

Deliverable D3.3

Proposal for updates of major industrial data models

Work package
leader:

Nils Sandsmark
nils.sandsmark@posccaesar.org

Task
leader:

Géza Kulcsár
gkulcsar@ptc.com

Editor:

Péter Olszi
peter.olaszi@aitia.ai

Abstract

This document is Deliverable 3.3 of the Arrowhead FPVN project, the final report of Work Package 3. It presents concrete proposals for updating major industrial data model standards – including ISO 23726 (IDO), ISO 10303 (STEP), ISO 15926, DEXPI, CFIHOS, SysML v2, IEC 63278 (AAS), and DPP – to support ontology-based interoperability, validated through Use Cases 2.9 and 3.9.

Grant agreement no.	101111977
Project acronym	Arrowhead fPVN
Project full title	Arrowhead flexible Production Value Network
Dissemination level	PU
Due Date	31-05-2026
Date of Delivery	31-05-2026
Deliverable Number	D3.3
Deliverable Name	Proposal for updates of major industrial data models
AL / Task related	Task 3.3
Author/s	Péter Olszi (AITIA)
Contributors	Arrowhead fPVN WP3 partners
Reviewer	Jerker Delsing
Keywords	Industrial data models, semantics, ontology-based interoperability,
Abstract	This document is Deliverable 3.3 of the Arrowhead fPVN project, the final report of Work Package 3. It presents concrete proposals for updating major industrial data model standards – including ISO 23726 (IDO), ISO 10303 (STEP), ISO 15926, DEXPI, CFIHOS, SysML v2, IEC 63278 (AAS), and DPP – to support ontology-based interoperability, validated through Use Cases 2.9 and 3.9.

Table of contents

1	Introduction	5
2	Synergy	7
3	Enabling Semantic Interoperability Across Engineering Standards	7
3.1	Standards in the WP3 work program and tested in use-cases	9
3.1.1	ISO 23726: Ontology-based interoperability (OBI)	9
3.1.2	ISO 10303: Product data representation and exchange (STEP)	18
3.1.3	ISO 15926: Integration of life-cycle data for process plants including oil and gas production facilities.....	24
3.1.4	DEXPI: Data Exchange in the Process Industry	28
3.1.5	CFIHOS: Capital Facilities Information Handover Specification	32
3.1.6	SysML: Systems Modeling Language.....	38
3.2	Standards in the WP3 work program but not tested in use-cases	43
3.2.1	IEC 63278: Asset Administration Shell	43
3.2.2	DPP: Digital Product Passport	46
3.3	Standards WP3 originally not included in the work program	50
3.3.1	AFO: Arrowhead Framework Ontology	50
3.3.2	IMF: Information Modelling Framework	55
4	Arrowhead fPVN use cases.....	58
4.1	Use Case 2.9: Digital Twin	58
4.2	Use Case 3.9: Process industry	59
4.2.1	Mapping and alignment of Balas model elements	59
4.2.2	Mapping of streams, ports and material components	62
4.2.3	Mapping to Chemical Entities of Biological Interest identifiers	62
4.2.4	The applicability of the PLM-RDL2 and DEXPI Process standards.....	62
5	Discussion.....	64
5.1	The alignment architecture: strengths and limits	64
5.2	Maturity varies across standards	64
5.3	The role of tooling	65
5.4	Semantic web technologies in an industrial context.....	65
5.5	Standardization engagement.....	65
5.6	Toward industrial deployment.....	66
6	Conclusion.....	66
7	References.....	68
8	Revision history	71
8.1	Contributing and reviewing partners	71



8.2	Amendments	71
8.3	Quality assurance	71

1 Introduction

Work Package 3 of the Arrowhead flexible Production Value Network (fPVN) project is responsible for defining and selecting data model languages, providing guidelines and tools for the use cases. The WP3 deliverable series spans three major reports that progressively build on each other.

Deliverable D3.1 [31] surveyed and catalogued the major standardized industrial data models relevant to the project's use cases. It identified the foundational properties, similarities, and dissimilarities of those standards – covering domains as diverse as process plant engineering (ISO 15926, DEXPI), product lifecycle management (ISO 10303), automation and control (International Electrotechnical Commission (IEC) 61131-3, IEC 61499), and energy systems (IEC 61850). That work established a shared baseline across the WP3 partners and the use case leaders.

Deliverable D3.2 [6] advanced the analysis by exploring how semantic web technologies can extend and bridge these standards. It demonstrated that ISO 23726-3 – the Industrial Data Ontology (IDO) – can serve as a shared upper ontology to which the other major standards can be aligned, and validated this through concrete ontology alignment work and the Arrowhead fPVN Use Cases 2.9 (Digital Twin) and 3.9 (Process Industry). The key finding was that ontology-based interoperability across heterogeneous industrial standards is technically feasible and delivers measurable value.

The present document, D3.3, is the third and final deliverable of WP3. Its title – *Proposal for updates of major industrial data models* – signals the shift in emphasis: from analysis to recommendation. D3.3 builds on the semantic foundations established in D3.2 and formulates concrete proposals for how the major industrial data model standards could be updated or extended to support the kind of interoperability demonstrated in the project. It also reflects on the maturity, limitations, and future evolution of each standard from this perspective, and documents the project's direct engagement with the relevant standards bodies.

Task 3.3 focuses on the use of a heterogeneous data model in the microservices paradigm – a key element of the Arrowhead framework – including both solution engineering and its maintenance and evolution. The microservices dimension is addressed concretely in Section 3.3.1, which presents the Arrowhead Framework Ontology (AFO) – an OWL 2 DL representation of the Arrowhead microservices architecture itself – and its validation in Use Case 2.9 (Section 4.1). The document examines how standards can be expressed as W3C (World Wide Web Consortium)-compatible OWL 2 DL (Web Ontology Language 2, Description Logic profile) ontologies and integrated into a shared, multi-knowledge-graph architecture; what alignment patterns and modeling guidelines have been developed, validated, and contributed back to standards organizations; and what lessons the validated use cases offer for future standardization and industrial deployment.

The following standards form the core of WP3's technical work in D3.3 and have been directly validated through the project's use cases:

- ISO 23726: Ontology-Based Interoperability (OBI), including the Industrial Data Ontology (IDO) as upper ontology;
- ISO 10303: Product data representation and exchange (STEP);
- ISO 15926: Integration of life-cycle data for process plants including oil and gas production facilities;
- DEXPI: Data Exchange in the Process Industry;
- CFIHOS: Capital Facilities Information Handover Specification;
- SysML v2: Systems Modeling Language.

In addition, two further standards in the WP3 work program – IEC 63278 (Asset Administration Shell) and the Digital Product Passport (DPP) – are covered but have not yet been directly tested in use cases. Two additional frameworks that were not part of the original WP3 baseline are also presented: the Arrowhead Framework Ontology (AFO), which was later incorporated into WP3 and validated in Use Case 2.9, and the Information Modelling Framework (IMF), which remained at study level during the project period.

The project's engagement extended beyond standards bodies to two closely related industry initiatives that became collaboration partners. Digitalization – Industrialization – Standardization – Collaboration (DISC) is a collaboration between the four major Norwegian companies in the Norwegian oil and gas industry: Aibel, AkerBP, AkerSolutions, and Equinor. The Digitalization and Information Standards Subcommittee (DISC) is a sub-committee of the Standards Committee of the International Association of Oil & Gas Producers (IOGP), formed in 2015. Its mission is to drive the availability, maturity, and utilization of preferred information standards and digitalization solutions across the oil and gas industry. Among the standards actively supported by DISC are DEXPI, CFIHOS, Joint Industry Programme (JIP) 33, and ISO 15926. CFIHOS Semantics is an IOGP initiative – operating within the JIP 36 (CFIHOS) framework – that develops semantic models, ontologies, and data structures compliant with IDO to augment the CFIHOS standard and make it applicable in semantic and knowledge-graph-based environments. Both initiatives share the goal of using IDO as a common upper ontology for aligning oil and gas industry data standards, which made them natural partners for Arrowhead fPVN WP3 throughout the project period.

The document is organized as follows. Section 2 summarizes the synergy activities and collaborative events of the reporting period. Section 3 forms the technical core, providing a structured analysis of each standard – including scope, the project's engagement with the relevant standards organization, methods applied, results achieved, and proposed updates or extensions. Section 4 presents the use case demonstrations. Section 5 discusses cross-cutting findings, and Section 6 concludes the deliverable.

2 Synergy

Building on the foundations established in the previous period, collaboration with DISC and CFIHOS Semantics continued through regular joint meetings. The cooperation deepened around the shared goal of using IDO as an upper ontology, with CFIHOS Semantics contributing property mapping work and alignment patterns, and DISC providing input on lifecycle modeling and STEP integration.

Regular weekly online meetings continued throughout the second half of 2025 and into 2026, progressively advancing the technical depth of the IDO guidelines. The focus shifted from conceptual alignment to concrete modeling patterns, covering property classification, lifecycle stage modeling, identifier patterns, and reference data integration. A decision log was established in June 2025 to track topics, discussions and decisions.

We had technical breakthroughs, too: the IDO core group's agreement on a new extension pattern distinguishing intrinsic properties from assigned attributes, the adoption of QUDT (Quantities, Units, Dimensions and Types) as the primary ontology for units of measurement, and the definition of five property types to be modeled in the domain ontology. The introduced pump lifecycle story served as the principal validation use case throughout this period.

In the final months of the project, significant effort was invested in establishing a sustainable documentation infrastructure. The team agreed on producing a set of documents titled "The Arrowhead fPVN Handbook on industrial semantics – Using IDO as upper ontology" [39]. The title was chosen deliberately. During development the document set had been informally referred to as the *IDO Handbook*; however, the group recognized that their work – while firmly grounded in IDO – had not achieved full alignment with every aspect of the specification, making that label an overstatement. The adopted title instead positions the output as an Arrowhead fPVN project deliverable that *uses* IDO as its upper ontology. The complementary artifacts include the Guidelines (as a brief overview), the Handbook (the comprehensive reference); the Patterns (a reusable pattern library for data modelers); and the Use Cases. AsciiDoc was selected as the authoring format to enable version control and multi-format publication. A dedicated GitHub repository was created under the POSC Caesar Association (PCA) to host the handbook, ensuring long-term maintenance beyond the project's lifetime [38]. PCA was identified as the maintenance agency for the resulting guidelines.

3 Enabling Semantic Interoperability Across Engineering Standards

Despite the availability of information model standards in industry since the mid-1990s, and the high expectations from standardization organizations, seamless interoperability has not been achieved. This gap is mainly due to several structural challenges:

- Different vocabularies are used across standards, even within the same standard family;
- Diverging model architectures and different modelling languages have been applied;
- Standardization groups have often prioritized the success of their own standards over broader business needs;
- Newly developed standards often duplicate existing concepts, but in different ways.

The Arrowhead fPVN initiative proposes a fundamentally different approach based on World Wide Web (W3C) standards and semantic technologies (Figure 1).

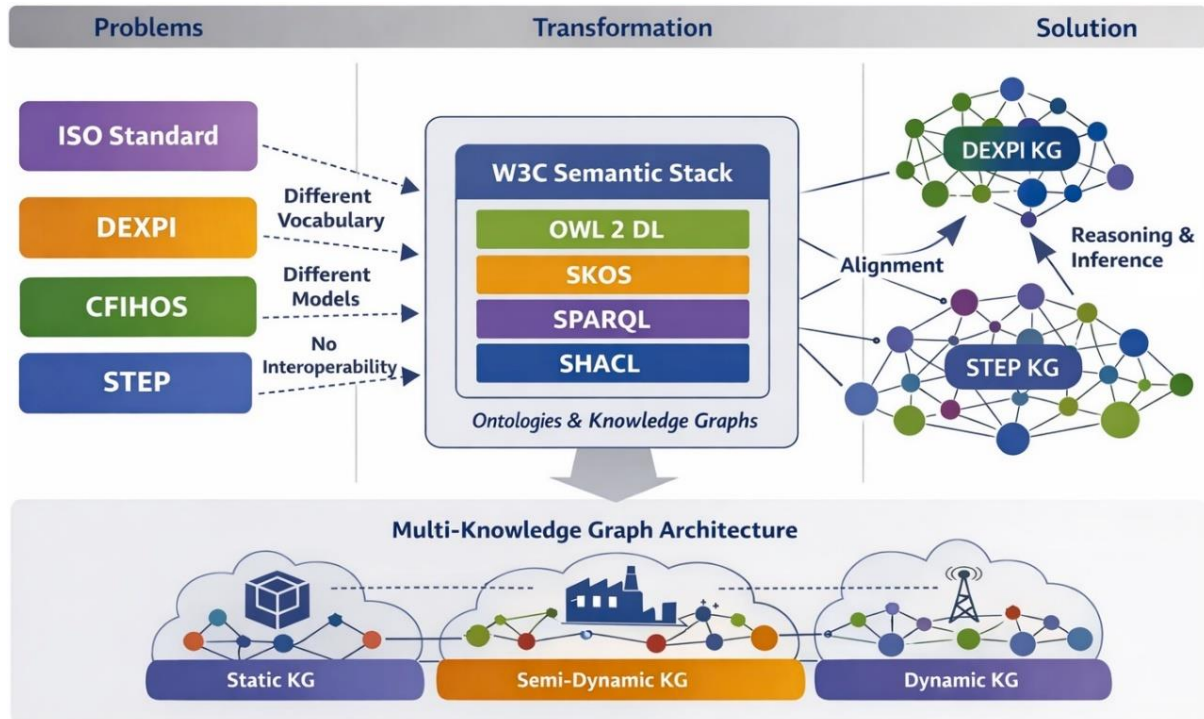


Figure 1. Semantic interoperability via knowledge graph (KG) alignment

All engineering standards are transformed into a shared semantic transformation framework, based on W3C ontologies and knowledge graphs: OWL 2 DL (a formal ontology representation), SKOS (with controlled vocabularies and concept schemes), SPARQL (as query language), and SHACL (for data validation). This eliminates differences among modelling languages and technology stacks, and paves the way for a unified semantic foundation. These transformations preserve the original semantic meaning of each standard, and enable reasoning and inference, in turn making previously implicit (“hidden”) knowledge explicit.

Rather than creating a single, fully integrated data model, the approach uses multiple knowledge graphs, each representing a standard or domain. The semantic alignment between graphs allows faster realization of business value and reduces the need for large and costly mapping projects. It also improves incrementally as more alignments are added. The flexible multi-knowledge graph architecture supports different types of knowledge graphs: static graphs (e.g., standards and reference models), semi-dynamic graphs (e.g., engineering design data), and dynamic graphs (e.g., operational and real-time data). These coexist and interact within a flexible and extensible ecosystem. The result is scalable interoperability, continuous evolution of systems, and integration across lifecycle phases.

By leveraging W3C semantic web technologies and alignment-based integration, it becomes possible to finally achieve the level of interoperability that traditional standards alone have not delivered – while preserving their value and extending their capabilities.

3.1 Standards in the WP3 work program and tested in use-cases

The following subsections describe the standards that formed the core of WP3's technical work program and were directly validated through the project's use cases. Each standard is presented according to a common structure: scope, the project's engagement with the relevant standards organization, methods applied, results achieved, and proposed further work. The standards covered are ISO 23726 (IDO), ISO 10303 (STEP), ISO 15926, DEXPI, CFIHOS, and SysML v2. Although these standards originate from different communities, serve different lifecycle phases, and employ different modeling formalisms, they share a common role in the project: each was transformed into an OWL 2 DL representation and aligned with IDO as the shared upper ontology, enabling cross-standard queries and reasoning within a unified knowledge graph architecture. The degree of alignment and the maturity of results vary across standards, reflecting the different starting points and available resources in each case. Together, however, they constitute a coherent and operationally validated body of work demonstrating that ontology-based interoperability across heterogeneous industrial standards is technically feasible and delivers measurable value.

3.1.1 ISO 23726: Ontology-based interoperability (OBI)

ISO 23726, titled *Industrial automation systems and integration – Ontology-based interoperability (OBI)*, is a multipart standard developed by ISO/TC 184/SC 4 (Industrial Data). Its foundation is the Industrial Data Ontology (IDO), an OWL 2 DL upper ontology that makes the meaning of industrial data explicit and machine-interpretable across engineering disciplines, industries, and lifecycle phases. IDO addresses a fundamental problem in industrial systems integration: as the number of systems exchanging data grows, point-to-point integrations multiply combinatorically, and meaning is lost because it is encoded implicitly in software, documentation, or institutional knowledge rather than in the data itself.

IDO resolves this by providing a shared semantic center – a star topology to which all systems align – so that any new system needs to be integrated once rather than connected to every other system individually. Arrowhead fPVN was named as a supporting project when the New Project (NP) proposal was submitted to ISO in 2022, and the standard reached Final Draft International Standard (FDIS) status in April 2026.

In the Arrowhead fPVN WP3 context, ISO 23726 serves as the semantic foundation for all data model interoperability work. IDO is used as the upper ontology to which the other major standards analyzed in this document – ISO 10303 (STEP), ISO 15926, DEXPI, and CFIHOS Semantics – are aligned. The complementarity between IDO and SysML v2 as a design-time engineering language is analyzed in Section 3.1.6.

3.1.1.1 Scope of the standard

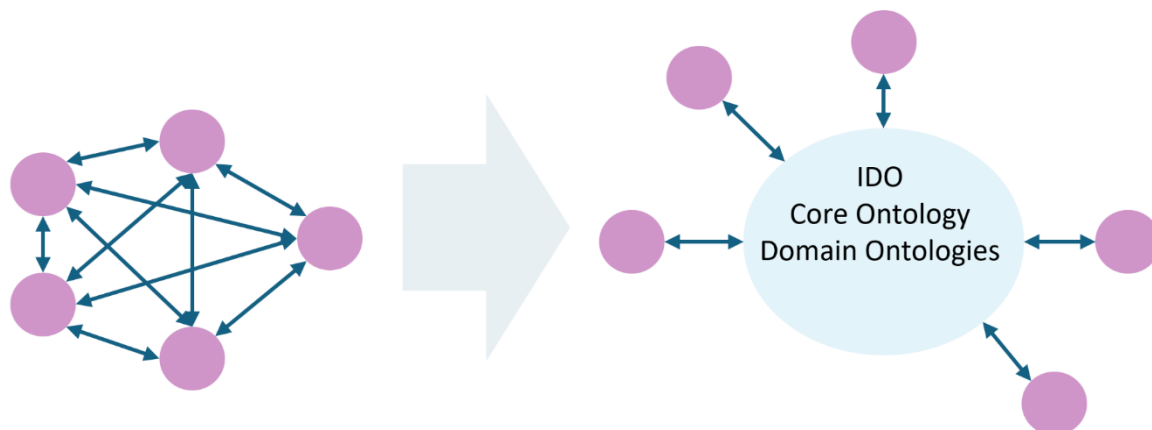
ISO 23726 is a multipart standard family under ISO/TC 184/SC 4 [20]. Its parts and their roles in the OBI ecosystem are summarized in Table 1.

Table 1. Parts of the ISO 23726 series

Part	Title	Role
ISO 23726-1	Overview and fundamental principles	Defines the vision for the OBI ecosystem and the compliance principles for all artefacts in the series. Expected publication 2028.
ISO 23726-2	Vocabulary	Defines terms and definitions used throughout the standard family.
ISO 23726-3	Industrial Data Ontology (IDO)	The upper ontology; the semantic foundation of the entire series. DIS approved December 2025; FDIS v4.2 finalized April 2026.
ISO 23726-100	Schedule Data Ontology	The first domain ontology built on IDO; provides vocabulary for schedule data and work planning.

ISO 23726-3, the Industrial Data Ontology, specifies an OWL 2 ontology for industrial data and information. Its normative scope is: (1) general overview of IDO and its main principles; (2) normative definitions of IDO classes and relations; (3) informative examples of the use of IDO terminology in context; and (4) reusable representation patterns. The following are explicitly outside the scope of ISO 23726-3: specification of data modelling languages (such as RDF or OWL), and specification of methods for reasoning with ontologies. IDO is intended for use in all lifecycle phases of industrial assets and processes, from design and procurement through construction, operation, maintenance, and disposal.

The motivation for the standard originates in a long-standing integration problem in industrial systems. As the number of systems exchanging data grows, the number of required integrations grows combinatorically. Each integration is effectively a mapping, but a mapping is not the same as understanding: it captures syntax, not semantics. The root cause is that meaning is hard-coded in software, locked in documentation, or buried in institutional knowledge – in other words, implicit. Systems can exchange identical data and still interpret it differently. IDO resolves this by making meaning explicit: rather than connecting every system to every other system, all systems align to a shared semantic center (a star topology). Any new system needs only a single alignment to this center to become interoperable with all others (Figure 2).



Point-to-point integration

Ontology-based integration

*Figure 2. Point-to-point integration (left) versus IDO star-topology alignment (right)
Source: OBI - IDO training sessions 2026 [22]. Reproduced with permission.*

IDO provides this center through three pairwise disjoint top-level categories that jointly cover all industrial entities:

- **Object:** Independent entities that persist through time (e.g., physical objects, locations, information, organizations);
- **Dependent:** Inherent qualities or realizable potentials of an object (e.g., a pump's potential to increase pressure);
- **Temporal:** Entities relating to time (e.g., activities, temporal regions).

