Create Tenant (POST with tenant-ID) ⟶ Deploy Participant Profile (bind to Cell + DID + Dataspace Profiles)
- Four sequential Activity Agents: IDP Agent (OAuth2), PDC Agent (participant contexts), Registration Agent (credential issuer), Onboarding Agent (VCs + storage)

5.7 Operational Tooling and DevOps

VisionsTrust Platform: comprehensive operational interface:

- Data Catalog, Data Spaces, Connections, Transfers, Compliance (DPP, CSRD, AI Act, LkSG, EUDR, DGA), Governance, Policy Engine, Graph View

Infrastructure: PostgreSQL, NATS messaging, Kubernetes, Secret Store

Observability: Prometheus metrics, log aggregation, alerting, Connector Observability API

Core principle: For file/JSON sharing over HTTP, the only custom code is a minimal Data Plane. Invest in configuration, not reimplementation.

Module 6: Implementation and Integration Journey (hands-on, cross-functional)

> **Didactic Goal:** *Move from theory to practice: turn design into working prototypes and integrated services using real reference cases.*

> 📚 **Training Material:** Developer Training slides 29–32 (E2E validation, failure patterns, toolkit) | Data Scientist Training slides 26–29 (data flow validation, failure patterns)

Anchored by the concrete reference case used across all trainings: Schülerkarriere/Skill Analytics Service/imc (skill gap analytics), explored from each role's perspective.

6.1 Translating Requirements to Technical Specifications

- Converting use case and governance rules into system requirements
- Defining data models, policies, and interfaces
- Collaboration between PM, Data, and Dev roles

Using the Offering Definition Card and boundary definitions (internal vs. shared data, personal data triggers, service-chain roles, operational ownership)

6.2 Reference Cases for Implementation

Developer reference case: Skill Gap Analytics: Implementation Perspective

  - Schülerkarriere (Provider) publishes anonymized skill profiles via PDC and Catalogue
  - Skill Analytics Service (Transformation) consumes, transforms, and re-offers enriched data
  - imc (Consumer) receives personalized skill gap analysis via Contract Service and Data Plane
  - Trust model: Verifiable Credentials, no API keys, no manual approvals

Data Scientist reference case: Skill Gap Analytics across Organizations

  - Schülerkarriere shares anonymized student skill profiles
  - Skill Analytics Service maps to ESCO, identifies competency gaps
  - imc uses skill gap analytics for training design
  - Data flows through semantic transformation and quality checks; results re-offered as new assets

6.3 End-to-End Validation

Developer validation checklist:

- Can a Consumer discover your offering in the catalog?
- Are credentials correctly presented and verified during negotiation?
- Does the Contract Agreement get created successfully (FINALIZED)?
- Can data actually be transferred via the Data Plane (COMPLETED)?
- Are audit records generated for compliance?

Data Scientist validation checklist:

- Can data be discovered via catalog with semantic search?
- Are semantic standards correctly applied (xAPI, JSON-LD)?
- Does the Contract Agreement get created successfully?
- Are usage conditions and consent properly enforced?
- Does the pipeline produce valid, re-offerable insights?

6.4 Where Implementation Often Fails

Common developer failures:

- Deploy containers without configuring credentials

- Forget to register the Data Plane
- Write policies referencing non-existent credential types
- Hardcode trust instead of using the PDC protocol stack
- Treat catalog as static documentation

Common data integration failures:

- Ingest data without applying semantic standards
- Skip metadata enrichment and catalog registration
- Ignore data quality checks and provenance tracking
- Treat transformation as one-off instead of reusable pipeline
- Build analytics in isolation without re-offering results

6.5 DevOps and Scaling Strategies

- CI/CD pipelines for data space services
- Containerization and Kubernetes deployment
- Versioning and rollback strategies
- Scaling horizontally and vertically

Core deployment principle: config-based isolation, single deployment for all tenants, shared identity provider, health and readiness APIs

Module 7: Operations and Scaling (PM/Business deep, Dev technical, Data analytical)

> **Didactic Goal:** *Ensure the data space can run, evolve, and scale sustainably in the real world beyond the pilot stage.*

> 📚 **Training Material:** Developer Training slide 27 (VisionsTrust Platform) | Data Scientist Training slide 24 (operational interfaces)

7.1 Operational and Governance Models

- Transition from project to operational service
- Roles in operations (operator, facilitator, support)
- Governance integration in daily operations

VisionsTrust Platform as the central operational interface: Data Catalog, Data Spaces browser, Connections management, Transfer monitoring, Compliance tracking, Governance classification, Policy Engine, Graph View

7.2 Licensing Schemes (Builder, Collaborator, Service Provider)

- Overview of role-based licensing models
- How participants generate and capture value
- Usage-based or subscription-based models
- Supporting innovation and service marketplaces