Table 2. IDO standardization timeline

Year	Milestone
2008	Development started in the IOHN project
2012–2016	Refined in the EU-funded Optique project; first large industrial implementations (Aibel, Aker Solutions, others)
2017	Proposed as ISO 15926 Part 14 at ISO/TC 184/SC 4 meeting in Oslo
2018	ISO/DTR 15926-14 approved as Technical Report
2020–2022	READI (Requirement Asset Digital lifecycle Information) JIP proposes an IS (not TR) under the READI project; NP submitted November 2022
2023	NP re-submitted as ISO 23726-3 to ISO/TC 184/SC 4; project approved. Arrowhead fPVN named as supporting project.
December 2024	Draft International Standard (DIS) v4.1 submitted for ballot
December 2025	DIS ballot approved — IDO becomes an official ISO standard
April 2026	Final Draft International Standard (FDIS) v4.2 finalized (50 classes, 102 object properties)

These categories are mutually exclusive: any entity belongs to exactly one. Cross-category relations (such as ``lis:hasPotential``, ``lis:hasQuality``, ``lis:participatesIn``, and ``lis:realizedIn``) bridge the categories and give IDO its modeling power.

IDO originated in 2008 in the IOHN (Integrated Operations in the High North) joint industry project and was first formalized as ISO/DTR 15926-14. Development and standardization history is summarized in Table 2.

The compliance framework governing all artefacts in the OBI series is defined in ISO 23726-1. Four principles apply to every OBI-conformant ontology:

1. Direct Semantics: The ontology must have a consistent interpretation under OWL 2 Direct Semantics, verifiable by a standard DL reasoner (e.g., HermiT or ELK).
2. RDF conformance: The ontology must be serializable and publishable as RDF, enabling web-scale distribution.
3. Digital resource requirement: Every OBI artefact must be identified by a dereferenceable Internationalized Resource Identifier (IRI). IDO is maintained by the POSC Caesar Association (PCA) [21].
4. Alignment rule: Every class in a domain or reference ontology claiming OBI conformance must be declared a direct or transitive subclass of an IDO class. This rule is what makes IDO the common semantic spine – two OBI-conformant ontologies that have never been explicitly mapped to each other are nonetheless jointly queryable, because both are grounded in the same upper layer.

These principles have been applied throughout the Arrowhead fPVN WP3 work: every aligned standard (STEP, ISO 15926, DEXPI, CFIHOS Semantics) produced an alignment ontology that satisfies the alignment rule, and all knowledge graphs are published in RDF.

3.1.1.2 Our work and cooperation with the standards organization

Arrowhead fPVN has been a named supporting project of the ISO 23726 standardization effort since the New Project proposal was submitted to ISO/TC 184/SC 4 in 2022 and approved in 2023. This relationship is documented in the NP submission itself, which describes Arrowhead fPVN as one of the R&D projects and standardization collaborations that motivated and validated the need for an IS-level upper ontology for industrial data.

Active technical collaboration with the IDO core development team has been ongoing throughout the second half of the project (2025–2026). The primary collaboration forums were regular weekly online meetings involving members of the IDO core group, DISC, and the CFIHOS Semantics team. These meetings progressively advanced the technical depth of IDO alignment guidelines, covering property classification, lifecycle stage modeling, identifier patterns, and reference data integration. A joint in-person meeting in Oslo (June 9–11, 2026) is expected to bring the working groups together to review and consolidate results across all three collaborating initiatives.

On April 30, 2026, WP3 hosted an IDO Training session [22] presented by the IDO core team for the broader Arrowhead fPVN WP3 community. The session covered the structure and content of IDO, the status of the ISO 23726-3 standardization process, and the technical changes

introduced in FDIS v4.2 relative to the DIS ballot version. This provided the project with up-to-date, first-hand knowledge of the standard immediately before the submission of D3.3.

At the SC 4 level, the project also engaged with the broader ISO/TC 184/SC 4 community. The meeting with SC 4 chair Kenneth (Kenny) Swope of Boeing in the summer of 2024, referenced in Section 3.1.2, identified ISO 23726 (IDO) as a candidate normative reference point for shared semantics across the SC 4 standards family. WP3 results related to ontology-based interoperability were also presented to the SC 4 community during the meeting week in Nagasaki (26–31 October 2025). Further details on the project's formal cooperation with the ISO 23726 standardization body are provided in [21] and [23].

3.1.1.3 Methods used

The central methodological commitment of WP3 is the adoption of IDO as the upper ontology for all semantic interoperability work. This means that every domain or reference ontology developed or aligned within the project – covering STEP, ISO 15926, DEXPI, CFIHOS Semantics, and others – is required to place each of its classes as a direct or transitive subclass of an IDO class. This alignment pattern does not require modifying the normative content of the aligned standards; it only adds subclass declarations in a separate alignment ontology file. The effect is that all aligned knowledge graphs share a common semantic foundation, enabling cross-standard queries and reasoning without a central integration model.

OWL 2 DL was adopted as the common formalism throughout. All standard-to-ontology transformations (ISO 10303, DEXPI, ISO 15926) produce OWL 2 DL artifacts, and all instance data in the project's knowledge graphs is represented in RDF. The choice of OWL 2 DL is deliberate: it is the sublanguage of OWL 2 that guarantees decidable automated reasoning while remaining expressive enough for industrial use cases. This enables automated consistency checking, inference of implicit subclass hierarchies, and derivation of new facts – tasks performed by standard OWL DL reasoners (HermiT, Pellet, ELK) without human intervention.

A key methodological tool is the use of ontology modeling patterns. A pattern is a reusable solution to a recurring modeling problem, analogous to design patterns in software engineering. The project validated its IDO alignment work in a canonical validation use case: the pump lifecycle story, which traces a pump from its functional specification through procurement, installation, operation, and maintenance, including the representation of its qualities (pressure, temperature), potentials (the function to increase pressure), and the activities it participates in. This story was deliberately chosen as a cross-cutting example that exercises all major IDO categories and relations, and served as the reference model for evaluating whether modeling decisions made during the collaboration were consistent and operationally sound.

Figure 3 illustrates how the pump lifecycle story exercises all three IDO categories and the modular layering strategy simultaneously. The diagram shows three horizontal layers: the IDO upper ontology (Object, Dependent, Temporal) at the top; a core/mid-level layer of shared concepts (functional object, physical object, activity, quality) in the middle; and application- or project-specific subclasses at the bottom. The pump, its pressure-increase potential, and the commissioning activity are each anchored to their respective IDO category, while the cross-category relations (`lis:hasPotential`, `lis:realizedIn`, `lis:participatesIn`) connect them. Domain ontologies derived from STEP, DEXPI, and CFIHOS Semantics slot into the

mid and bottom layers, each placing their classes as subclasses of the appropriate IDO class. This layered architecture is the mechanism that makes cross-standard queries possible at runtime without a central integration model.

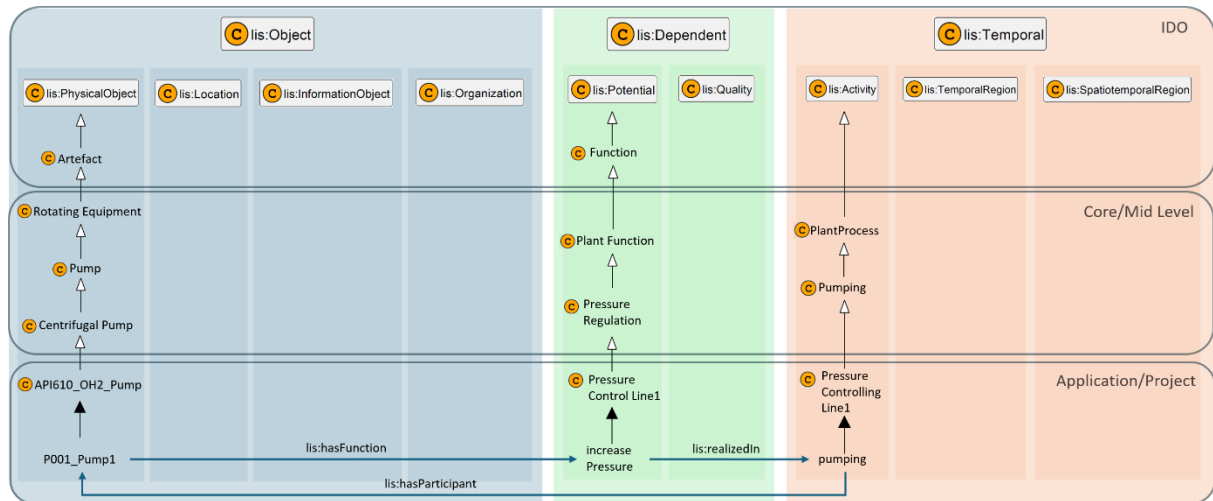


Figure 3. Pump lifecycle story – layered modular ontology architecture across the three IDO categories.
Source: OBI - IDO training sessions 2026 [22]. Reproduced with permission.

For property modeling, five property type categories were defined for use in domain ontologies built on IDO: intrinsic physical quantities (e.g., mass, rated power), assigned attributes (e.g., tag identifiers, document references), lifecycle states (e.g., design, as-built, operational), temporal data (durations, timestamps, operational periods), and relational attributes (properties that are meaningful only in context, such as a pump's pressure at a specific nozzle). QUDT (Quantities, Units, Dimensions and Types) was adopted as the primary ontology for units of measurement, providing a well-maintained, machine-readable vocabulary for physical quantities that integrates cleanly with IDO's `lis:Quality` and `lis:PhysicalQuantity` classes.

SHACL (Shapes Constraint Language) was used as a complementary validation mechanism alongside OWL reasoning. While OWL 2 DL operates under the open-world assumption and is suited to inference, SHACL operates under the closed-world assumption and is suited to validation of specific data completeness and structural constraints. SHACL shapes were defined for the key modeling patterns to enforce, for example, that every `lis:PhysicalQuantity` individual carries a unit of measure, and that every IDO-aligned TAG has at least one associated functional object. These shapes can be applied at the knowledge graph level as quality gates, independent of the upstream standard.

The overall knowledge graph architecture used in the project supports three coexisting graph types: static graphs (representing standard ontologies and reference models, updated only when a standard is revised), semi-dynamic graphs (representing engineering design data, updated during project phases), and dynamic graphs (representing operational and real-time data). The IDO upper ontology acts as the common semantic spine across all three layers, making it possible to formulate queries that span engineering design data and operational history within the same semantic framework.

3.1.1.4 Results

ISO/DIS 23726-3 was submitted for ballot in December 2024 and approved in December 2025, making IDO an official ISO Draft International Standard. The FDIS (Final Draft International Standard), version 4.2, was finalized in April 2026 and comprises 50 classes and 102 object properties. This represents the first OWL 2 DL upper ontology for industrial data to achieve IS-track ballot approval within the ISO/TC 184/SC 4 family.

The DIS ballot attracted 182 national body comments. Of these, 142 were accepted and incorporated into FDIS v4.2, 20 were deferred to a subsequent edition, and 20 were rejected with documented rationale. The resolution process was technically substantive: FDIS v4.2 introduced 12 new entity types (including `lis:TemporalDatum`), modified 186 existing entities, and added the `lis:consumes`/`lis:consumedBy` and `lis:siteDeterminedBy`/`lis:siteInsideOf` relations. These additions directly resolved modeling gaps identified through the project's use-case validation work.

A comprehensive documentation set, titled "The Arrowhead fPVN Handbook on industrial semantics – Using IDO as upper ontology" was produced by the IDO core group with active participation from the project. The complementary artifacts include the Guidelines (as a brief overview); the Handbook (the comprehensive normative-style reference); the Pattern Library (a catalog of reusable modeling patterns); and the Use Cases (worked examples grounding the patterns in realistic industrial scenarios). All artifacts are authored in AsciiDoc to enable version control, diff-based review, and multi-format publication. They are hosted in a dedicated repository under the POSC Caesar organization on GitHub, with POSC Caesar Association (PCA) identified as the long-term maintenance agency.

Four major industrial standards were aligned with IDO during the project period, each producing a separate OWL 2 DL alignment ontology that declares the relevant classes of the aligned standard as subclasses of IDO classes: ISO 10303 STEP Core (see Section 3.1.2), ISO 15926-4 Ed. 4 (Section 3.1.3), DEXPI Plant (Section 3.1.4), and CFIHOS Semantics. None of these alignments modify the normative content of the aligned standards; they are maintained as standalone ontology files. The result is a shared semantic layer over which cross-standard SPARQL queries and OWL DL reasoning can be performed without a central integration model.

The IDO-based interoperability architecture was validated in two Arrowhead fPVN use cases. In UC 2.9 (Digital Twin for process plants), the combined STEP Core, DEXPI Plant, and ISO 15926 knowledge graphs – all aligned with IDO – were used to demonstrate cross-standard queries over engineering design and operational data within the same semantic framework. In UC 3.9 (Process Industry), the Balas process simulation model was mapped to IDO-aligned reference data, enabling integration of process simulation outputs with engineering knowledge graphs. Both use cases confirmed that the alignment architecture is operationally sound at the instance data level, not only at the schema level.

The IDO approach has already been adopted or evaluated by several industrial organizations independently of this project. Aibel used IDO-aligned knowledge graphs to detect duplicate material master data entries through OWL DL reasoning – a task that previously required manual reconciliation. Siemens Energy adopted IDO as the top-level ontology in their enterprise

knowledge management platform for industrial assets. TotalEnergies identified IDO as a candidate foundation for integrating digital technical data across engineering disciplines. Grundfos developed a large-scale product model using OWL reasoning grounded in IDO classes. Equinor, through its internal DISC programme, is building the Norwegian continental shelf digitalization platform on IDO as its semantic spine.



Figure 4. ISO 23726 industrial consortium (2022–2026)

The breadth of the industrial constituency behind the standard is illustrated in Figure 4. The consortium spans energy (AkerBP, Equinor, TotalEnergies), maritime (Aibel), manufacturing (Aker Solutions, Siemens, Grundfos), aerospace and automotive (SEIIA.SE), mining, academic institutions (University of Western Australia, University of South Australia, Universitetet i Oslo), and standardization bodies (Standards Norway, POSC Caesar Association, CFIHOS).

Beyond direct adoption of IDO, several organizations have used IDO as the upper ontology for newly developed domain-specific ontologies – the pattern defined by the OBI alignment rule. The most significant example is the Information Modelling Framework (IMF), developed by Equinor in collaboration with AkerBP, Aibel, and Aker Solutions as a domain ontology for functional system modeling on the Norwegian continental shelf (Section 3.3.2). IMF is grounded directly in IDO classes and represents the first large-scale industrial deployment of the OBI compliance pattern at enterprise level. Its adoption by multiple competing operators and contractors on the same infrastructure demonstrates that the alignment architecture is not only technically sound but also commercially viable in a multi-stakeholder setting.

An empirical result from UC 2.9 is that significant interoperability benefits can be achieved from a small number of class-level alignments between standards. Even with only a subset of STEP and DEXPI classes explicitly aligned to IDO, the knowledge graphs produced cross-standard query results that would have required a major point-to-point integration project under traditional approaches. This finding – consistent with the corresponding result reported for STEP in Section 3.1.2.4 – confirms that the star-topology alignment strategy delivers early and growing value: initial benefits are achievable with limited alignment effort, and the value increases incrementally as more classes are aligned over time.

3.1.1.5 Proposed further work

The DIS ballot process surfaced a set of structural tensions in the IDO model that were examined carefully during the resolution of national body comments. Three issues are worth noting as open design questions that remain relevant for implementors and for the next edition.

The first concerns the modeling of streams in process industry contexts. A stream (e.g., a flow of material between two equipment items) is representable in IDO as a sequence of connected functional objects, but the DIS text does not provide explicit normative guidance on how stream identity is preserved across lifecycle changes. This was raised during ballot and deferred to a future edition.

The second concerns the `lis:Feature` class. As defined in the DIS, features (e.g., a bore, a flange face) are modeled as physical objects with spatial extent. One ballot comment observed that this forces every feature to have a mass, which is ontologically inappropriate for purely geometrical features such as holes. This tension between the physical-object-centric ontology and engineering geometry modeling has been deferred.

The third concerns the `lis:ActivityProfile` class, which was observed to be misclassified under the *Temporal* category when it arguably describes a pattern or specification of activity (a *Dependent* or *Object* entity) rather than a temporal entity itself. This issue was accepted in principle but the resolution was deferred pending a broader review of the *Temporal* category hierarchy.

FDIS v4.2 resolved the most operationally significant DIS issues. The introduction of `lis:TemporalDatum` provides a clean class for timestamps and reference time instants, filling a gap that had forced implementors to use workarounds. The `lis:consumes`/`lis:consumedBy` relations now formally support material and energy flow modeling. The `lis:siteDeterminedBy`/`lis:siteInsideOf` relations support spatial containment, which is important for facility and location modeling. Several informative annexes covering alignment mappings to other standards were moved to online resources to allow faster updating outside the ISO ballot cycle.

The four alignments completed in this project (STEP, ISO 15926, DEXPI, CFIHOS Semantics) cover the standards directly in scope for the Arrowhead fPVN use cases. The alignment architecture is designed to scale: additional standards – including IFC (ISO 16739) for the built environment, IEC CDD (Common Data Dictionary) for components, and AAS submodels for Industry 4.0 assets – can be onboarded using the same pattern. Each new alignment requires only a separate alignment ontology file; the IDO upper ontology and existing alignments are unaffected. Extending the alignment ecosystem to cover these standards is recommended as a next step, building directly on the governance approach and the Arrowhead fPVN Handbook developed during this project.

The Arrowhead fPVN Handbook, Patterns, Use Cases, and Guidelines artifacts produced under this project provide the documentation infrastructure for IDO adoption, but governance mechanisms for the long-term maintenance of alignment ontologies remain to be formalized. As IDO editions are revised (a second edition addressing the deferred DIS issues is anticipated), alignment files for each aligned standard will need versioned updates. A recommended further step

is the definition of a governance framework – covering versioning policy, change impact assessment, and release coordination – for the family of IDO alignment ontologies. This could be contributed to ISO/TC 184/SC 4 or maintained by PCA as part of its ongoing IDO stewardship role.

The validation performed in UC 2.9 and UC 3.9 demonstrated correctness of the alignment architecture and the feasibility of cross-standard queries, but both use cases operated at a scale appropriate for a research project. Industrial adoption at full scale – covering complete plant models with tens of thousands of tagged objects, complete lifecycle histories, and live operational data – requires toolchain integration work beyond the scope of this deliverable. This includes integration of IDO-aligned ontologies into mainstream engineering tools (e.g., AVEVA, Intergraph, Bentley), SPARQL endpoint provisioning for plant knowledge graphs, and SHACL-based data quality gates embedded in engineering workflows. These are practical engineering tasks that are now technically unblocked by the results of this project and are recommended as the immediate next industrial implementation phase.

ISO 23726-1, covering overview and fundamental principles of the OBI ecosystem, is projected for publication in 2028. The compliance principles defined in Part 1 – including the Direct Semantics requirement, the RDF conformance rule, and the alignment requirement that every class in an OBI-conformant ontology must subclass an IDO class – have been applied in practice throughout this project. The project's experience with applying these principles across four different standards constitutes a directly relevant empirical basis for the Part 1 content. A recommended further step is a formal contribution to the ISO 23726-1 editing process, drawing on the alignment work and the Arrowhead fPVN Handbook material produced here.

The Digital Product Passport (DPP) initiative under CEN/CLC/JTC 24 [30] (Section 3.2.2) is moving toward a standardized API layer for accessing product lifecycle data. IDO provides a natural semantic foundation for DPP data models, and the alignment architecture demonstrated in this project could be applied to make DPP data semantically interoperable with STEP, DEXPI, and ISO 15926 knowledge graphs. A concrete further step is to define an IDO-aligned OWL ontology for the DPP data model and a SHACL-based conformance layer, enabling DPP instances to participate in the same cross-standard reasoning environment established in this project.

3.1.2 ISO 10303: Product data representation and exchange (STEP)

ISO 10303, commonly known as STEP (Standard for the Exchange of Product model data), is the primary international standard for the computer-interpretable representation and exchange of product manufacturing information. It covers the full industrial lifecycle – from design and engineering analysis through manufacturing, configuration management, and long-term archiving – and is applied across aerospace, automotive, shipbuilding, construction, and other engineering-intensive industries. STEP is maintained by ISO/TC 184/SC 4 (Industrial Data) and is one of the most comprehensive standardization efforts in the ISO catalogue, with Application Protocols (APs) tailored to specific domains and use cases. Its significance lies in enabling tool-independent, vendor-neutral exchange of rich product data, which underpins digital continuity across complex supply chains and engineering programs.

3.1.2.1 Scope of the standard

In the modernized STEP architecture, a harmonized information model (the STEP “Core model”) is developed in SysML and intended to be shared across multiple STEP Application Protocols (APs). This strengthens semantic consistency and reduces redundant modelling across related standards in the 10303 family (Figure 5). ISO 10303 can be understood relative to other industrial data standards – such as IFC (ISO 16739), ISO 15926, or DEXPI – to highlight its breadth, depth, and cross-industry interoperability role.

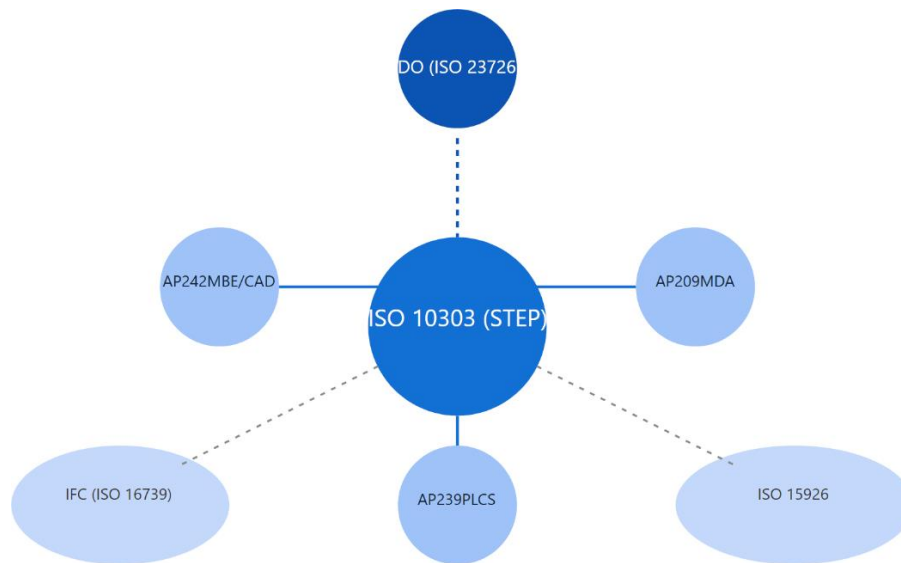


Figure 5. Standards Landscape around ISO 10303 STEP

The ISO 10303 committee was founded in 1984. At inception, no established modeling languages existed for this kind of work, so the committee developed its own formalism – *Express* – together with associated architectures and tooling. This long heritage has produced one of the largest and most ambitious standardization efforts within ISO: a family of more than 800 standard numbers covering industrial domains, modeling approaches, and implementation methods. SC 4 standards are cited as driving economic benefits through improved interoperability, and the committee maintains harmonization with IEC on electrical and automation standards.

SC 4 is actively advancing the strategic direction of the STEP family toward ontology-based interoperability, digital twins, and data-driven compliance, with the explicit goal of leading Industry 4.0 innovations. This includes the ongoing modernization of STEP around a harmonized SysML-based Core model, and growing interest within the committee in W3C Semantic Web technologies as a complement to established STEP exchange formats.

Table 3. Relevant ISO 10303 STEP Application Protocols

AP ID	Name	Why it matters
AP242	Managed Model-Based 3D Engineering	AP242 supports the exchange and long-term archiving of product structure, configuration control, and multiple geometry representations.
AP209	Multidisciplinary Analysis and Design (MDA)	AP209 enables exchange of information between CAD and CAE/FEA domains with traceability between design and analysis.
AP239	Product Life Cycle Support (PLCS)	AP239 defines lifecycle support information including configuration control, structures, interfaces, requirements, planning, and documentation.

3.1.2.2 Our work and cooperation with the standards organization

In the summer of 2024, representatives from the project met with the chair of ISO/TC 184/SC 4, Kenneth (Kenny) Swope of Boeing. In follow-up correspondence, Swope highlighted the strategic value of improving integration between the ISO 15926 and ISO 10303 families and suggested using ISO 23726 (IDO) as a future normative reference point for “mutual content” across industrial data standards.

An important observation motivating this work is that XMI – the XML Metadata Interchange format used to serialize SysML models – is a syntax, not a semantic layer. XMI captures model structure with deterministic, parser-friendly artifacts and stable identifiers, but does not provide machine-interpretable semantics or support for automated reasoning. Transformation to OWL 2 DL is therefore necessary to unlock semantic interoperability and integration with knowledge graphs. The work leveraged the fact that relevant STEP Core content is available in SysML form, making it possible to export the model as an XMI file. ISO provides established transformation guidance for SysML-based XMI representations, which provided the basis for a structured transformation pipeline from SysML XMI to OWL 2 DL. Manual tests were performed on the portion of the STEP Core model relevant for the Arrowhead fPVN project. A first semi-manual proof-of-concept (POC) OWL 2 DL ontology was created. The POC integrated well with other OWL 2 DL-based knowledge graphs.

This work was presented to the fPVN Arrowhead review in Brussels in June 2025 and later to the ISO/TC 184/SC 4 community during the SC 4 meeting week in Nagasaki (26–31 October 2025). In July 2025, the Arrowhead fPVN WP3 subproject decided to automate the transformation from XMI to OWL 2 DL for improved repeatability and scalability.

3.1.2.3 Methods used

The work followed a model-to-ontology transformation approach, using a SysML XMI export of STEP Core content as the source representation and OWL 2 DL as the target formalism (Figure 6). XMI was a well-suited starting point for this pipeline: as a standardized serialization of SysML models, it produces deterministic, parser-friendly artifacts with stable element identifiers and consistent structural representation. This predictability is essential for rule-driven transformation, as it allows mapping rules to be defined once and applied across all modules of the STEP Core model.

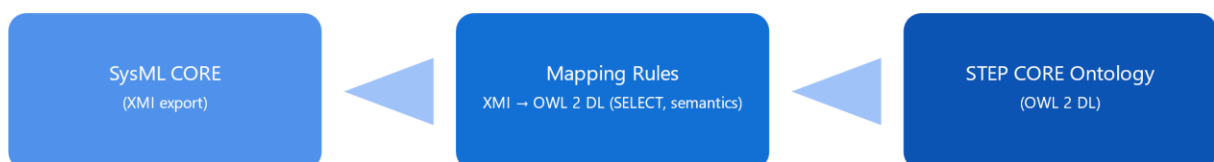


Figure 6. The model-to-ontology transformation pipeline

The first step was to identify and scope the relevant parts of the STEP Core model for the UC 2.9 use cases. A custom rule set was then developed to map SysML constructs to their OWL 2 DL equivalents. This included handling of SELECT types, which are a distinctive feature of the STEP modeling language: a SELECT type defines a named choice between two or

more types, and was mapped to OWL using a union class construct (`owl:unionOf`) together with a reified pattern that makes the choice explicit through named individuals. Modules were transformed one at a time, which allowed each output to be checked and verified before being merged into the growing ontology. The resulting OWL artifacts were then validated for other OWL-based knowledge graphs used in the project and aligned with ISO 23726 (IDO).

In the final stage, the entire STEP Core XMI was processed using a Python-based tool that combined standard XML parsing libraries with RDF read and write capabilities. This automated run produced an ontology containing 12,610 axioms, 1,075 classes, 2,113 object properties, and 3,374 subclass relationships. The transformation preserved the engineering intent of the STEP Core model to a very high degree.

3.1.2.4 Results

Semi-manual proof-of-concept: The initial proof-of-concept produced a STEP Core OWL 2 DL ontology that integrated well with other OWL-based knowledge graphs and aligned with project ontologies. The modular structure of STEP Core enabled module-by-module transformation (activity, analysis, breakdown, document management, PDM, configuration, geometry reference patterns), which simplified verification.

Automatically generated ontology: The transition to a fully automated transformation pipeline succeeded. Mapping rules adapted from STEP XMI→XSD guidance to XMI→OWL 2 DL preserved engineering semantics. Since the entire STEP Core XMI was processed, the resulting ontology is broader than what is needed for UC 2.9. Validation focused on UC 2.9, with no signs of reduced quality elsewhere.

Lightweight alignments: The Upper Ontology approach enables alignment across large industrial ontologies. In UC 2.9, two major engineering standards were used, yet only a small number of classes were aligned. Despite this, the interoperability between the knowledge graphs already produced significant business benefits. This indicates that early results and value can be achieved without large-scale alignment projects, and benefits will grow as more classes are aligned over time.

Interoperability positioning (Figure 7): The STEP Core → OWL transformation is positioned alongside other standards-to-OWL initiatives as part of a broader OWL 2 DL-based interoperability strategy.

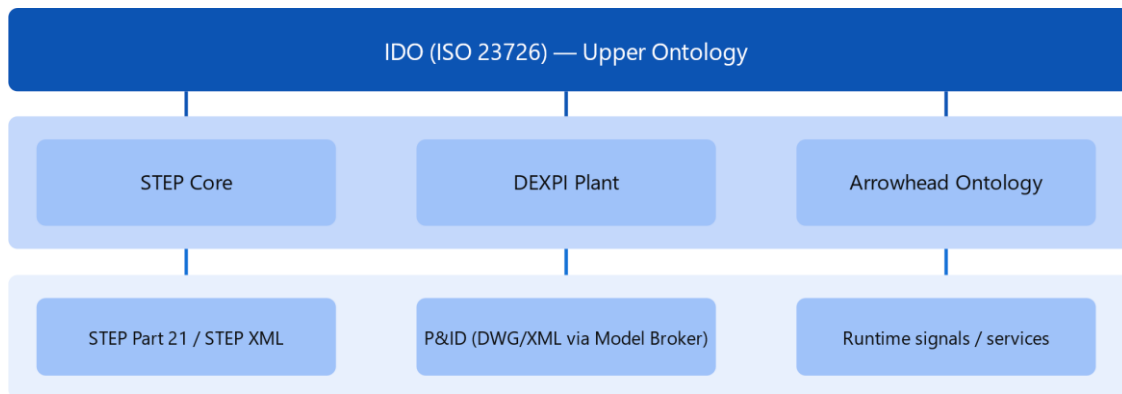


Figure 7. Alignment Layers – Interoperability Positioning

3.1.2.5 Proposed further work

The proof-of-concept demonstrated that transformation of STEP Core content into OWL 2 DL is feasible, while preserving engineering semantics. Further work is expected to continue in coordination with ISO/TC 184/SC 4 to ensure consistency with the ongoing development, maintenance, and harmonization of the ISO 10303 standard family and connected standards (Figure 8). This work should be understood as a contribution to the broader SC 4 standardization context, supporting long-term semantic interoperability across ISO 10303, ISO 15926, ISO 23726 (IDO), and related industrial data standards.

	Now	3-6 months	6-12 months
Tooling	Part 21/XML PoC	Ingestion MVP	Prod ingestion
Rules & Tests	Publish rules v1	Regression tests	Versioning policy
Pilots	Pilot AP239	Pilot AP242	Industrial E2E
Governance	Steering cadence	SHACL gates	Release KPIs

Figure 8. IDO roadmap

We need collaboration on model representations and semantic formalisms. Engagement with SC 4 working groups is expected to continue regarding the role of SysML-based representations of STEP Core models and their relationship to formal semantic representations. This includes consideration of how SysML/XMI artifacts can support interoperable semantic models without altering the normative content of the standards.

Formalization and governance of transformation principles is also necessary. Further work may include contributing to SC 4 discussions on the documentation, validation, and governance of transformation rules and patterns (e.g., SysML/XMI to OWL 2 DL). Such work should ensure traceability to the original standard artifacts and support consistency across implementations and over time.

A concrete further step in this governance work is the introduction of SHACL (Shapes Constraint Language) as a mechanism for validating semantic conformance of the transformed

OWL artifacts. SHACL shapes can be defined within the STEP Core ontology to specify structural and semantic constraints – for example, that every class derived from a STEP entity must carry a defined identifier, or that SELECT-type union classes must be correctly bounded. Automated SHACL validation gates can then be integrated into the transformation pipeline, running at each release to detect constraint violations before the ontology is published or consumed by downstream knowledge graphs. This would provide a principled and repeatable governance mechanism, directly analogous to the kind of schema validation already established in the STEP exchange file formats, but operating at the ontology level.

The work carried out in this project provides a foundation for two potential contributions to SC 4 as ISO Technical Specifications: one covering SysML Models in OWL 2 DL, documenting the transformation rules, mapping patterns, and recommended practices for deriving OWL 2 DL ontologies from SysML-based information models such as the STEP Core; and a second covering Constraint Validation using SHACL, defining conformance shapes and validation patterns applicable to OWL 2 DL ontologies derived from SC 4 standards. Contributing these deliverables as ISO Technical Specifications within the SC 4 framework would formalize the pipeline developed in this project, give it normative standing within the ISO 10303 family, and make it available to the wider SC 4 community as a reusable foundation for future ontology-based work.

Continued collaboration with SC 4 may support the positioning of ISO 23726 (IDO) as a reference framework for “mutual content” across standards. This includes alignment between ISO 10303, ISO 15926, Industry Foundation Classes (IFC), and other relevant standards with ISO 23726 (IDO) as a reference for shared semantics.

Further investigation may be undertaken within SC 4 to explore approaches for representing STEP instance data (e.g., ISO 10303-21 and ISO 10303-28) in semantic form. This includes consideration of how instance-level data can be consistently related to schema-level representations within a unified semantic framework. Incremental validation may be carried out in collaboration with SC 4 domain experts using selected Application Protocols (e.g., AP239, AP242). This supports verification that domain-specific semantics, lifecycle aspects, and configuration management concepts are preserved and correctly represented.

The growing cooperation between ISO/TC 184/SC 4 and IEC – particularly IEC/TC 65 on industrial-process measurement, control and automation [28] – is an important context for this work. SC 4 and IEC/TC 65 share an increasing overlap in the subject matter (covering digital twins, smart manufacturing, and Industry 4.0 architectures), and have established joint work to reduce duplication and ensure consistency across standards families. The ontology-based interoperability approach developed in this project spans both domains: STEP (SC 4) and Asset Administration Shell (AAS, IEC/TC 65) address complementary aspects of the industrial digital ecosystem, and semantic alignment between them is both technically feasible and strategically desirable. Further work should take this ISO–IEC cooperation into account, seeking opportunities to contribute results that are applicable across both standards families and that support the shared goal of machine-readable, semantically interoperable industrial data.

All further work should remain consistent with the evolving SC 4 architecture, including the STEP Core model, modularization principles, and ongoing harmonization efforts across the ISO 10303 family.

3.1.3 ISO 15926: Integration of life-cycle data for process plants including oil and gas production facilities

ISO 15926 is one of the most mature and comprehensive data integration standards in the process industry. It was developed in response to the need for improved data exchange and integration capabilities between systems involved in major process plant, and oil and gas (O&G) project development and operations. Two major initiatives contributed to its creation: EPISTLE (European Process Industries STEP Technical Liaison Executive), driven by the European process industry, and PCA (POSC Caesar Association), representing the O&G industry. PCA is a cooperation between POSC (Petrotechnical Open Software Corporation), that started in 1990 focusing on subsurface and drilling, and the Norwegian initiative Caesar Offshore focusing on the topside of offshore platforms. Their cooperation, established in 1995, resulted in the publication of ISO 15926-2:2003, the foundational data model of the standard.

Within Arrowhead fPVN, ISO 15926 plays a dual role: as a source of reference data that can be mapped to IDO, and as a framework whose extension with process-industry-specific reference data (including pulp and paper) demonstrates cross-industry applicability. The standard's lifecycle scope, semantic depth, and active reference data library make it relevant to the project's goal of ontology-based interoperability.

3.1.3.1 Scope of the standard

The scope of the standard is reflected in its title, “Integration of life-cycle data for process plants including oil and gas production facilities”. ISO 15926 defines the representation and exchange of product-related information across the complete industrial lifecycle. The standard covers data required for product design, engineering, manufacturing, operation, and maintenance. Its primary purpose is to ensure that datasets describing products, processes, and supporting information remain interoperable across tools, organizations, and long operational life spans.

The parts of ISO 15926 relevant for the current work in fPVN are:

- ISO 15926-1:2004, Overview and fundamental principles
- ISO 15926-2:2003, Data Model
- ISO/TS 15926-4 Ed. 4, Initial reference data (RDL) – an early version applicable for fPVN
- ISO 15926-6:2024, Methodology for the development and validation of reference data

ISO 15926-2 is a generic model; its O&G orientation is reflected in the reference data of ISO 15926-4, not in the data model itself. This means the standard can serve other process plant industries, including pulp and paper, without modification to its foundational layer. Figure 9 shows the relationship of the ISO 15926-2 data model to internal and external models (see also ISO 15926-1:2004(E):5.2).

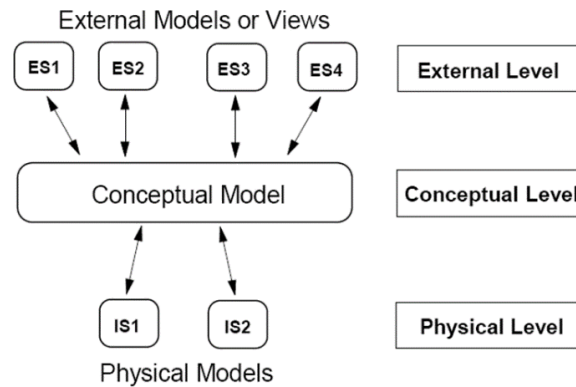


Figure 9. ISO 15926-2: Data model relationship to internal and external models

The data model is represented using entity data types organized in a universal subtype/supertype hierarchy. ISO 15926 was originally not designed to support reasoning, but as the development of reference data progressed, it became clear that an ontological representation was required to provide precise and unique class definitions. Reference data defined according to ISO 15926-4 Ed. 4 can therefore serve as a basis for reasoning, provided it is exposed through an appropriate OWL representation.

Both ISO 15926-2:2003 and ISO 23726-3 (IDO: Industrial Data Ontology) are upper ontologies that exclude domain-specific data. Their alignment – making ISO 15926 classes subclasses of IDO classes – forms the semantic bridge through which ISO 15926 reference data becomes accessible within the IDO-based interoperability architecture pursued in fPVN.

Reference data in ISO 15926 is subdivided into the following categories (see Figure 10):

- Core classes (ISO 15926-1:2004, 3.1.4)
- Proprietary classes (3.1.16)
- De facto classes (3.1.8)
- Standard classes (3.1.20)
- Manufactured product classes (3.1.14)
- Commodity product classes (3.1.2)
- Proprietary product classes (3.1.17)

ISO 15926-4 Ed. 4 includes only *Core classes*, which serve as a required extension of the ISO 15926-2:2003 data model and as the basis for all other reference data categories. Further development of reference data can then build on concepts more familiar to practicing engineers, without requiring direct engagement with the underlying modelling language of ISO 15926-2.

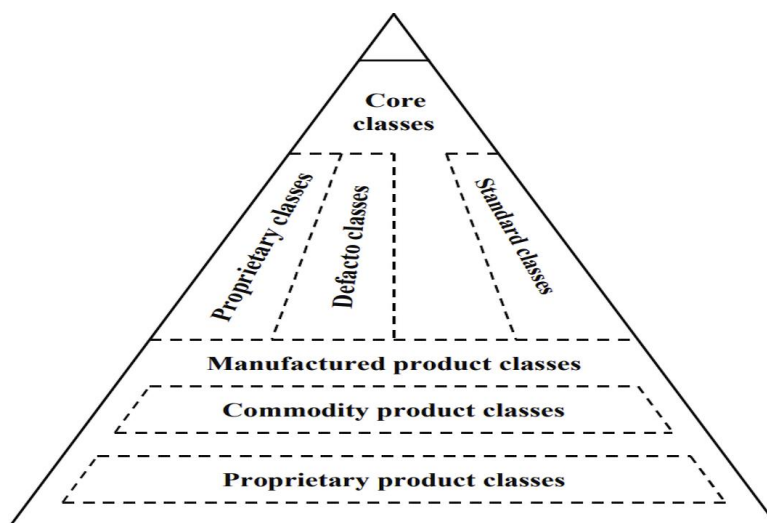


Figure 10. Types of ISO 15926-1:2004(E) classes

ISO 15926-4 Ed. 4 is published in a spreadsheet following the format specified in ISO CD-TS 15926-6:2007 Annex D, with columns including: *designation*, *text definition*, *source*, *notes*, *superclass 1–3*, *ISO 15926-2 entity*, and *classification*. The upper ontology classes are identified by the classifications `ISO15926-2 ENTITY TYPE` and `ISO15926-12 ENTITY TYPE` in the *Classification* column.

3.1.3.2 Our work and cooperation with the standards organization

The fPVN work on ISO 15926 pursues two complementary objectives: to extend the standard's reference data to cover pulp and paper industry concepts, and to align that reference data with IDO to enable ontology-based interoperability across the fPVN use cases. The scope of the current task is to develop a Reference Data Library (RDL) for fPVN, based on and extending ISO 15926-4 Ed. 4, for use with IDO. The extended RDL will be proposed for inclusion in a future ISO 15926-4 Ed. 5. The working file is based on `ISO15926-4ed4\_ALL\_basis for fPVN-2`, dated 2025 May 27. Subsequent updates to `iso15926-4ed4\_ALL\_spreadsheet.xlsx` were tracked, with the voting on the final version concluding in 2026 April.

Collaboration with the ISO/TC 184/SC 4 standardization community has been ongoing. The cross-industry scope of the reference data – covering both O&G and pulp and paper (P&P) – was used to demonstrate that lifecycle reference data can be shared across industries. Engagement with the PCA and the DISC working group informed both the content and the methodology of the reference data extension work.