7.3 Business Continuity and SLA Management

- Ensuring availability and reliability
- SLA definition and monitoring
- Incident response and escalation processes
- Risk management and disaster recovery

7.4 KPIs and Value Tracking

- Usage metrics and adoption monitoring
- Data quality and service performance indicators
- Measuring economic and societal impact
- Continuous feedback and optimization cycles

7.5 Community Building and Ecosystem Growth

- Onboarding new participants using PDC Tenant Manager provisioning
- Communication and knowledge-sharing channels
- Encouraging collaboration and co-creation
- Scaling across domains, sectors, and countries
- Long-term sustainability and evolution

Available Training Materials

The following training presentations have been developed and released as part of this curriculum:

Code	Training	Version	Slides	Primary Audience
ES-01	Project Manager Training	v10.5 (2026-07)	20	PM / Business
ES-02	Developer Training	v10.5 (2026-07)	34	Software / DevOps
ES-03	Data Scientist Training	v10.5 (2026-07)	31	Data / Analytics

The Developer and Data Scientist trainings share a common foundation (slides 1–13); the PM Training covers the same foundational concepts in slides 1–8 of its 20-slide deck. All three trainings then diverge into role-specific deep content. The shared foundation covers data space definition, the Prometheus-X/EDGE-Skills architecture, critical distinctions (onboarding vs. deployment), service chain thinking, four essential elements, and the training reference case.

References

- Resources from EDGE-Skills project
- Resources from ARIANE project
- prometheus-x.org: concepts, architecture, learning paths
- VisionsTrust Platform: operational platform for catalog, governance, compliance, data exchange
- Prometheus-X Building Blocks: Connector (PDC), Identity Hub, Data Plane SDKs
- VisionsTrust: SaaS reference implementation for discovery, contracts, consent, and data exchange
- Trust Frameworks (Gaia-X, EBSI): authorization, authentication, identification

Glossary

AI Act: EU Artificial Intelligence Act: EU Regulation 2024/1689, the world's first comprehensive legal framework for AI, applying a risk-based approach to AI systems.

API: Application Programming Interface: A defined interface through which software components communicate.

B2B: Business to Business: Data sharing between two organizations.

B2C: Business to Consumer: Services or data exchanges directed at individual end users.

B2G: Business to Government: Data sharing from a private company to a public authority.

CI/CD: Continuous Integration / Continuous Deployment: A DevOps practice of automating the building, testing, and deployment of software.

CSRD: Corporate Sustainability Reporting Directive: EU Directive 2022/2464, which requires large companies to report on environmental and social sustainability impacts.

DCAT: Data Catalog Vocabulary: A W3C vocabulary (v3) for describing datasets and data services in catalogues, enabling interoperability between data catalogues. In Prometheus-X, catalogue resources are described using DCAT v3 and serialized as JSON-LD.

DCP: Decentralized Claims Protocol: An Eclipse Foundation specification that defines how data space participants communicate identities and credentials to secure data access. DCP is an interoperable overlay to DSP for conveying organizational identities and establishing trust while preserving privacy.

DGA: Data Governance Act: EU Regulation 2022/868, in force since September 2023, which establishes a framework for trusted data intermediaries, reuse of protected public-sector data, data altruism organizations, and governance mechanisms for European data spaces.

DID: Decentralized Identifier: A W3C-standard identifier that enables verifiable, self-sovereign digital identities without relying on a central authority. In the curriculum, DIDs are created and managed by the Identity Hub. The did:web method is specifically referenced.

DPCP: Data Processing Chain Protocol: A Prometheus-X protocol and Node.js framework for orchestrating distributed data processing workflows. DPCP allows users to create and manage a hierarchical system of chains, nodes, and pipelines, enabling scalable and modular integration with external services.

DPP: Digital Product Passport: A standardized digital record containing information about a product's lifecycle, materials, and sustainability data, required under emerging EU regulations.

DS: Data Scientist: Curriculum shorthand for the Data Scientists / Data Engineers / Semantic Specialists learner role and their training materials.

DSP: Dataspace Protocol: The IDSA-defined open protocol governing how data space connectors communicate with each other. It covers catalogue discovery, contract negotiation, and transfer coordination. In Prometheus-X, DSP is implemented by the PDC as an interoperability protocol and is extended to support personal data, consent, and multi-party contracts.

DSSC: Data Spaces Support Centre: An EU-funded initiative that provides a framework, toolbox, and reference architecture for European data spaces, maintained by a consortium including IDSA, Gaia-X, and others.

E2E: End-to-End: Refers to validation or testing that covers the full flow from initial discovery through contract negotiation, transfer initiation, data receipt, and audit record generation.