Regular weekly online meetings throughout the second half of 2025 and into 2026 provided the primary collaboration forums. These meetings progressively advanced the technical depth of the IDO alignment guidelines, including property classification, lifecycle stage modeling, identifier patterns, and reference data integration.

3.1.3.3 Methods used

All editions of ISO 15926-4 to date have been documented in a spreadsheet based on the format specified in ISO CD-TS 15926-6:2007 Annex D. The current working version, ISO 15926-4 Ed. 4, includes several columns added beyond the original format: *URI*, *UniqueNumber*, *Synonyms*, and additional *Superclass* and *Classification* columns.

For the fPVN extension, further columns were added to include: (1) The URI of the corresponding class in the PCA Production RDL, to facilitate future harmonization; (2) Finnish and Swedish names, to meet the multilingual requirements of the fPVN use cases; and (3) additional cross-references useful for the project scope.

The alignment of ISO 15926-4 Ed. 4 with IDO follows a subclass pattern: classes in ISO 15926-4 are declared subclasses of their corresponding IDO classes. This makes the ISO 15926 reference data directly accessible through IDO-based queries and reasoning, without modifying the normative content of either standard. Instance models derived from ISO 15926-4 then conform to the IDO model for interoperability with STEP, DEXPI, and CFIHOS knowledge graphs within the same semantic architecture (see Figure 16).

The methodology for developing and validating reference data follows ISO 15926-6:2024. It requires at minimum the identification of at least one superclass; a definition that distinguishes the class from its superclass(es) and sibling classes; and formal representation in the required columns.

3.1.3.4 Results

The fPVN extension of ISO 15926-4 Ed. 4 demonstrates that reference data can be shared between the O&G industry (refineries and topsides) and the pulp and paper industry. The review of ISO 15926-4 Ed. 4 content confirmed a considerable overlap at the *Core Class* level between these two industries, as expected given their shared process plant character. P&P-specific classes were identified and added to the working file, extending the reference data scope without disrupting the existing O&G content.

At the *Standard Class* level, differences between industries emerge, reflecting the divergent normative bases: O&G is predominantly aligned with US standards, while European P&P relies more on EN and ISO standards. IEC classes are common to both, providing a natural harmonization point.

The IDO alignment approach – making ISO 15926 classes subclasses of IDO – proved straightforward to apply at the *Core Class* level and integrates well with the broader OWL 2 DL-based knowledge graph architecture used in fPVN. Combined with the STEP *Core* ontology and the DEXPI *Plant* ontology (see Sections 3.1.2 and 3.1.4), the resulting semantic graph enables cross-standard queries over process, asset, and lifecycle information.

These results indicate that the approach is generalizable beyond the O&G and P&P industries. Any process plant industry whose concepts can be organized under ISO 15926-2 entity types – and then aligned with IDO – can benefit from the same reference data infrastructure.

3.1.3.5 Proposed further work

The current set of fPVN reference data in the working file covers the first step of the ISO 15926-6:2024 methodology: class identification and superclass assignment. To fully integrate this content with a future ISO 15926-4 Ed. 5, the following further work is required:

1. Definitions must be updated to fully comply with ISO 15926-6:2024, with formal requirements added and all relevant columns completed.
2. The content should undergo the same editorial review procedure applied to ``iso15926-4ed4_ALL_spreadsheet.xlsx``.
3. The reference data should be extended in a shared environment involving O&G (both refineries and topsides) and P&P industry participants, leveraging the demonstrated overlap at the *Core Class* level.
4. A next phase should address *Standard Classes*, which will require resolving the divergent normative bases between US-standard-aligned O&G and EN/ISO-aligned European P&P, while retaining IEC classes as a shared foundation.
5. The subclass mappings from ISO 15926-4 to IDO should be published as a formal OWL alignment file, to be maintained alongside both standards. SHACL shapes should be defined to validate conformance of ISO 15926-derived knowledge graphs for IDO structural constraints, consistent with the governance approach proposed for the STEP *Core* ontology (Section 3.1.2).
6. Future work should prepare for alignment between the fPVN RDL and the PCA Production RDL, using the URI cross-reference column already introduced in the working file.
7. The fPVN extension work should be formally proposed for inclusion in ISO 15926-4 Ed. 5 through the ISO/TC 184/SC 4 process. This contribution would demonstrate a replicable pattern for cross-industry reference data development under the ISO 15926 framework.

The broader context for this work is the strategic alignment between ISO/TC 184/SC 4 and IEC/TC 65, which is relevant because ISO 15926 reference data – once aligned with IDO – can also support interoperability with AAS submodels and IEC CDD entries, connecting the process-industry lifecycle layer with the Industry 4.0 layer addressed by IEC 63278 (see Section 3.2.1).

3.1.4 DEXPI: Data Exchange in the Process Industry

Data Exchange in the Process Industry (DEXPI) is an industry-driven initiative to develop a neutral and open standard for the exchange of engineering data, with piping and instrumentation diagram (P&ID) interoperability. It defines a representation of process plant information with a focus on process design and functional plant definition. It does not aim to cover the full lifecycle of a plant, nor the detailed physical realization of plant components.

The *DEXPI Process* part describes the engineering view of the process, independent of a specific installation. It captures the functional and logical aspects of the system, including process functions, material and energy streams, equipment definitions, control and instrumentation logic, and design conditions.

The *DEXPI Plant* part represents the functional plant structure based on tagging (functional locations). It includes the identification and structuring of apparatus/machines and instrumentation as defined in P&IDs, but remains at a functional and specification level rather than a physical one.

3.1.4.1 Scope of the standard

The scope includes support for process modelling and functional system definition; it provides a tag-based functional–location view of plant components; and enables consistent representation of P&ID semantics and engineering data. It does not define how functional components are physically realized or installed, but includes the specified objects. Neither does it cover detailed lifecycle aspects, such as manufacturing, construction, or operation.

This focus allows DEXPI to provide a clear and interoperable representation of engineering intent, while leaving physical realization and lifecycle management to complementary standards.

3.1.4.2 Our work and cooperation with the standards organization

DEXPI operates as a collaborative ecosystem where process industry companies, software vendors, engineering firms, and research organizations work together to ensure consistent and practical data exchange (Figure 11). DEXPI e.V. is a non-profit association and working group within ProcessNet [24]. It was established in 2011 as an industry initiative addressing interoperability challenges between engineering tools. The headquarters are located in Germany, closely linked to DECHEMA in Frankfurt am Main. DEXPI has an open membership model including multiple [stakeholder groups](#). Its mission and role is to develop a neutral and open data exchange standard for process engineering, enable the transition from document-centric to data-centric engineering, with a focus on semantic interoperability of engineering data, especially P&IDs.

DEXPI e.V. is not a formal standardization body such as ISO, but rather an industry consortium that develops its specifications on the foundation of established international standards, including ISO 10209, ISO 10628, IEC 62424, and IEC 61987. Rather than operating in isolation, DEXPI collaborates with related initiatives such as CFIHOS and OPC UA, and contributes input and validation to broader standardization efforts. Its specifications are shaped through a collaborative process involving working groups and broad industry participation. This ensures that the specifications are continuously refined based on actual use cases and tool implementations. The reliability and interoperability of these specifications are further ensured through cross-vendor testing and implementation projects.

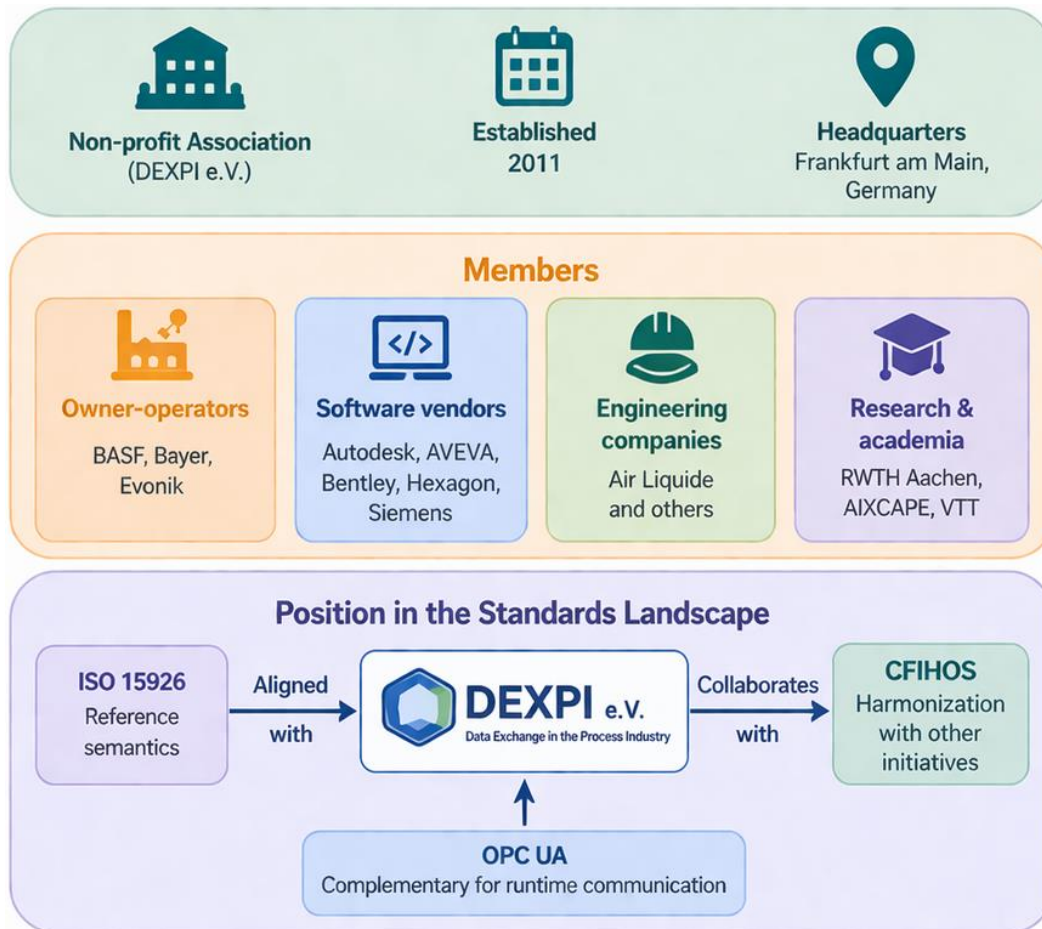


Figure 11. DEXPI in the standards landscape

From the beginning of the fPVN Arrowhead WP3 work Heiner Temmen of DEXPI e.V. has been involved in the technical work between the DEXPI organization and several of the WP3 member organizations. Even though DEXPI has not developed any official *DEXPI Plant* Ontology, they have been very helpful to produce such an ontology and to provide know-how to populate it with instance data. Lately an unofficial ontology has been agreed on within the DEXPI community which has been very important for the success of the development of the semantic *DEXPI Plant. Semantum*, member of both Arrowhead fPVN WP3 and DEXPI.org has been the key resource to automatically transform piping and instrumentation diagrams represented in PDF documents into P&ID Knowledge Graphs – one of many important result based on existing major engineering standards and W3C technologies.

3.1.4.3 Methods used

From the beginning of the cooperation between fPVN and DEXPI.org, there was only a limited interest within the DEXPI community in transforming the DEXPI UML model into an ontology, as well as transforming the *Proteus* file format for instance data into a DEXPI ontology-based knowledge graph. For Arrowhead fPVN, however, the transformation to OWL 2 DL and the use of knowledge graphs was considered fundamental. The WP3 group therefore decided to use the DEXPI UML model in XMI (XML Metadata Interchange) format as the basis for generating a *DEXPI Plant* ontology. The company Plant & Bytes (Germany) played a key role

in this process by performing the transformation of the DEXPI 1.4 XMI model into an OWL 2 DL ontology.

Semantum used its Model Broker product to generate Proteus files from P&IDs within scope. Proteus is an XML-based format, and by using standard XML technologies, TBH Konsult AB generated instance data conforming to the DEXPI 1.4 OWL 2 DL ontology. Although this approach was labor-intensive, it proved the feasibility of creating a semantic model based on DEXPI classes and relations. During 2025, DEXPI.org released version 2.0 of the DEXPI Plant model, and Semantum updated its Model Broker technology to support this new version.

In parallel, the DEXPI community defined a draft *DEXPI Plant* ontology, representing a significant step forward towards the principles promoted within Arrowhead fPVN. By the end of 2025, Semantum released an RDF-based output from the Model Broker. This enabled a fully integrated pipeline from P&IDs in DWG and PDF formats to knowledge graphs in RDF format, directly importable into the GraphDB environment used in fPVN (see the Automatic Pipeline and RDF Knowledge Graph sections in Figure 12). Throughout this pipeline, all semantics, topologies, and connectivity information are preserved.

3.1.4.4 Results

The combined two-step approach, consisting of an initial manual transformation process followed by the development of an automatic transformation pipeline, proved to be highly successful. The manual approach was essential in building a deep understanding of the DEXPI model, its structure, and semantics. This understanding was subsequently applied in the design and implementation of the automated pipeline.

Classes in DEXPI that carry a TAG identifier were mapped to CFIHOS TAGs and aligned with the IDO TAG class (a subclass of the IDO Functional Object). In a similar way, the STEP `BreakdownElementVersion` was aligned with the IDO TAG class. This alignment enables interoperability between DEXPI and STEP, as both the DEXPI tagged elements and the STEP BreakdownElementVersion represent the same conceptual entity within the IDO framework.

The topology of the DEXPI model defines the relations and connectivity of the system. This provides insight into how individual components interact to deliver the required system functionality at the system level. The combination of the CADMATIC 3D viewer with the system definition captured in the DEXPI knowledge graphs makes implicit information in the 3D model explicit. This significantly improves the ability to perform analysis, simulation, and system-level reasoning.

3.1.4.5 Proposed further work

Based on the results achieved, it is recommended that DEXPI.org establishes an official transformation of both the *DEXPI Process* and *DEXPI Plant* models into a formal ontology representation, aligned with the principles of the Industrial Data Ontology and the approaches demonstrated within fPVN. Such an ontology should provide (1) a native OWL 2 DL representation of the DEXPI models; (2) Include well-defined semantics for classes, relations, and attributes; (3) Support direct generation of RDF-based knowledge graphs; and (4) ensure align-

ment with related standards such as ISO 15926, CFIHOS, and STEP through a common semantic framework. An official ontology-based representation would reduce ambiguity in interpretation of the DEXPI models; enable consistent and interoperable implementations across tools and stakeholders; facilitate integration into digital twin architectures and data-centric engineering environments; and eliminate the need for ad hoc or project-specific transformations.

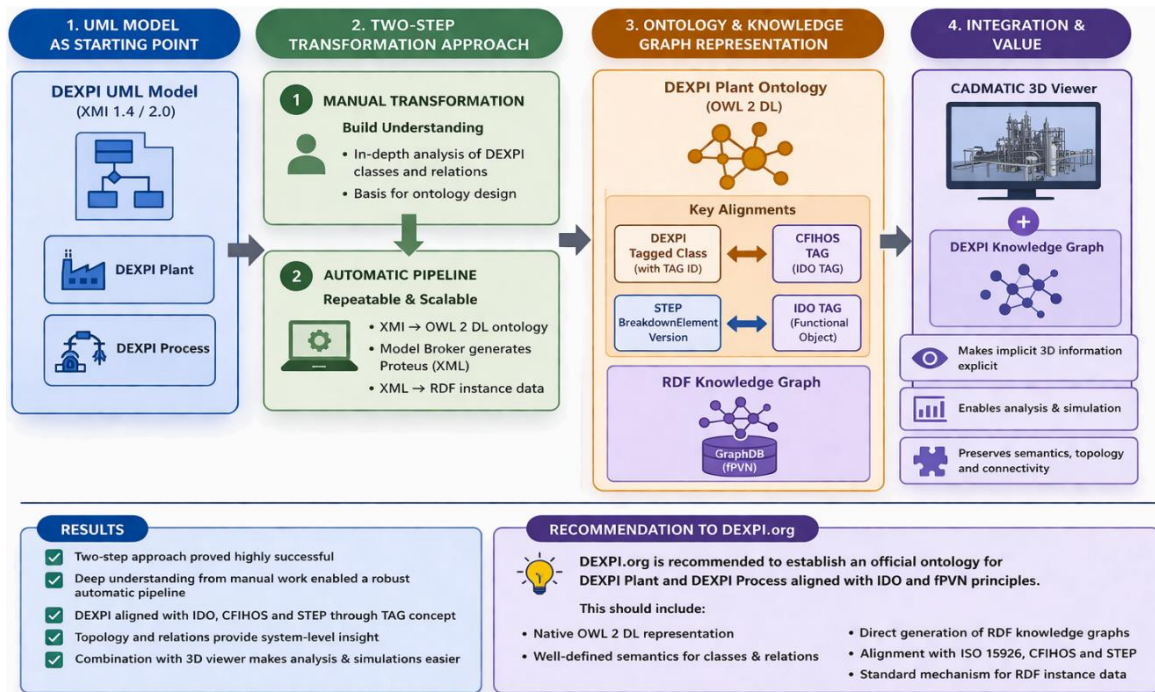


Figure 12. Graphical summary of integrating DEXPI into ontology-based industrial standards

Furthermore, it is recommended that DEXPI.org defines standardized mechanisms for representing instance data in RDF, complementing or extending existing formats such as Proteus. By taking these steps, DEXPI can strengthen its role as an enabler of semantic interoperability in the process industry and better support the transition towards knowledge graph-based engineering ecosystems. Figure 12 gives a graphical summary of the current section.

3.1.5 CFIHOS: Capital Facilities Information Handover Specification

The Capital Facilities Information Handover Specification (CFIHOS) is an industry standard for the structured exchange of information throughout the lifecycle of capital facilities, covering engineering, construction, operations, and decommissioning. Originating in 2012 as a Shell initiative under USPI-NL, governance transferred in 2020 to the International Association of Oil & Gas Producers (IOGP) as Joint Industry Project 36 (JIP36) [25]. It targets the full supply chain – owner/operators, EPC contractors, fabrication and commissioning contractors, and equipment suppliers – and is applicable across industries including oil and gas, chemical, nuclear, power, and water.

CFIHOS publishes three main deliverables: a Narrative Document Library (scope and procedure document, technical specification, implementation guides, and contracting scenario templates); a logical Relational Data Model covering the full facility lifecycle; and a Reference Data Library (RDL) providing standardized terminology, equipment classes, properties, lookup

lists, and document types. Despite membership being funded by annual company fees, the standard and all its deliverables are freely available to the public under a Creative Commons 4.0 license. As of Version 2.0 (published October 2024), CFIHOS comprises 78 member organizations and over 370 contributing subject matter experts.

3.1.5.1 Scope of the standard

The scope of CFIHOS is data handover for process and power plants across their full lifecycle. Figure 13 depicts a plant lifecycle activity model; within it CFIHOS is represented by the yellow block.

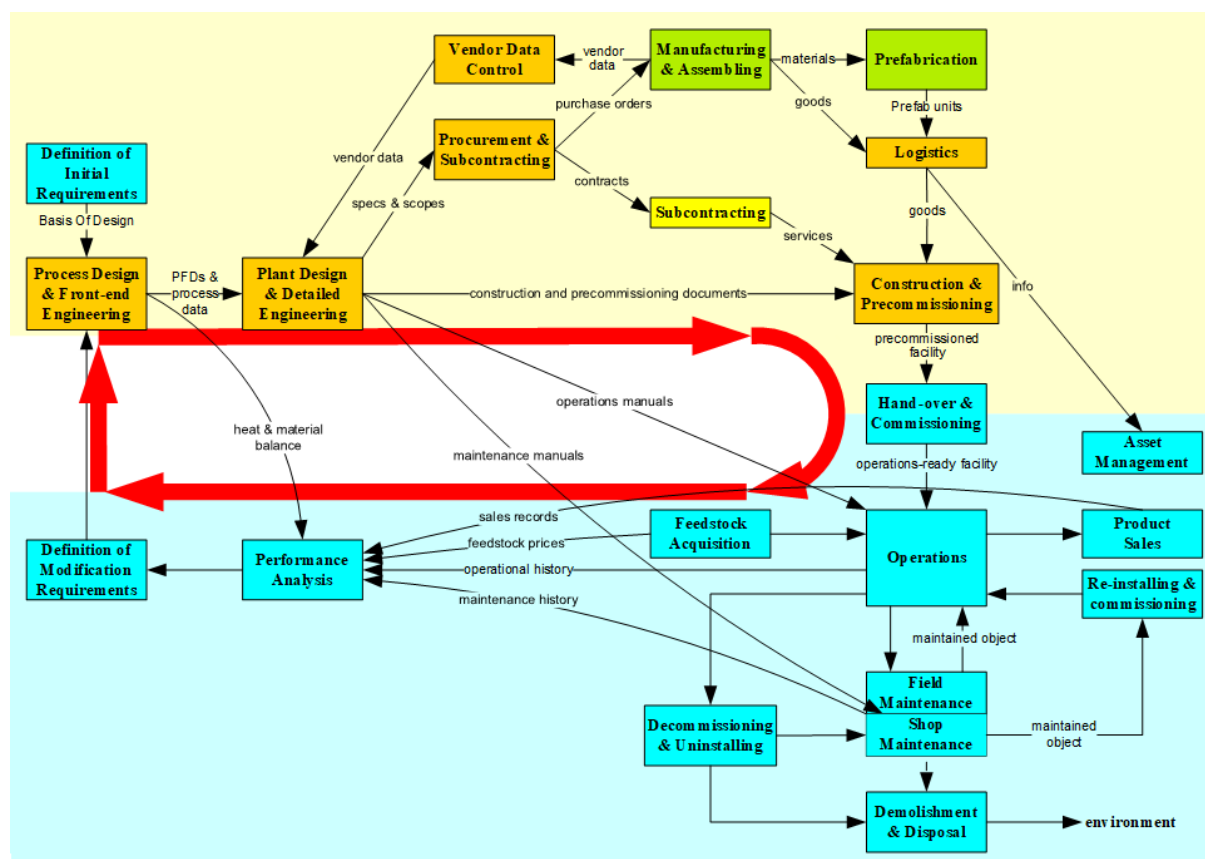


Figure 13. Plant Life Cycle Activity Model

CFIHOS provides practical, standardized specifications for information handover, covering the entire lifecycle of a facility from engineering and construction through operations and decommissioning. Its data models define entities, attributes, and relationships so that facility information is interpreted consistently across tools and organizations, enabling different systems to communicate without manual intervention. The standard applies across a wide range of industries – oil and gas, chemical, nuclear and other sustainable energy sectors, data centers, pulp and paper, and other capital-intensive industries – and supports multiple languages, including English and Chinese.

CFIHOS is intentionally focused on the information needed to own, operate, maintain, and demonstrate regulatory compliance for a facility. It does not aim to be an exhaustive catalogue of all collectable data, nor does it prescribe specific tools or software. Figure 14 illustrates an

example of a part of the data model; blue entities are CFIHOS standard content (tabs of the RDL spreadsheet), green is data to be provided by the EPC contractor, and yellow by the Owner/Operator.

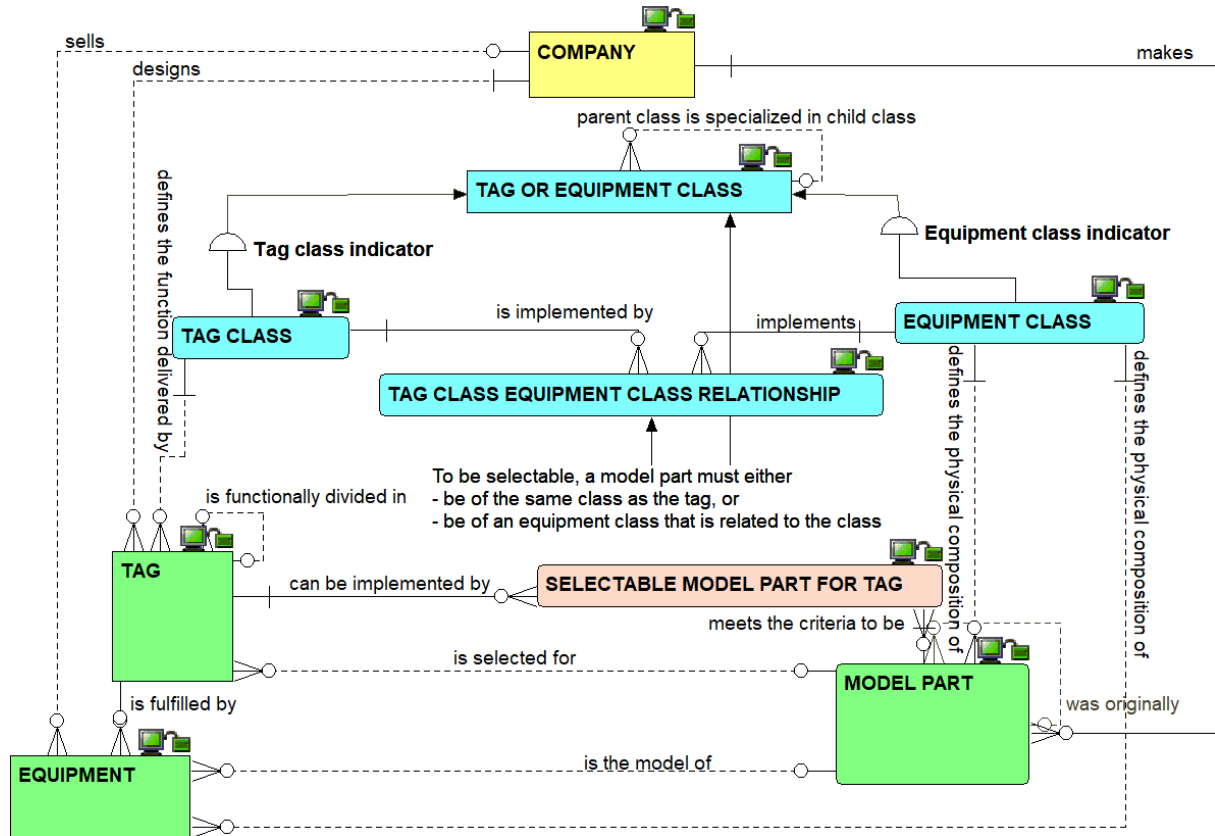


Figure 14. An example of a part of the data model

3.1.5.2 Our work and cooperation with the standards organization

CFIHOS is governed through an "onion model" of three concentric layers. The outer layer handles governance, reporting to the IOGP Standards Committee and an Engineering Leadership Council of twelve senior engineering executives from major oil and gas companies. The middle layer is a Steering Committee whose members represent each of the five membership groups rather than their own organizations. The inner layer consists of working groups that deliver content, drawn from across the member organizations. These cover the Reference Data Library, data model development, document requirements, software, implementation guidance, and marketing. The CFIHOS Semantics group – which focuses on the semantic representation of CFIHOS using IDO and on mapping to related ontologies – was established in early 2024 and is the primary interface between CFIHOS and the Arrowhead fPVN WP3 work. It is led by Onno Paap (FLUOR Engineering), who has been the main CFIHOS collaborator throughout the project.

In CFIHOS, every entry in the Reference Data Library must be accompanied by a reference to a source standard; the same discipline applies to JIP33. A mapping from CFIHOS to ISO 15926-4 Ed. 3 has been completed. Feature Requests have been approved for further mappings: to ECLASS [34], to the IEC Common Data Dictionary (CDD) [35], and to IDO.

CFIHOS also maintains active memoranda of understanding with DEXPI, PCA (POSC Caesar Association), and CII (Construction Industry Institute) to prevent duplication and ensure that domain-specific data models extend CFIHOS rather than replicate it.

Collaboration between fPVN WP3 and CFIHOS was initiated through a four-session training series organized within the WP3 IDO training program during November 2024 and January 2025. The first session (8 November 2024) was presented by Peter Townson, Project Director of CFIHOS, and provided a high-level overview of the standard, its governance, and Version 2.0. The second session (15 November 2024), presented by Jason Roberts (Shell), covered the logical data model in technical depth. The third and fourth sessions (24 and 31 January 2025), presented by Onno Paap, covered CFIHOS Semantics, the IDO alignment approach, and the SousLeSens tooling [36].

Following the training series, deeper technical collaboration took place at two face-to-face meetings in Oslo: on 19–20 March 2025 and 6–7 May 2025. These brought together representatives from Arrowhead fPVN WP3, the CFIHOS Semantics group, and the DISC initiative.

3.1.5.3 Methods used

CFIHOS specifies a relational database model, and the Reference Data Library (RDL) is used to populate the standardized data. The functional data model consists of 142 entities, while the RDL spreadsheet contains 39 tabs with data to fill the standard tables. The data model was created in ER/Studio, which can generate create-scripts for up to 1,000 database brands, available upon request. SQL Server is the commonly requested output. XMI format is also available, and a pre-filled MS Access database.

The CFIHOS Semantics ontology is a direct transformation of the CFIHOS RDL into OWL 2, aligned with IDO as the upper ontology. For the class hierarchy, CFIHOS plant objects – transmitters, pumps, valves, pipelines, and similar – are mapped to `ido:InanimatePhysicalObject`` as their primary IDO parent, with a secondary mapping to `ido:Artifact`` (objects that are the result of a deliberate activity). This dual-parent pattern follows ISO 15926 practice and was adopted after discussion within the CFIHOS Semantics group, since CFIHOS itself uses single-parent hierarchies.

A key modeling decision concerns the separation of functional and physical perspectives. The CFIHOS RDL uses a single `tag_or_equipment_class`` table with indicator flags distinguishing tag classes (functional, required items) from equipment classes (purchased, physical items). In the semantic model, each CFIHOS class is represented as two distinct OWL classes: one anchored in the TAG lifecycle activity dataset (the functional view, subclassing `ido:FunctionalObject``) and one in the EQUIPMENT lifecycle activity dataset (the physical view, subclassing `ido:InanimatePhysicalObject``). These two individuals are linked by an `isFulfilledBy`` relation, preserving the CFIHOS design intent while conforming to IDO's separation of functional and physical concepts.

For property modeling, the CFIHOS Semantics ontology follows the IDO aspects pattern. A property such as `design pressure`` is not represented as a plain data attribute but as a structured quality carrying an aspect (in this case a scope aspect: *design*). The recognized aspect categories are: *place* (e.g., entry, external, wall), *range* (e.g., average, minimum, nominal), *origin* (e.g.,

actual, calculated, measured), *function* (e.g., pumping, short circuit), *scope* (e.g., allowable, ambient, design, operating), and *context* (e.g., at rated temperature). The base property – pressure, temperature, flow rate – is placed after all applicable aspects in a canonical ordering, producing a *Standard Property* name. This normalization ensures that equivalent properties from different source standards can be recognized and merged.

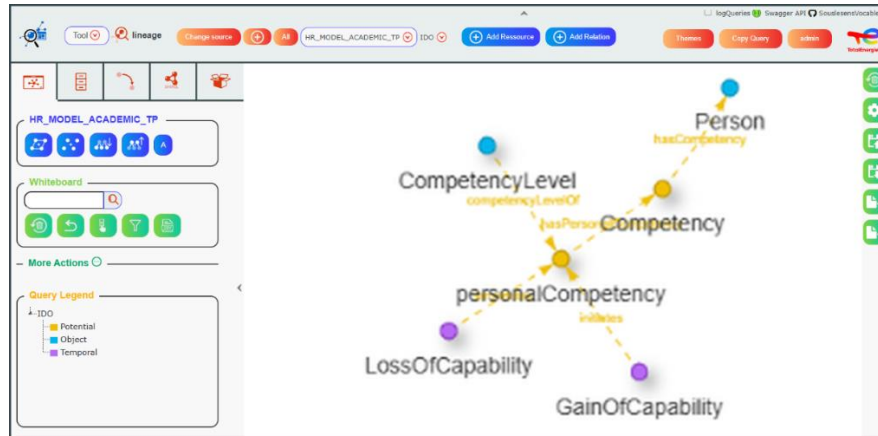


Figure 15. SousLeSens, an ontology editor used by the CFIHOS Semantics team

The CFIHOS Semantics team uses SousLeSens (Figure 15) as the primary ontology environment. SousLeSens is an open-source, web-based ontology editor written in TypeScript, originally sponsored by TotalEnergies and hosted by the Knowledge Graph Alliance [36]. It provides a graphical *Weaver* view for ontology navigation, an *Onto Creator* for authoring, and a SPARQL endpoint for querying. Rather than authoring the CFIHOS taxonomy manually in this tool, the team generates the ontology programmatically: a Python program reads the CFIHOS RDL spreadsheet, applies the IDO alignment rules, and produces a Turtle (`.ttl`) file that is then loaded into SousLeSens. This automated pipeline ensures that every release of the RDL can be reflected in the ontology without manual rework.

3.1.5.4 Results

The primary deliverable of the CFIHOS Semantics work within fPVN is the CFIHOS Ontology: an IDO-aligned OWL 2 representation of the full CFIHOS Reference Data Library, generated automatically from the RDL spreadsheet using the Python pipeline described above. The ontology covers the complete CFIHOS class hierarchy – approximately 500 tag and equipment classes – together with their associated properties and lookup value lists. It is hosted in the SousLeSens environment via the Knowledge Graph Alliance platform, with a public SPARQL endpoint providing open query access.

To support adoption of the CFIHOS Ontology and the broader IDO alignment approach, a dedicated handbook was produced within the fPVN project. The handbook provides step-by-step practical guidance on how to develop a domain ontology from the CFIHOS RDL, how to apply the IDO upper ontology as the alignment layer, and how to handle the key modeling decisions – including the TAG/EQUIPMENT lifecycle dataset split and the IDO aspects property pattern. It is intended as a reusable reference for any organization wishing to build or extend an IDO-compliant ontology from CFIHOS content.

On the standards alignment side, the mapping from CFIHOS to ISO 15926-4 Ed. 3 was completed during the project period. Feature Requests for further mappings – to ECLASS [34], to the IEC Common Data Dictionary (CDD) [35], and to IDO – have been approved and are in the implementation queue for 2026. These mappings will be reflected in both the CFIHOS RDL and the CFIHOS Ontology, with the IDO mapping automated through the same Python pipeline. The modeling work carried out jointly with the CFIHOS Semantics group and the DISC initiative at the Oslo meetings contributed directly to Deliverable D3.2 of the Arrowhead fPVN project.

3.1.5.5 Proposed further work

A number of Feature Requests are prioritized for implementation in 2026. The most substantial is the mapping of the approximately 55 JIP33 Procurement Data Sheets (PDS) [37], each containing hundreds of properties organized into property groups. The database of all PDS property metadata totals approximately 12,000 property names. These records will be mapped to CFIHOS using the bearer/aspect normalization approach described in the Methods section above, reducing the effective property count significantly by merging synonyms and structurally equivalent properties across equipment classes. The resulting standardized properties will be integrated into both the CFIHOS RDL and ISO 15926-4 Ed. 6. The CFIHOS Ontology will be updated automatically once the improved RDL is available. An algorithm will also ensure alignment with ISO 14224 (which defines Object Information Models for process industry equipment), enabling CFIHOS classes and their associated bearer components to be merged with the ISO 14224 equipment boundary definitions.

On the infrastructure side, a public internet server is planned to provide a dereferenceable CFIHOS namespace, a persistent SPARQL endpoint, and example plant instance data. This will make the CFIHOS Ontology a first-class linked data resource, accessible to any tool or organization without requiring a local installation of SousLeSens or any other software.

The approved Feature Requests for mappings to ECLASS [34] and the IEC Common Data Dictionary (CDD) [35] are scheduled for 2026, alongside the IDO mapping. A further planned mapping target is ISO 81346 (the international standard for structuring principles and reference designations), which would strengthen CFIHOS interoperability with maintenance and asset management systems. All completed mappings will be published as part of the CFIHOS RDL and reflected in the CFIHOS Ontology.

A newly added strategic theme in CFIHOS is the development of a data maturity model – a framework for specifying which information is required at which project milestone or contract event, covering both major handover points and more granular intra-contract triggers. A dedicated working group is being formed. Existing international standards will be surveyed before any new framework is defined.

Finally, CFIHOS is progressing toward formal ISO standardization. The CFIHOS Technical Specification document has been restructured in Version 2.0 as a true requirements document with unique, machine-readable identifiers and REQIF export, bringing it into alignment with ISO requirements management conventions. This work positions CFIHOS for submission as an ISO standard, which would significantly broaden its adoption and normative standing across industries beyond oil and gas.

3.1.6 SysML: Systems Modeling Language

Industrial systems engineering is undergoing a profound transformation. Systems are no longer developed as static, stand-alone artifacts, but rather as evolving systems of systems embedded in complex digital ecosystems. These systems must be specified, verified, and evolved across long lifecycles, while continuously interacting with operational data, services, and external stakeholders. In this context, traditional document-centric engineering approaches are increasingly insufficient, and a machine-interpretable, formally grounded modeling language becomes essential.

Systems Modeling Language (SysML) is a general-purpose modeling language for systems engineering, developed and maintained by the Object Management Group (OMG) [26]. SysML v2.0, published as a formal specification in September 2025 [27], is the next-generation version of the language, providing enhancements over SysML v1 in terms of precision, expressiveness, usability, interoperability, and extensibility. SysML v2 is based on the KerML meta-model, which is grounded in formal semantics. It includes a textual syntax in addition to graphical syntax, and provides a standard application programming interface (API) for interoperability with other tools and applications across the digital engineering ecosystem.

Discussions with OMG are under way on introducing an Arrowhead-based library as a domain extension for the next version of SysML v2. This builds on earlier SysML v1 work that modeled service-oriented systems of systems. The Arrowhead Framework's concepts for service orchestration, automation, and secure interoperability provide a mature starting point, while the specific contribution of the current work package is the alignment of system models with industrial data models to achieve semantic interoperability. Integrating Arrowhead semantics directly into SysML v2's more rigorous, machine-processable foundation would support consistent, lifecycle-spanning representations of services, capabilities, and data – and with them, the scalable engineering of complex interconnected industrial systems.

3.1.6.1 Scope of the standard

SysML v2 is used to specify requirements, behavior, structure, analysis, and verification while maintaining consistency and traceability across all aspects of the system model. Compared to SysML v1, SysML v2 introduces a cleaner semantic foundation and more expressive constructs for composition, specialization, and constraint definition. This makes SysML v2 particularly well suited for automation, model validation, and systematic reuse in large engineering programs.

At the same time, SysML v2 deliberately remains a design-time language. It follows a predominantly closed-world assumption and focuses on correctness and completeness within a bounded model scope. While this is essential for engineering validation, it limits the ability of SysML v2 to support semantic integration across heterogeneous data sources and standards once systems move beyond the design phase. This is where ontology-based approaches, and in particular ISO 23726-3, become critical.

3.1.6.2 Our work and cooperation with the standards organization

Within the Arrowhead fPVN project, SysML v2 is used as a primary language for capturing engineering intent, including system structure, behavior, interfaces, and service interactions. SysML v2 models serve as authoritative design representations that can be shared across organizational and tool boundaries while remaining sufficiently formal to support downstream processing.

SysML v2 and IDO are not competing modeling paradigms, but complementary layers in a digital engineering stack. Both aim to improve interoperability, consistency, and lifecycle integration, but they operate at different abstraction levels and serve different purposes. SysML v2 excels at expressing *how* a system is designed: its decomposition into parts and subsystems, its functional and behavioral responsibilities, and the constraints that govern its correct realization. IDO, defined as part of ISO 23726-3, focuses instead on expressing *what* things are and how they are semantically related across domains and standards. As an OWL-based ontology, IDO enables reasoning, inference, and open-world knowledge integration in ways that are outside the intended scope of SysML. This separation of concerns is not a weakness but a strength. SysML v2 provides rigor, structure, and engineering discipline, while IDO provides semantic richness, extensibility, and cross-standard alignment.

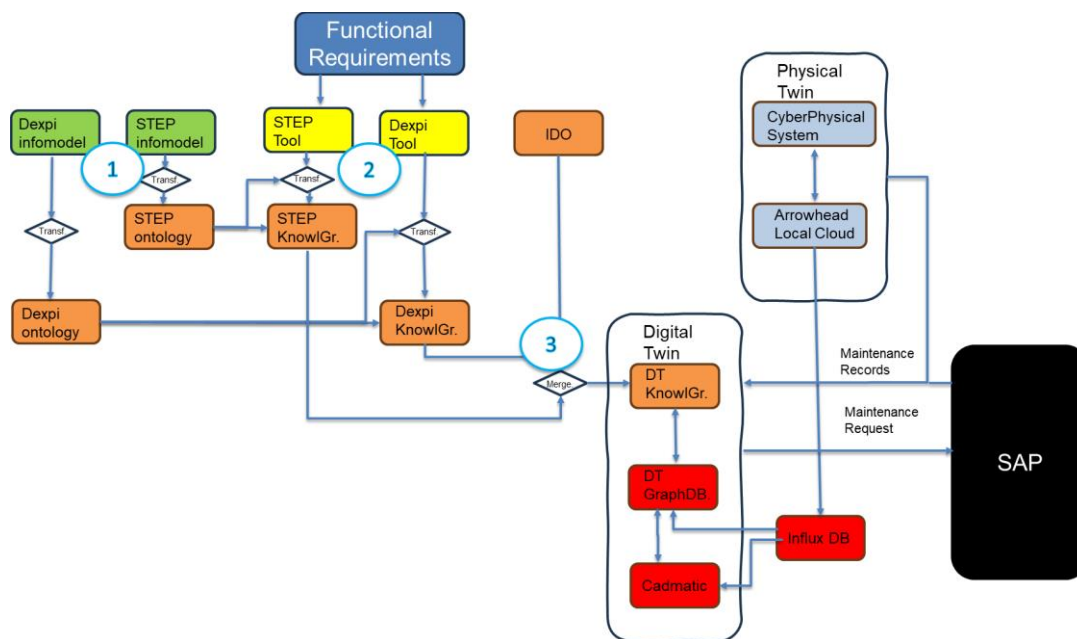


Figure 16. Conceptual architecture of interoperability among various standards

The conceptual architecture of interoperability among various standards is shown in Figure 16, while Figure 17 illustrates how this architecture is extended with SysML v2 lifecycle management, digital threads, domain libraries, and system-of-systems models.