EBSI: European Blockchain Services Infrastructure: A European Commission initiative providing blockchain-based infrastructure for cross-border public services, including authentication, identification, and authorization.

EDR: Endpoint Data Reference: In EDC-interoperable data transfers, a token issued by a provider's Control Plane to a consumer after successful contract negotiation and transfer initiation. The EDR grants temporary, scoped access to data at a specific endpoint by providing the endpoint URL of the provider's Data Plane and an authorization token.

ESCO: European Skills, Competences, Qualifications and Occupations: A multilingual EU classification framework maintained by the European Commission that maps approximately 13,890 skills/competences and 3,008 occupations in a structured, machine-readable taxonomy.

EUDR: EU Deforestation Regulation: EU Regulation 2023/1115, requiring companies to ensure products sold in the EU have not contributed to deforestation.

GDPR: General Data Protection Regulation: EU Regulation 2016/679, the foundational European data protection law. Key obligations include lawful basis for processing, consent management, data minimization, purpose limitation, data subject rights, and accountability.

IDSA: International Data Spaces Association: The non-profit organization that defines the IDS Reference Architecture Model, the Dataspace Protocol (DSP), and related standards for secure, sovereign data exchange in data spaces.

IoT: Internet of Things: Networked physical devices that collect and transmit data.

ISO/IEC 20151: An International Standard (FDIS stage, publication expected H2 2026) that provides a formal, globally recognised definition of a data space and its essential characteristics. Full title: "Information technology – Cloud computing and distributed platforms – Dataspace concepts and characteristics."

ISO/IEC 27560: An ISO/IEC standard for consent records in ICT systems (published 2023), providing a framework for documenting, exchanging, and managing consent information.

JSON-LD: JSON for Linked Data: A W3C standard method for encoding Linked Data using JSON, adding a @context to make JSON documents machine-interpretable as semantic graphs.

KPI: Key Performance Indicator: A measurable metric used to evaluate success against objectives.

LkSG: Lieferkettensorgfaltspflichtengesetz: German Supply Chain Due Diligence Act (effective 2023) requiring companies to conduct due diligence on human rights and environmental risks in their supply chains.

ML: Machine Learning: A subset of AI in which systems learn from data to improve performance on tasks.

NATS: Neural Autonomic Transport System: A high-performance open-source messaging and event-streaming system used for distributed systems communication.

OAuth2: Open Authorization 2.0: An open authorization framework that enables secure, token-based delegated access.

ODRL: Open Digital Rights Language: A W3C-standard policy expression language used to describe permissions, prohibitions, and obligations over data assets. In the Prometheus-X stack, ODRL policies are attached to Assets at publication time and evaluated at runtime by the Connector's policy engine.

OEM: Original Equipment Manufacturer: A company that produces components or products used in another company's end product.

PDC: Prometheus-X Dataspace Connector: The open-source software agent that represents a participant in the Prometheus-X data space ecosystem. The PDC implements DSP as an interoperability protocol and extends it for personal data, consent, and multi-party contracts.

PDC Tenant Manager: A component of the Prometheus-X open source stack that enables multi-tenant deployment of PDC instances at scale. A tenant is an isolated participant context provisioned within a shared deployment.

PDI: Personal Data Intermediary: A Prometheus-X core service that enables consent-driven personal data sharing. The PDI acts as a trusted intermediary: individuals manage their consents from a single place; the PDI does not store the personal data itself (Separation of Powers Principle). The PDI is Prometheus-X's most distinctive component.

PM: Project Manager: Curriculum shorthand for the Project Managers / Business Ecosystem Leaders learner role.

RDF: Resource Description Framework: A W3C standard framework for representing information as subject-predicate-object triples, forming the foundation of knowledge graphs and the Semantic Web.

ROME: Répertoire Opérationnel des Métiers et des Emplois: A French occupational classification framework maintained by France Travail, describing approximately 531 job groups with associated skills and competences.

SDK: Software Development Kit: A set of libraries, tools, and APIs that enable developers to build specific types of applications.

Service Chain Protocol: See DPCP (Data Processing Chain Protocol). The slide decks use the term "Service Chain Protocol" as an alternative name for DPCP.

SLA: Service Level Agreement: A contract that defines the expected level of service between a provider and consumer, including availability, performance, and incident response targets.

VC: Verifiable Credential: A W3C-standard digital credential that can be cryptographically verified. VCs allow participants to prove membership, attributes, or roles without a central authority. In the Prometheus-X stack, VCs are stored in the Identity Hub.

xAPI: Experience API: A learning technology specification (IEEE 9274.1.1) that defines a standard format for recording and sharing learning experiences as statements.