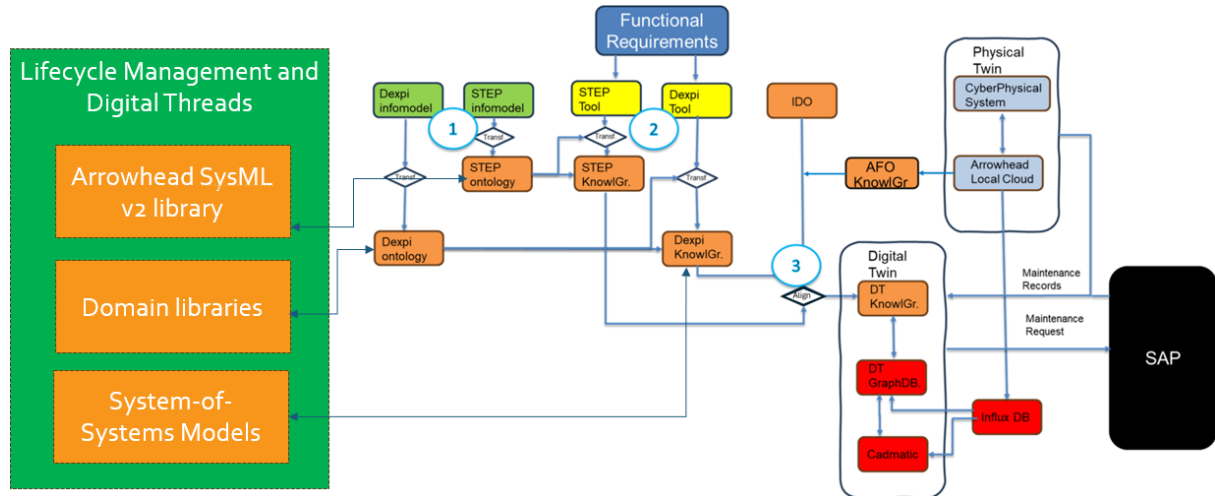


Figure 17. Conceptual architecture extended with SysML v2 lifecycle management and digital threads

3.1.6.3 Methods used

The primary method applied in this work is the use of SysML v2 as a disciplined, machine-interpretable representation for capturing engineering intent, and the systematic alignment of those representations with the Industrial Data Ontology (IDO).

```
#lcd processPlantD {
  #sysd heatExchangerD :> systems {
    #sd temperature : SensorReading :> services;
  }
  #sysd preheaterFlowRegulatorD :> systems {
    #sd flowConfig : Configure :> services;
  }
  #sysd preheaterControllerD :> systems {
    #sd temperature : SensorReading :> services;
    #sd conf : Configure :> services;
  }
}
```

Figure 18. SysML v2 textual model of a process plant subsystem

SysML v2 is used to model domain concepts such as process equipment, functional units, connections, and flows – illustrated here with DEXPI as the primary use case (see Section 3.1.4 for DEXPI-specific details). SysML v2 provides a disciplined way to capture not only the structural relationships between these elements, but also the engineering intent behind them. For example, a component can be explicitly modeled as fulfilling a specific process function within a defined system context, as shown in Figure 18 and Figure 19. This modeling approach allows domain information to be integrated with broader system architectures, including service-oriented and automation-related aspects emphasized in Arrowhead, and makes assumptions and dependencies explicit – which is essential for verification and evolution over time.

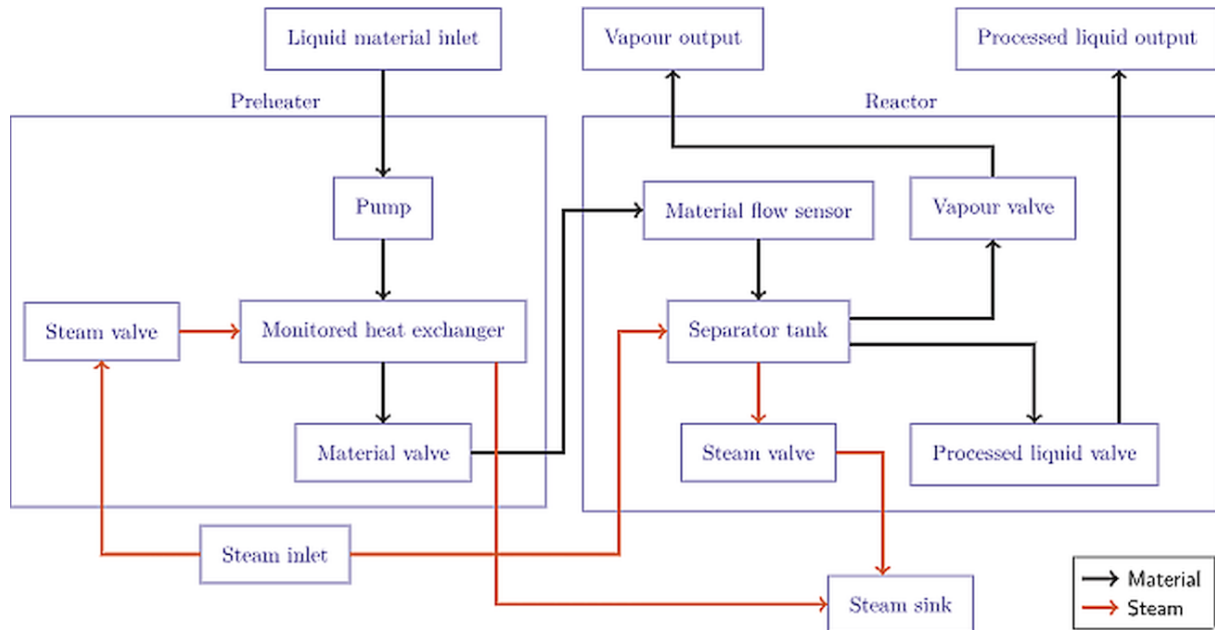


Figure 19. System architecture of the Preheater and Reactor subsystems, showing material and steam flows

Transforming SysML v2 models into ontology representations is a non-trivial task, but one that is becoming increasingly feasible with tool support and emerging AI-assisted techniques. The core challenges include identifying appropriate abstraction levels, flattening meta-model structures, and resolving the conceptual gap between closed-world design models and open-world ontologies. The transformation steps typically involve distinguishing design-level types from instance-level knowledge, mapping SysML constructs to OWL TBox and ABox elements, aligning SysML domain concepts with existing IDO vocabularies, and detecting or resolving semantic inconsistencies. The overall transformation pipeline – from .sysml source files through an LLM-assisted transformation prompt to a layered ontology output – is illustrated in Figure 20.

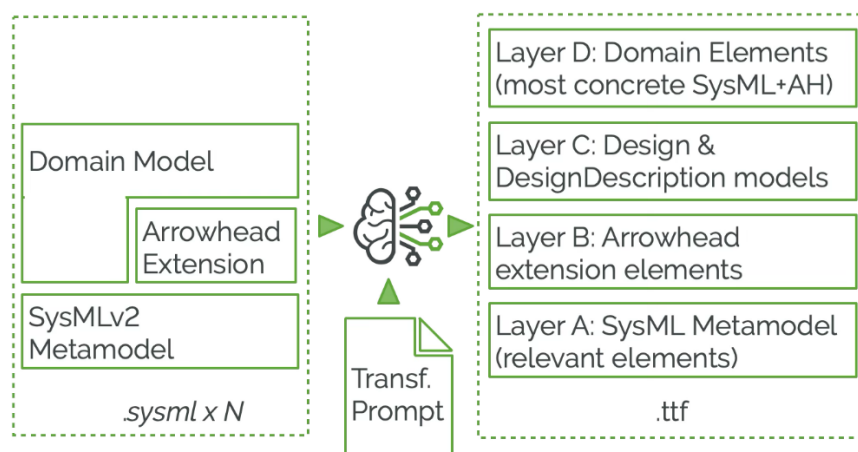


Figure 20. LLM-assisted transformation pipeline from SysML v2 models to layered IDO knowledge graph output

Recent experiments suggest that LLM-assisted workflows can support much of this work, particularly the repetitive and semi-formal steps involved in model transformation and ontology alignment. In this role, SysML v2 acts as a high-quality, structured knowledge source, while IDO becomes the convergence layer for semantic interoperability across systems and standards.

3.1.6.4 Results

A major benefit of aligning SysML v2 models with IDO is the ability to establish *continuous digital threads* across lifecycle phases and standards. SysML v2 models already offer strong internal traceability: requirements can be traced to functions, functions to components, and components to interfaces and services. However, this traceability typically remains confined within the modeling environment. By mapping SysML v2 elements to IDO concepts, these design-time relationships can be lifted into a *shared semantic space*, where SysML-derived knowledge can be combined with data from other standards such as STEP, ISO 15926, or DEXPI, enabling a unified lifecycle view. In practical terms, this makes it possible to perform consistency checking across heterogeneous models and datasets, to validate design assumptions against operational or domain knowledge, to execute semantic queries that span multiple engineering disciplines, and to assess the impact of lifecycle changes in a principled and traceable way. In effect, SysML v2 provides the structured input, while IDO provides the semantic substrate that allows this structure to participate in broader knowledge graphs and reasoning processes.

DEXPI provides an ideal use case to demonstrate the practical value of aligning SysML v2 with IDO: DEXPI models are inherently complex, lifecycle-spanning, and closely intertwined with other industrial standards, particularly ISO 15926, and the domain of process plant engineering demands both structural correctness and semantic clarity. Applied to this use case (see Section 3.1.4 for full details), SysML v2 elements such as parts and relationships were mapped to IDO individuals and relations grounded in a common semantic framework. As a result, DEXPI data is no longer isolated within a single standard but becomes interoperable with other lifecycle and domain models. This enables the semantic integration of DEXPI plant models with ISO 15926 reference data, supports cross-standard reasoning over process, asset, and lifecycle information, and allows DEXPI knowledge to be reused in digital twins and analytics environments.

Crucially, SysML v2 ensures that the mapped information remains faithful to the original engineering design, while IDO ensures that it is usable in a broader semantic ecosystem. Together, SysML v2 and ISO 23726-3 provide a pathway from disciplined system design to semantically integrated lifecycle knowledge – reducing manual data translation, improving consistency across standards, and laying the foundation for scalable digital engineering platforms. More generally, the results suggest that SysML v2 and ISO 23726-3 together can serve as cornerstones of future industrial digital ecosystems, supporting both rigorous engineering and flexible, ontology-based interoperability.

3.1.6.5 Proposed further work

The work presented here demonstrates the feasibility of using SysML v2 as a structured source for ontology-based knowledge integration, with DEXPI as the primary use case. A natural next step is to extend this approach beyond DEXPI to other industrial standards and domains represented in the Arrowhead fPVN work program. Applying the SysML v2–IDO alignment pattern to standards such as CFIHOS, ISO 15926, and STEP would broaden the scope of semantic interoperability and strengthen the overall multi-standard knowledge graph architecture.

The transformation work between SysML v2 and IDO currently relies on a partly manual and exploratory process. Further work should aim to formalize and automate this pipeline, drawing

on the experience gained with the STEP Core XMI-to-OWL transformation and the DEXPI Proteus-to-RDF pipeline. Establishing a repeatable, tool-supported workflow for SysML v2-to-OWL conversion would improve scalability and allow the approach to be applied consistently across different domain models and project contexts.

The experiments with LLM-assisted model transformation and ontology alignment represent a promising but early-stage direction. Further investigation is needed to understand where AI-assisted workflows can reliably support the transformation process and where human review remains essential. Developing appropriate validation mechanisms and quality assurance practices for LLM-generated mappings will be an important area of further work.

Now we have a good opportunity to contribute insights from the Arrowhead fPVN work back to the standardization process. In particular, the practical experience of aligning SysML v2 with OWL-based ontologies may inform future versions of the standard or associated tooling guidance. Engagement with OMG on the topic of SysML v2 interoperability with semantic web technologies is a natural avenue to pursue.

Finally, as SysML v1 remains in widespread industrial use, further work may address the practical challenges of supporting a transition toward SysML v2 in existing engineering environments. This includes understanding how legacy SysML v1 models can be progressively migrated or aligned with IDO-based knowledge representations, and how tool vendors and standards bodies can coordinate to ease this transition for industrial practitioners.

3.2 Standards in the WP3 work program but not tested in use-cases

Not all standards examined within the WP3 work program were directly validated through the Arrowhead fPVN use cases. The two standards presented in this section – IEC 63278 (Asset Administration Shell) and the Digital Product Passport (DPP) – were studied, analyzed, and architecturally positioned within the project's semantic interoperability framework, but did not serve as primary data sources or query targets in UC 2.9 or UC 3.9. This does not diminish their strategic relevance: both are central to the broader industrial digitalization landscape that Arrowhead fPVN addresses, and both have been shown to be technically compatible with the IDO-based knowledge graph architecture established in the project. The AAS provides an engineering-oriented interface layer for asset information that is architecturally complementary to the Digital Twin Knowledge Graph, while the DPP represents a cross-domain, regulation-driven information view that can be realized as a query-driven projection over semantically integrated data sources. The results presented here are therefore best understood as architectural positioning and proof-of-concept work that establishes the conditions for full integration in subsequent projects and industrial deployments.

3.2.1 IEC 63278: Asset Administration Shell

The Asset Administration Shell (AAS) represents a digital representation of an asset designed to enable interoperable data exchange across systems, organizations, and lifecycle stages. Within the context of Industry 4.0, it forms the digital counterpart of an asset, together constituting an Industry 4.0 component [29].

The AAS provides a structured and machine-readable representation of asset information through a modular architecture based on submodels. These submodels encapsulate different

aspects of an asset, such as identification, technical properties, operational data, and services. The approach is technology-neutral: it supports multiple serializations such as JSON, XML, and OPC UA, and enables integration across both IT and OT domains.

From an architectural perspective, the AAS can be understood as an interface layer that exposes asset-related information in a structured and engineering-oriented manner, while the underlying data and logic may originate from multiple heterogeneous sources.

3.2.1.1 Scope of the standard

The AAS specifications define a common framework for representing assets digitally, including information models, communication interfaces, and exchange mechanisms. The scope covers identification, semantic referencing, submodel structures, and access mechanisms. The scope of the AAS includes (1) representation of asset information across the full lifecycle; (2) definition of interoperable interfaces and APIs; (3) integration between IT systems and operational technologies; and (4) support for semantic interoperability through concept descriptions and external dictionaries.

At the same time, the AAS does not enforce domain-specific models or implementation choices. This allows flexibility across industries but also means that consistency and interoperability depend on conventions, templates, and governance outside the core specification.

3.2.1.2 Our work and cooperation with the standards organization

The work has been carried out in alignment with developments within the Industrial Digital Twin Association (IDTA), which governs and maintains the AAS specifications and submodel templates. Rather than contributing directly to the formal standardization process, the work focuses on implementation-oriented research and validation of architectural approaches that complement the AAS ecosystem.

In collaboration with academic partners – in particular with the Luleå University of Technology –, the work explores how AAS can be integrated with semantic technologies and distributed system frameworks, such as the Arrowhead Framework. The contributions can be summarized as development and evaluation of architectural patterns combining AAS and knowledge graphs; exploration of automated AAS generation and synchronization mechanisms; and identification of challenges related to scalability, governance, and interoperability in heterogeneous environments.

3.2.1.3 Methods used

The methodological approach combines the AAS concept with knowledge graph technology and the Arrowhead Framework to address interoperability challenges in heterogeneous industrial environments.

A key element is a layered architecture separating semantic intelligence from engineering representation (Figure 21). The Digital Twin Knowledge Graph constitutes the intelligence layer, where assets, services, and relationships are represented using semantic technologies such as RDF and OWL. This layer enables reasoning, consistency checking, and integration of heterogeneous data sources. The AAS layer acts as an engineering-oriented interface that exposes

relevant information through structured submodels and APIs, making it accessible and usable for engineers and applications. This separation allows complex and expressive semantic models to be maintained independently of how information is presented and consumed in engineering contexts.

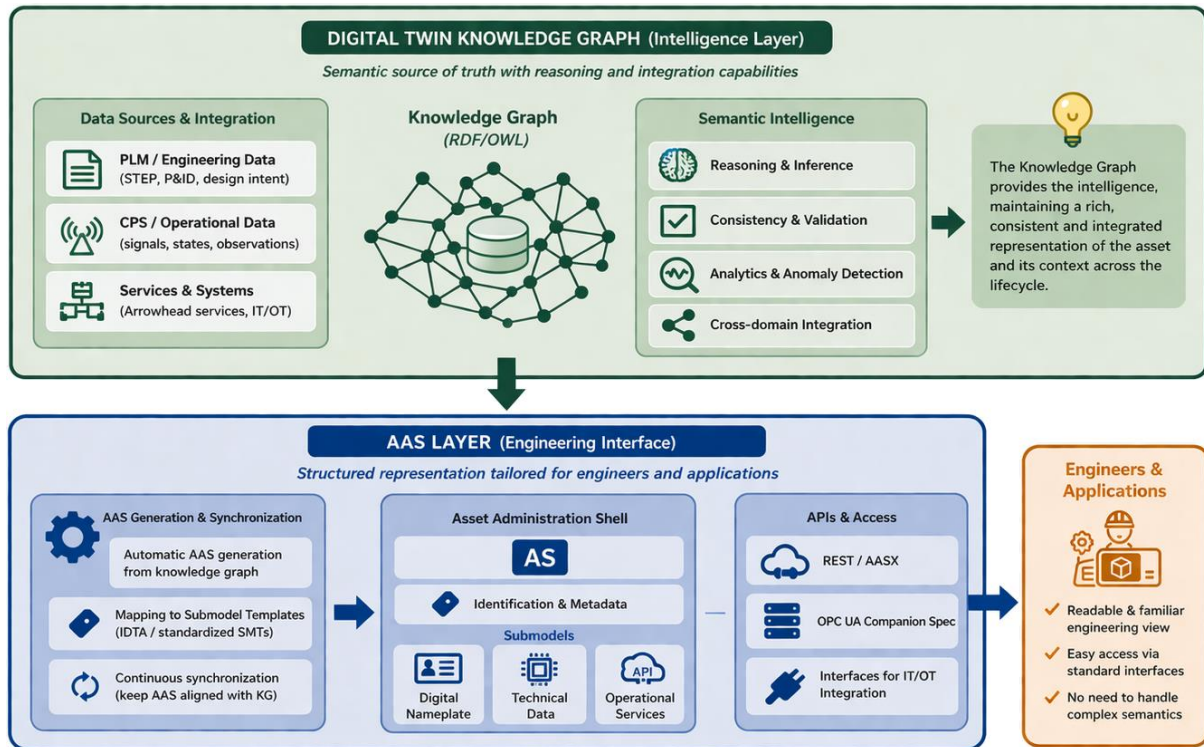


Figure 21. The AAS architecture separates the intelligence layer and the engineering interface

The implementation approach uses a knowledge graph as the primary source of truth for asset information. It then programmatically generates the AAS models from the knowledge graph. Synchronization mechanisms ensure consistency between the knowledge graph and AAS instances. The implementation uses GraphDB for semantic data management, and FA³ST Service as an AAS server. The approach aims to create a unified and coherent digital representation across the asset lifecycle by combining engineering data from PLM systems with operational data from cyber-physical systems.

3.2.1.4 Results

The results of this work are primarily conceptual and methodological, with limited experimental validation. The proposed architecture has been partially implemented and demonstrated in a controlled environment, but it has not yet been validated in large-scale industrial settings. The work indicates that a knowledge graph can function as a coherent and extensible source of truth for asset-related information, that the AAS models can be generated and maintained automatically from such a source – at least in principle –, and that the integration of semantic technologies with AAS has the potential to reduce manual configuration and improve interoperability.

A central architectural result is the clear separation between the intelligence layer and the engineering interface. The Digital Twin Knowledge Graph provides capabilities such as reasoning, validation, and integration of heterogeneous data. The AAS layer translates this into a

structured representation that is aligned with engineering practices and easier to understand and use. This architecture enables complex semantic information to be encapsulated and exposed in a form suitable for engineering workflows, bridging the gap between advanced semantic technologies and practical industrial usage.

A limited proof-of-concept implementation, including tests and a small-scale demonstration setup related to the Arrowhead Framework Ontology (AFO), indicates technical feasibility. Again, these results are based on controlled conditions and a limited number of assets. The results should therefore be interpreted as a validated architectural concept in a prototype-level implementation, not a fully tested or industrially validated solution.

3.2.1.5 Proposed further work

Given the limited validation, future work should focus on experimental verification, scalability, and practical deployment in industrial contexts. This includes extensive testing in larger and more heterogeneous industrial environments; evaluation of scalability and performance for large numbers of assets; alignment of knowledge graph models with AAS submodel templates and emerging standards; development of robust toolchains for automated AAS generation and lifecycle management; integration across full lifecycle scenarios, including engineering, operation, and maintenance; and establishment of governance models for data quality, versioning, and trust.

Further work should also focus on the layered architecture itself. In particular, research is needed to understand how the interaction between the Digital Twin Knowledge Graph as the intelligence layer and the AAS as the engineering interface can be standardized, scaled, and adopted in industrial practice. This includes improving how semantic insights are translated into engineering-relevant views, so that advanced capabilities such as reasoning, analytics, and decision support can be utilized without requiring deep expertise in semantic technologies.

3.2.2 DPP: Digital Product Passport

The Digital Product Passport (DPP) is a structured, digital representation of product-related information across the entire lifecycle, as introduced under the EU Ecodesign for Sustainable Products Regulation (ESPR). It enables transparent access to data such as materials, environmental impact, circularity, compliance, and lifecycle events.

Within the flexible product value network (fPVN) context, the DPP can be understood as a standardized, machine-readable view or report of product information, dynamically derived from distributed data sources rather than maintained as a centralized dataset. From an ISO 23726 perspective, the DPP represents an information-level projection built upon a shared semantic reference model (ontology). This ensures consistent meaning across domains and systems.

3.2.2.1 Scope of the standard

The purpose of the DPP is to support circular economy strategies (reuse, repair, recycling); comply with regulations and reporting; enable traceability across supply chains; and provide digital continuity of product information across lifecycle phases. The DPP spans multiple domains, including engineering and design, manufacturing and supply chain, operations and

maintenance, and end-of-life and recycling. Unlike traditional engineering standards such as STEP or DEXPI, the DPP is not limited to a specific lifecycle phase or domain. Instead, it provides a cross-domain representation of product-related information. The DPP does not define a single monolithic data model – instead, it relies on federated data sources, existing domain standards, and shared semantic definitions. Consequently, the DPP should be understood as a standardized representation layer that exposes information derived from semantically integrated sources.

3.2.2.2 Our work and cooperation with the standards organization

The Digital Product Passport is driven primarily by European Union regulation, in particular the Ecodesign for Sustainable Products Regulation (ESPR). The ESPR defines what information must be available, which product categories are affected, and how information shall be accessed and used.

Several organizations are involved in standardization activities supporting the DPP. The European Commission defines regulatory requirements and delegated acts. CEN/CENELEC develops harmonized European standards for implementation [30], while ETSI contributes to interoperability, identifiers, and data exchange. ISO and IEC provide domain standards and semantic frameworks. Relevant standards and initiatives include ISO 23726 (Industrial Data Ontologies), which establishes semantic reference model and ontology principles; ISO 10303 (STEP), which covers product structure and engineering data; DEXPI, which addresses plant and process data models; ISO 15926, which supports lifecycle data integration for process industries; and the W3C standards RDF, OWL, and SPARQL, which provide the foundation for knowledge graph technologies.

Our work aligns with and builds upon these efforts. We apply ISO 23726 principles to establish a shared ontology layer and represent domain knowledge as OWL 2 DL ontologies. Existing standards such as STEP and DEXPI are reused as structured data sources, and the resulting data is integrated into a federated knowledge graph within the fPVN. The DPP is then realized as a standardized, queryable view on this integrated data.

This approach supports a layered architecture in which each layer has a distinct role. EU regulation determines what information is required. Domain standards define how that information is structured. Ontologies define what the information means. The fPVN manages where the data resides and how it is connected across distributed systems. The DPP exposes a consistent and regulated view of this integrated information to external stakeholders.

3.2.2.3 Methods used

The method follows a layered semantic architecture consistent with ISO 23726 (Figure 22). At the top level, a shared ontology defines the concepts, relationships, and semantics that apply across all domains. This layer establishes a common understanding of core entities such as product, material, lifecycle event, and compliance attribute, ensuring that these terms carry the same meaning regardless of which domain or system uses them.

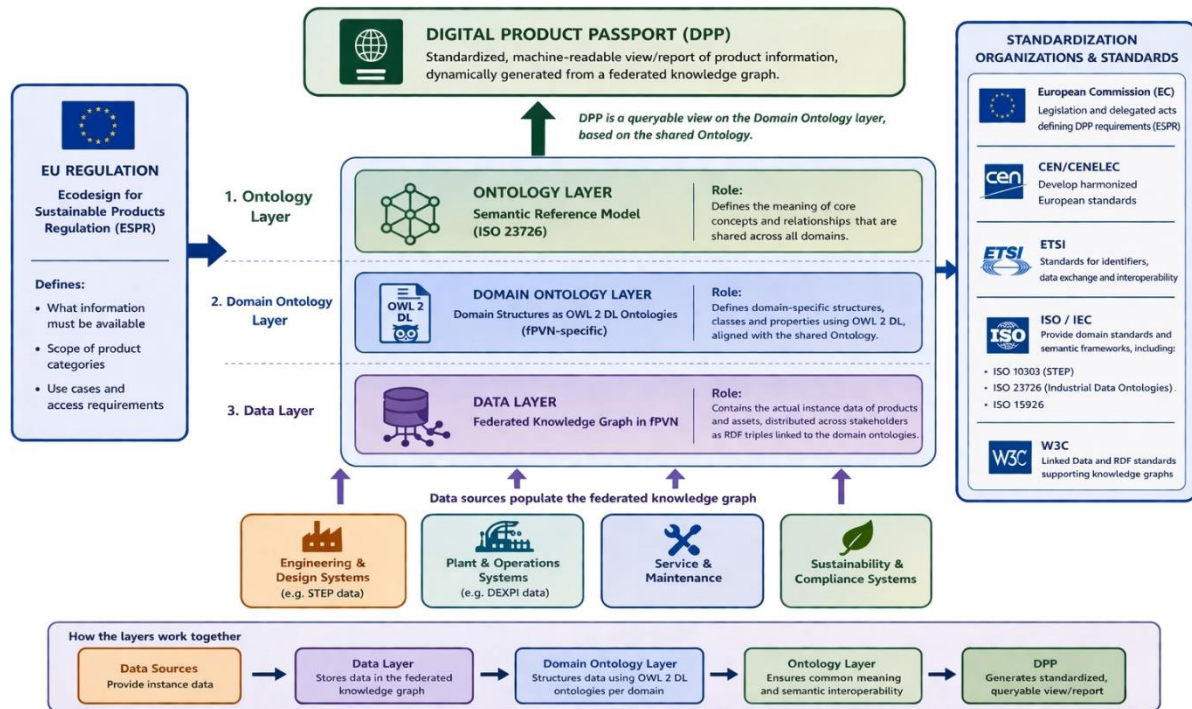


Figure 22. Relation of the DPP semantic layered architecture to EU regulation and standards organizations

Below this shared layer, domain-specific ontologies define structures that are particular to individual fields or standards. These are expressed as OWL 2 DL ontologies, which specialize and extend the shared ontology while remaining aligned with it. Ontologies derived from domain standards such as STEP and DEXPI serve as representative examples of this layer.

The lowest layer consists of instance data, represented as RDF triples and distributed across the systems and organizations that own that data. This data is not collected into a central repository. Instead, each data owner retains control of their own data, while the semantic structure provided by the ontology layers makes it possible to treat the distributed data as a coherent whole.

Integration across these layers involves mapping domain standards to the shared ontology, transforming structured data into RDF representations, and establishing semantic links between data from different domains. The result is a set of distributed but semantically connected datasets that can be queried as if they formed a single source.

The DPP is generated from this integrated data as a query-driven, machine-readable view, produced for example through SPARQL queries or dedicated APIs. It represents a projection of semantically integrated data across domains, formatted as a standardized representation aligned with EU regulatory requirements. Because the underlying data remains live and distributed, the DPP is dynamic and can be updated as product information changes, while remaining consistent for all stakeholders who access it.

3.2.2.4 Results

The work demonstrates that the DPP can be realized as a federated, ontology-based view rather than a centralized data repository. Concretely, this involved establishing a shared semantic reference model aligned with ISO 23726, defining domain ontologies using OWL 2 DL, and integrating heterogeneous data sources into a federated knowledge graph. Domain standards such as STEP and DEXPI were mapped to this common semantic framework, making it possible to generate consistent, machine-readable DPP representations from data that remains distributed across its original sources.

The approach yields several practical benefits. Because all domains share the same semantic foundation, data from different systems and organizations can be interpreted and combined without ambiguity, enabling interoperability across domains. Existing standards are reused rather than replaced or duplicated, which reduces the cost of adoption and preserves investments already made in domain-specific data models. Regulatory requirements are addressed directly through the structure of the semantic model, keeping the technical implementation aligned with EU obligations. Product information views are dynamic and can be updated as the underlying data changes, and the semantic links across the knowledge graph support traceability throughout the full product lifecycle.

Taken together, these results show that the DPP can serve as a standardized interface through which external stakeholders can access semantically integrated product data, without requiring that data to be extracted from its original context or consolidated into a single system.

3.2.2.5 Proposed further work

Several areas require further work to mature and standardize the approach. Domain standards need to be more fully harmonized with the shared ontology in accordance with ISO 23726, and formal governance arrangements for the ontologies and semantic models must be established. The mechanisms through which the DPP is queried and accessed, including APIs, require standardization. In federated environments, clear rules for access control, data ownership, and trust between participating parties also need to be defined. Finally, the approach should be validated through pilot implementations at industrial scale to confirm that it performs as intended under realistic conditions.

Continued collaboration with standardization bodies is equally important. Ontology-based approaches should be aligned with emerging European standards as they develop, and consistency between EU regulatory definitions and the underlying semantic models must be maintained. Active participation in standardization work can contribute to harmonized specifications for DPP implementation and help promote interoperability across industries and domains.

The overarching objective is to establish the DPP not as a standalone data model, but as a standardized view on distributed product data that is grounded in shared semantics and fully aligned with both regulatory requirements and industrial interoperability frameworks.

3.3 Standards WP3 originally not included in the work program

The two frameworks presented in this section – the Arrowhead Framework Ontology (AFO) and the Information Modelling Framework (IMF) – were not part of the original WP3 baseline. Their trajectories in the project, however, differ. AFO was developed during Arrowhead fPVN within WP4, then incorporated into the WP3 technical work and validated in Use Case 2.9 as a runtime ontology for local-cloud deployment topology and service dependencies. It therefore constitutes a project-validated result, despite not being planned in the initial work program. IMF remained at study level during the project period. It is architecturally relevant and compatible with the IDO-based stack, but was not integrated as a validated WP3 work item. The sections below reflect this distinction.

3.3.1 AFO: Arrowhead Framework Ontology

The Arrowhead Framework Ontology (AFO) is a formal OWL 2 ontology for representing the runtime structure of Arrowhead framework local clouds and the service dependencies through which their constituent systems interact [15], [17]. Developed within the Arrowhead fPVN project in WP4, it models local clouds, systems, husks (minimal micro-operating systems), unit assets, hosts, servers, provided services, consumed services, and selected Arrowhead framework core-system roles. In contrast to more generic industrial ontologies, the AFO is positioned at the *operational layer*: its purpose is not only to document concepts, but to capture the current deployment state of a local cloud as a machine-queryable knowledge graph.

AFO was developed as a WP4 prototype and subsequently incorporated into WP3, where it was applied in Use Case 2.9 "Digital Twin" as the semantic layer for representing the demonstrator's runtime structure and service dependencies. Initial alignment work with IDO as upper ontology has also been undertaken, positioning the AFO as a specialized runtime ontology that can participate in the broader semantic interoperability architecture pursued jointly by WP3 and WP4.

3.3.1.1 Scope of the standard

Although the AFO is not a formal standard, its scope is clearly delimited. It covers the representation of local-cloud deployment topology, service provision and consumption, core-system roles, and temporal deployment traces, including the core entity types introduced above and the constraints connecting them [15].

Arrowhead framework-specific semantic resources already existed prior to the AFO. The I40GO ontology stack includes dedicated ``ArrowHead System`` and ``ArrowHead Service`` modules, while the *Digital Reference Arrowhead* ontology covers classes such as ``AH Local Cloud``, ``AH System``, and ``AH Service``, together with core-system concepts and documentation relations such as ``is documented by`` and ``refers to``. These works provide valuable conceptual and documentation-alignment resources. What distinguishes the AFO is its modeling objective: rather than emphasizing conceptual coverage or design-description artifacts (interface design descriptions, service descriptions, system design descriptions), the AFO is positioned at the *operational layer*, targeting the current runtime deployment state of a local cloud as a machine-queryable knowledge graph [19].

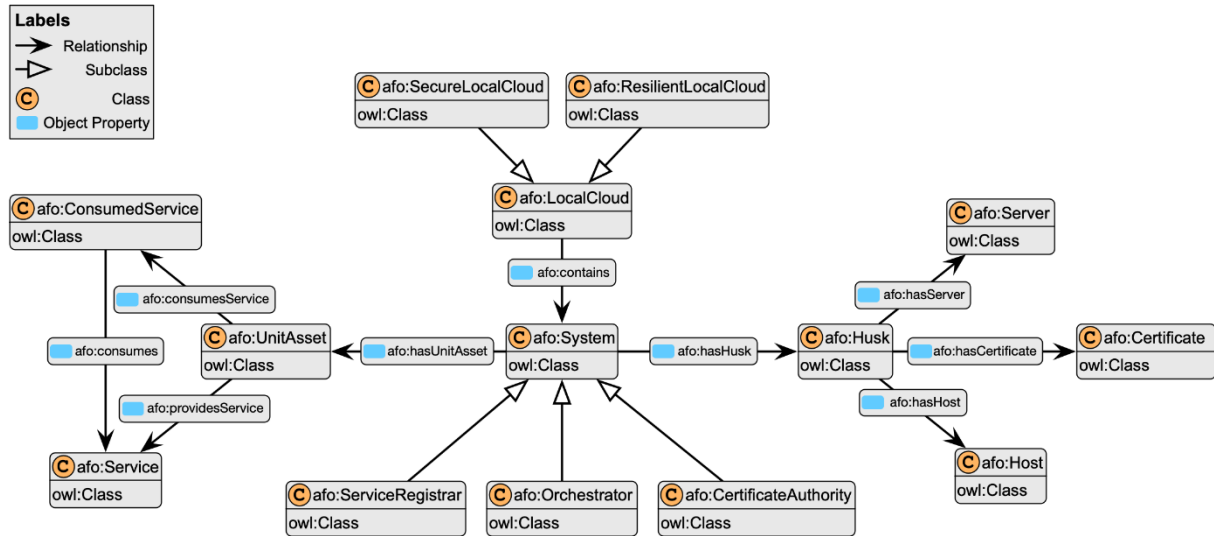


Figure 23. Graph view of the key AFO classes and object properties.
Source: Wintercorn and van Deventer [15]. Reproduced with permission.

The AFO does not aim to model detailed control logic, continuous telemetry streams, QoS constraints, complete authorization-policy semantics, or inter-cloud federation [19]. These concerns either belong to complementary standards and ontologies already relevant to WP3, or remain outside the intended scope of a deployment-topology ontology. This scoping is appropriate for WP3 because it keeps the AFO focused on what it does best: making the operational structure of an Arrowhead framework deployment explicit, queryable, and comparable over time. The key classes and object properties of the AFO are shown in Figure 23: `afo:LocalCloud` contains `afo:System` instances; each composed of exactly one `afo:Husk` and one or more `afo:UnitAsset` instances. Unit assets provide and optionally consume services. Subclasses `afo:ServiceRegistrar`, `afo:Orchestrator`, and `afo:CertificateAuthority` support automated classification of core-system roles through OWL axioms, though CA integration in the current mbaigo implementation remains partial.

3.3.1.2 Our work and cooperation with the standards organization

The AFO is not maintained by a formal standards organization. It is a research contribution developed within Arrowhead fPVN as part of WP4 by Oskar Wintercorn and Jan van Deventer at Luleå University of Technology [15], and subsequently incorporated into WP3. The ontology was developed in close alignment with the Arrowhead Framework architecture, and is accompanied by SHACL shapes for validation and by the KGrapher system for extracting runtime deployment data into RDF/Turtle snapshots [16].

KGrapher is itself implemented within the *Model-based Arrowhead framework implementation in Go* (mbaigo) [18], which applies Model-Based Systems Engineering (MBSE) to capture deployments as explicit models. This structured foundation makes it possible to automatically generate machine-interpretable knowledge graphs from live Arrowhead deployments, without modifying the underlying systems.

Within WP3, AFO was treated as an active work item and applied in Use Case 2.9 by Torbjörn Holm, Jan van Deventer, and Oskar Wintercorn, where it served as the semantic foundation for

representing the runtime structure and service dependencies of the use case demonstrator. In parallel, initial alignment work has been carried out with IDO as upper ontology.

3.3.1.3 Methods used

The AFO work followed a competency-question-driven ontology design approach [15]. The ontology was shaped around a set of operational questions that a deployment knowledge graph should be able to answer, including: which systems provide a given service; which consumed services are currently unsatisfied; on which host a given system's husk runs; whether a local cloud is resilient, degraded, or non-operational; and what the temporal history is of a device installed at a given functional location [15].

The operational pipeline begins with KGrapher, which queries the Service Registrar and the active systems in the local cloud, assembles an RDF/Turtle description of the current deployment, and stores that description in GraphDB as a timestamped named graph [15]. The two-stage query process – first retrieving the list of active systems from the Service Registrar, then issuing concurrent GET requests to each system – keeps snapshot generation fast even as the number of systems grows.

SHACL validation is then used as a conformance gate for structural and data-quality requirements, while OWL 2 RL reasoning materializes additional facts that improve queryability and completeness [15]. The combination is deliberate: OWL axioms provide schema-level semantics under the open-world assumption, enabling inference and classification, while SHACL provides a controlled closed-world check that can detect missing information and prevent silent drift in the instance model.

For deployment-specific context that should not alter the core AFO vocabulary, the approach introduces the Arrowhead Local Cloud extension ontology (``alc:`` prefix) as a stakeholder-defined layer. This layer classifies assets, rooms, functional locations, and other concepts specific to a given deployment. In the climate-control demonstrator, ``alc:Room`` represents sensor installation locations; in an industrial plant, analogous classes represent equipment tags or functional locations used in engineering and maintenance practice.

The AFO also formalizes operational status as OWL subclasses. A local cloud is classified as ``afo:ResilientLocalCloud`` when it contains at least two instances of each mandatory core system (*Service Registrar* and *Orchestrator*), enabling failover. If it additionally contains a *Certificate Authority*, it is classified as ``afo:SecureLocalCloud``. These subclasses specialize ``afo:LocalCloud`` with additional OWL axioms, enabling automated classification through reasoning rather than manual inspection [19].

A natural-language interface (Talk to Your Graph, TTYG) was integrated with GraphDB to expose competency questions grounded in the AFO through plain-language interaction. In the prototype, simple inventory-style queries could be translated into useful SPARQL, but the local LLM setup (Ollama with Llama 3.1) was not sufficiently reliable for operator-facing use in deeper dependency-traversal queries. This component is therefore reported as a preliminary usability exploration rather than a validated result [15].

3.3.1.4 Results

The reported results are best understood as a structured proof of concept rather than broad industrial validation. The demonstrator used was a climate-control local cloud consisting of at least six systems: *Service Registrar*, *Orchestrator*, *KGrapher*, *DS18B20*, *Parallax*, and *Thermostat*, deployed on Raspberry Pi hardware. The evaluation addressed five competency questions [19]:

CQ1: “Which systems provide a given service?” OWL-based inference was shown to derive additional knowledge about systems and services beyond what is asserted in the exported configuration: from 178 asserted statements, reasoning materialized an additional 54 inferred statements – a 31.4% increase – primarily through type propagation along class hierarchies, inverse-property completion, and rule-based service entailments.

CQ2: “Which consumed services are currently unsatisfied?” SHACL validation was shown to detect representative classes of deployment faults, including missing mandatory deployment links and metadata policy violations, and to localize them to actionable focus nodes and property paths in the validation report.

CQ3: “On which host does the husk of a given system run?” Edge deployability was confirmed: KGrapher generated a complete local-cloud snapshot in approximately ten milliseconds on a Raspberry Pi 5 (covering two stages of inter-system communication), OWL reasoning completed in less than one second on the same device, and ingestion of the resulting graph into GraphDB processed 2,000+ statements in less than one second [19].

CQ4: “What is the operational status of a local cloud (resilient/degraded/non-operational)?” The diagnostic capability was demonstrated through SPARQL queries that identify unsatisfied service dependencies and assess operational status at the SoS level – facts that are not obtainable from any single constituent system in isolation.

CQ5: “What is the temporal history of the device at a given functional location?” Digital thread reconstruction was demonstrated through timestamped named graphs, enabling traceability queries such as the functional location history of a sensor across replacement events.

A central architectural result is the clear separation between the *intelligence layer* (the Digital Twin Knowledge Graph, providing reasoning, validation, and heterogeneous data integration) and the *engineering interface* (the AAS or operator-facing views, exposing structured information in a form aligned with engineering practice). This separation is consistent with the layered architecture pursued more broadly in WP3.

The natural-language interface demonstrated that competency questions grounded in the AFO can be answered through plain-language interaction, though the prototype LLM setup (Ollama with Llama 3.1) was not yet sufficiently reliable for operator-facing use in deep graph-traversal queries [19].

	subject	predicate	object
1	urn:2026-02-20T14:07:38	rdf:type	afo:FunctionalLocation
2	urn:2026-02-20T14:07:38	alc:hasLocation	alc:Kitchen
3	urn:2026-02-20T14:07:38	afo:hasUnitAsset	alc:AlphaCloud_jeppcorn_ds18b20_28-00000f02986f
4	alc:AlphaCloud_jeppcorn_ds18b20_28-00000f02986f	rdfs:label	"28-00000f02986f"
5	urn:2026-02-23T08:41:22	rdf:type	afo:FunctionalLocation
6	urn:2026-02-23T08:41:22	alc:hasLocation	alc:Kitchen
7	urn:2026-02-23T08:44:37	rdf:type	afo:FunctionalLocation
8	urn:2026-02-23T08:44:37	alc:hasLocation	alc:Kitchen
9	urn:2026-02-23T08:44:37	afo:hasUnitAsset	alc:AlphaCloud_jeppcorn_ds18b20_28-00000f033350
10	alc:AlphaCloud_jeppcorn_ds18b20_28-00000f033350	rdfs:label	"28-00000f033350"

Figure 24. Result triples of a SPARQL CONSTRUCT query
Source: Wintercorn and van Deventer [15]. Reproduced with permission.

An additional result, relevant from the WP3 perspective, is that initial alignment work with IDO as upper ontology has begun to position the AFO concepts within the broader industrial semantic stack. Although still preliminary, this indicates that the AFO's deployment-oriented runtime knowledge can be connected to engineering lifecycle knowledge from other standards without losing its specific operational focus. The digital thread capability is illustrated concretely in Figure 24, which shows the result triples of a SPARQL CONSTRUCT query reconstructing the temporal history of a temperature sensor at a functional location, across three timestamped deployment snapshots. Highlighted rows correspond to two distinct physical sensors: the intermediate snapshot shows no `afo:hasUnitAsset` link, consistent with the period during which the sensor was offline.

3.3.1.5 Proposed further work

The demonstrator used for evaluation constitutes a proof of concept at small scale. Future work should focus on both broader validation and consolidation of the initial IDO alignment. This includes formalizing the IDO mappings, validating them for the Use Case 2.9 knowledge graphs, and using them to connect the AFO with the broader WP3 semantic architecture spanning STEP, DEXPI, CFIHOS, and AAS.

Further work should address validation in larger and more heterogeneous Arrowhead framework deployments, including multi-host and potentially multi-cloud scenarios, together with systematic evaluation of snapshot latency, graph growth, and reasoning overhead under more realistic industrial loads [15]. The natural-language interface should also be developed further, with appropriate validation mechanisms to ensure reliability for operator-facing use.

The AFO would benefit from a clearer governance model covering versioning, contribution rules, and compatibility with evolving Arrowhead framework implementations. Because deployment graphs reveal sensitive information about topology and service dependencies, any broader adoption should include work on filtered views, access-controlled query interfaces, and other mechanisms for controlled disclosure across organizational boundaries [15].

Because operational knowledge graphs make deployment topology and service dependencies machine-readable, they also expose sensitive business and operational information. Production

adoption of the AFO therefore requires strong authentication, authorization, and network segmentation around the knowledge graph repository and its query endpoints. Practical deployments should also consider separating internal operational graphs from filtered, purpose-specific views shared with external stakeholders for compliance or supply-chain collaboration. These security and controlled-disclosure mechanisms are identified as necessary prerequisites for broader industrial deployment.

From a standardization perspective, the architectural separation that the AFO enforces – between the operational knowledge graph as the intelligence layer and engineering or operator-facing views as the interface – maps naturally onto the layered architecture that has emerged from fPVN's broader work. The AFO could serve as an operational-runtime complement to the engineering lifecycle knowledge represented through STEP, DEXPI, and CFIHOS, connected through IDO as the common semantic reference layer.

3.3.2 IMF: Information Modelling Framework

The Information Modelling Framework (IMF) is a specification developed within the READI Joint Industry Project (JIP), with participants including TechnipFMC and Equinor, among others [10]. Development began in the early 2020s with the goal of transitioning the energy industry from document-centric information management to model-centric information management for industrial facility assets. The latest publicly available version is IMF Manual v0.3.0 (May 2024) [11], accompanied by a formal OWL ontology of the same version [12]. WP3 studied IMF as a potentially relevant framework, but IMF was not incorporated as a validated WP3 work item during the project period.

3.3.2.1 Scope of the standard

IMF addresses a fundamental problem in the engineering and operation of facility assets: information about an asset is fragmented across documents, disciplines, lifecycle phases, and supply chain actors. In current practice, the number of documents grows exponentially as a project progresses from early design through procurement, construction, and commissioning. Information is duplicated across these documents, relational context is lost, and the holistic description of the asset becomes increasingly difficult to maintain. When an asset transitions to operation, information often needs to be re-engineered from scratch. IMF is designed to replace this document-centric practice with incrementally built, machine-readable information models that accumulate rather than fragment knowledge [11].

The IMF approach rests on three pillars (Figure 25). The *Specification Model* resides in the engineering and subject matter expert (SME) domain. It is where engineers build a structured model of the asset using the ISO/IEC 81346 reference designation system, which organizes information according to functional, product (physical), and location aspects. The *Asset Model* resides in the ontology domain and provides the semantic integration layer that holds the three pillars together; this layer operates "under the hood" and is expressed as ontologies using W3C semantic technologies. The *Application Model* resides in the IT and digital domain, where legacy systems, engineering registers, and databases are connected to the model through International Resource Identifiers (IRIs). This three-part separation is deliberate: it minimizes disruption to both the engineering domain and the IT domain while enabling formal, machine-readable integration between them [11].

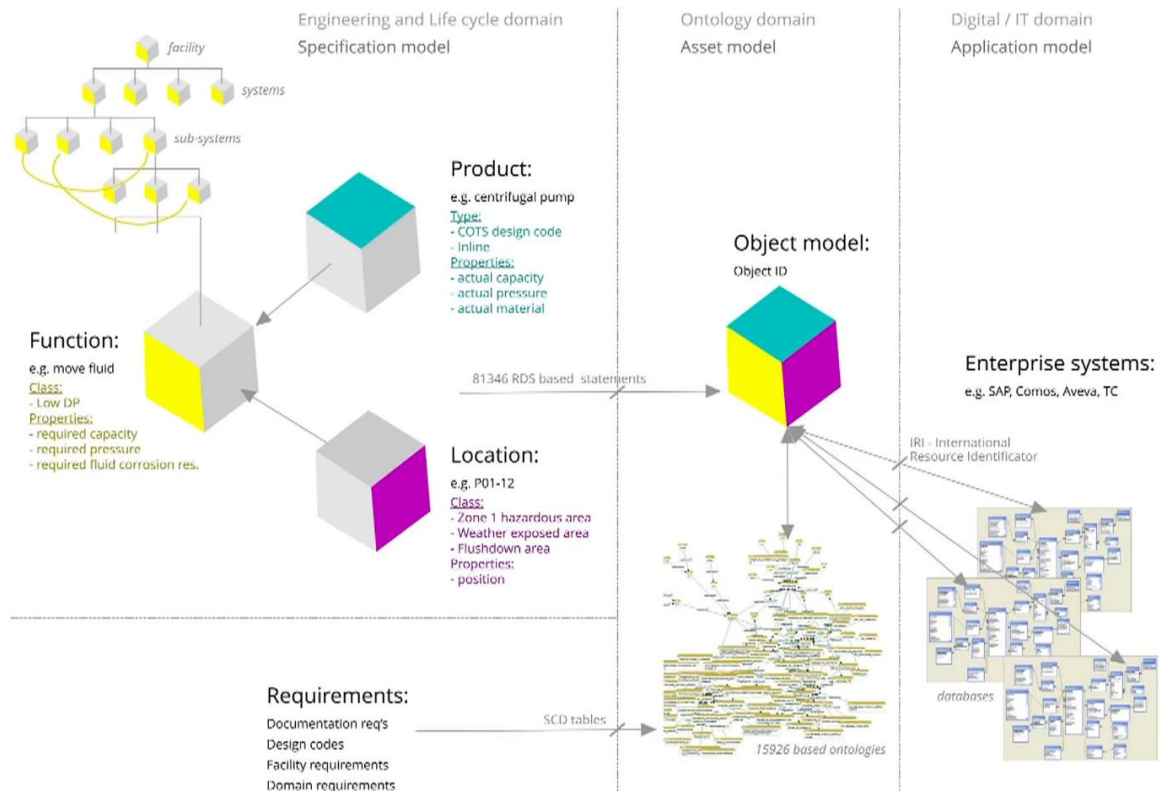


Figure 25. The IMF three-pillar architecture: Specification Model, Asset Model, Application Model

The IMF language defines three element types: *Block* (representing any system, assembly, or entity of interest), *Connector* (representing a topological connection between blocks and carrying information about the medium transferred), and *Terminal* (representing the point on a block's boundary where a medium may pass). Four aspects, aligned with the core aspects of ISO/IEC 81346, structure the modelling: the Function Aspect (‘=’, intended activity domain), the Product Aspect (‘-’, intended implementation domain), the Location Aspect (‘+’, intended space domain), and the Installed Aspect (‘:’, actual implementation, used to document physical assets as built and installed). Three fundamental relations organize the model: ‘partOf’ (hierarchical breakdown within an aspect), ‘connectedTo’ (topological connectivity and medium flow), and ‘fulfills’ (requirement–solution threading, which can cross aspect boundaries). Reusable *IMF Types* and *IMF Templates* are defined in reference data libraries and instantiated during modelling, supporting standardization and re-use of engineering knowledge [11].

IMF is formally grounded in several standards and technologies. The structuring principles follow ISO/IEC 81346-1 [13]. Reference data and semantic verification rely on ISO/DIS 23726-3 [14], which specifies the Industrial Data Ontology (IDO), designed to enable semantic interoperability across all life cycle phases of industrial assets. The IMF ontology is implemented using W3C recommendations: RDF for graph-based data representation, OWL for formal ontological axioms, and SHACL for shape-based validation [12]. The IMF language is designed to be available in four syntactic forms: a visual language for use by engineering SMEs in modelling tools; a formal RDF specification for knowledge engineers and application developers; an exchange protocol based on the Asset Administration Shell (AAS, IEC 63278), which is planned but not yet released; and a formal semantics expressed in first-order logic to support automated reasoning and integrity verification [11].

3.3.2.2 Our work and cooperation with the standards organization

WP3 identified IMF as a relevant initiative early in the project, primarily because of its shared reliance on semantic web technologies and its stated ambition to integrate lifecycle data for industrial assets using formal ontologies and reference data libraries. Both IMF and the Arrowhead fPVN approach target the same underlying problem – the fragmentation of asset information across organizational and lifecycle boundaries – and both rely on W3C standards (RDF, OWL, SHACL) as a common technical foundation. IMF's explicit use of ISO 15926-14 as a reference data and verification layer represented a direct point of contact with the IDO-based approach pursued in WP3.

Initial exchanges took place between WP3 participants and members of the READI JIP community to explore possible alignment. The IMF specification, its ontology, and its structuring principles were studied in the context of WP3's work on semantic interoperability. However, active cooperation did not progress to joint implementation within the Arrowhead fPVN project. Engagement between the two communities paused around the spring of 2025, and no formal contribution from WP3 to the IMF standardization process was made. The IMF ontology repository, which had been hosted on GitLab, was archived and migrated to GitHub in 2024, with the latest release (v0.3.0) dating from May 2024 [12]. As of the writing of this deliverable, version 1.0.0, which had been anticipated for the fall of 2024, has not been published.

3.3.2.3 Methods used

WP3's engagement with IMF did not proceed beyond a study phase. The methods applied were limited to reviewing the IMF specification and its supporting materials, including the IMF Manual v0.3.0 [11] and the OWL ontology published at the IMF namespace site [12]. The review assessed the conceptual architecture of IMF, its language constructs, and its integration with ISO 15926-14 and W3C semantic technologies, towards the goals and technical approach pursued in Arrowhead fPVN. Unfortunately, no joint modelling experiments, ontology alignment work, or implementation activities were carried out.

3.3.2.4 Results

The review did not produce joint technical results. It did, however, confirm that IMF and the IDO-based approach pursued in WP3 are not competing frameworks but address complementary concerns at different layers. IMF provides an engineering-facing modelling language – with structured vocabulary, aspect-based breakdown, and a visual syntax – designed to be used directly by subject matter experts. The IDO-based approach provides formal ontological foundations and reasoning capabilities for semantic interoperability across standards and systems. Notably, IMF explicitly depends on external reference data libraries for classifiers, attribute definitions, and medium types: roles that IDO-aligned ontologies, such as those developed within fPVN, could naturally fill. Similarly, IMF's RDF representation is designed to support integration with OWL-based verification tools – the same category of tooling that the fPVN architecture relies upon.

3.3.2.5 Proposed further work

There are concrete synergy points between IMF and the IDO-based approach that make re-engagement worthwhile when conditions allow. IMF's ``classifier`` references to external reference data libraries are an explicit integration point: IDO-aligned ontologies could serve as those libraries, giving IMF models a formally grounded semantic layer without requiring changes to the IMF language itself. IMF's RDF representation could be aligned with IDO ontologies to enable automated consistency checking and integrity verification of IMF models using standard OWL reasoning tools. IMF's ``fulfills`` relation, which threads requirement–solution chains across aspects and potentially across lifecycle phases, could enrich digital twin knowledge graphs with engineering traceability that complements the structured breakdown hierarchies already modelled in IDO. The planned IMF exchange protocol based on the Asset Administration Shell (AAS, IEC 63278) would further reinforce this potential, given the active AAS work within WP3.

From a standardization perspective, the architectural separation that IMF enforces – between the engineering domain, the ontology domain, and the IT domain – maps naturally onto the layered architecture that has emerged from fPVN's own work. IMF could serve as an engineering-facing interface to the IDO semantic layer, analogous to the relationship between the AAS and the Digital Twin Knowledge Graph described elsewhere in this deliverable. Future work should explore this integration in pilot scenarios involving facility asset lifecycle management, and consider contributing alignment findings back to the READI JIP and relevant ISO standardization activities.

4 Arrowhead fPVN use cases

The Arrowhead fPVN project includes industrial pilot use cases across several work packages. The two use cases described here – UC 2.9 (Digital Twin for pulp mills and carton board mills) and UC 3.9 (Interoperability for technical information exchange in process industry) – belong to Work Package 9 (Process industry) and were selected because they served as the primary industrial validation context for the WP3 standards work described in Section 3. The final results of both use cases are documented in D9.4 [9], which provides the complete demonstrator outcomes, partner contributions, and validation runs on which the summaries below draw.

4.1 Use Case 2.9: Digital Twin

Use Case 2.9, led by Stora Enso and WP9, developed a Digital Twin demonstrator for pulp mills and carton board mills. The technical work was carried out in four activities documented in D9.3 [33]: Activity 1 established the knowledge graph infrastructure, using GraphDB as the triple store and SPARQL as the query interface; Activity 2 transformed DEXPI P&ID data and STEP product data into RDF using the pipelines developed in WP3; Activity 3 integrated the Arrowhead Framework Ontology (AFO) to represent the local-cloud runtime structure and service dependencies of the demonstrator; Activity 4 demonstrated synchronization between live Arrowhead framework deployments and the knowledge graph through the KGrapher tool. The result was an end-to-end pipeline from engineering source data to a queryable semantic knowledge graph covering both plant engineering content and operational deployment state.

Use Case 2.9 was the primary industrial validation context for the WP3 standards work. The DEXPI-to-RDF and STEP-to-OWL transformations described in Sections 3.1.1–3.1.4 were applied and verified for UC 2.9 plant data, and the AFO alignment work described in Section 3.3.1 was carried out within the use case. An important finding, discussed further in Section 5, was that meaningful cross-standard query capability was achievable with a limited number of class-level IDO alignments. Full documentation is provided in D9.2 [32], D9.3 [33], and D9.4 [9] (all restricted to authorized project participants). D9.4 documents the final proof-of-concept results, including the six-activity pipeline from DEXPI/STEP source data to the GraphDB knowledge graph, the Cadmatic eShare 3D visualization interface, and the SAP Work Order integration via the Sinetiq Arrowhead implementation.

4.2 Use Case 3.9: Process industry

Use Case 3.9 [32] takes advantage of standardization and data harmonization across industry associations. Figure 26 presents the project way to refine and enrich process flow diagram information for standardization. VTT initiated research and technical analysis based on the existing DEXPI Process specification [1], and PLM-RDL2 [2], and the model information of Balas simulation software. Balas [7] is a product for pulp, paper and board applications. The information analysis was performed to prepare mappings between Balas calculation modules and DEXPI process steps and PLM-RDL2 process ontology classes. The main results were presented in Deliverable D3.2 [6]. The work is also internally validated with example models (Deliverable D4.2 [8]) and their transformations, documented in Deliverable D9.4 [9].

For inter-company validation the engineering company, AFRY, joined the work. In this step, the mappings were categorized and prioritized for standardization based on their usability in real plant engineering projects. Also, updates from the newer DEXPI 2.0 [3] standard were considered. The prioritized mappings are provided to the domain specific industry associations, THTH and SEIIA, to finalize a proposal for standardization. The industry associations also utilize their own background knowledge in the proposal finalization. PCA enables services and a platform for this effort. Thus, use case 3.9 is also a practical example of inter-industry cooperation to progress standardization.

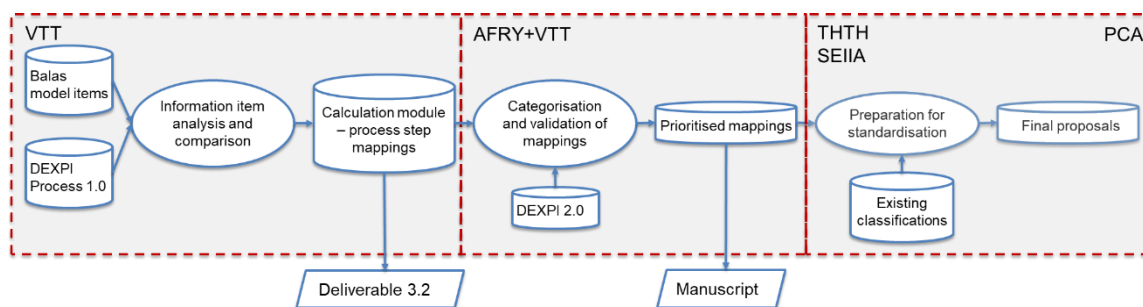


Figure 26. From a use case to standardization

4.2.1 Mapping and alignment of Balas model elements

Deliverable D3.2 presented in detail the mapping of Balas calculation modules to the corresponding concepts in DEXPI Process and PLM-RDL2. However, for serializing Balas models in a consistent way other Balas model items also need to be transformed to the standard formats.

This requires mapping of all relevant Balas model items to corresponding items in the standards, including calculation module parameters, streams, material components, and ports.

The mapping of the calculation module parameters to the corresponding standard items was not straightforward. Especially challenging was the mapping of Balas calculation module parameter to corresponding items in DEXPI Process. The main challenge was that Balas calculation module parameters can be completely different set from the attributes of a *Process Step*. Therefore, a workaround was used to account for the parameters when serializing in the DEXPI Process format. In the implementation, a Balas specific XML namespace was defined for Balas parameters (DEXPI Process models are serialized in XML format).

To account for Balas calculation module parameters in PLM-RDL2, the relevant ontology was extended with Balas calculation module parameter specific classes. The different parameter classes were aligned with different subclasses of the *ScalarQuantityDatum* class using the *subclassOf* relation. In analysis, the subclass relations were documented using MS Excel. The most suitable *ScalarQuantityDatum* subclass was identified based on the unit of measure (UOM). Table 4 lists an example set of Balas parameters and the aligned PLM-RDL2 class.

Out of the ca. 750 Balas parameters ~410 could be aligned with a suitable *ScalarQuantityDatum* class based on UOMs. Parameters that do not have a UOM were linked to dimensionless property datum class (see, e.g., parameter `Max control flow ratio` in Table 4). Parameters that could not be linked based on UOM were aligned directly with the *ScalarQuantityDatum* class. The final validation runs of the adapter software against five example models of varying complexity, together with the complete calculation module transformation table, are reported in D9.4 [9].

*Table 4. PLM-RDL2 mapping to Balas calculation module parameters.
The bold rows highlight some limitations in the supported UOMs of PLM-RDL2.*

<i>Balas parameter</i>	<i>Balas UOM</i>	<i>PLM-RDL2 class</i>	<i>PLM-RDL2 UOM</i>	<i>Conversion factor</i>
Heat losses	%	dimensionless property datum	%	
Heat losses	kW	power datum	kW	
Leakage air in	%	dimensionless property datum	%	
Heat loss coeff	%	dimensionless property datum	%	
Pressure drop	kPa	pressure datum	kPa	
Cake consistency	%	dimensionless property datum	%	
Solids loss	%	dimensionless property datum	%	
Temperature drop	%	dimensionless property datum	%	
Dry content	%	dimensionless property datum	%	
Clear filtrate	%	dimensionless property datum	%	
Outlet consistency	%	dimensionless property datum	%	
Cloudy filtrate	%	dimensionless property datum	%	
Efficiency	%	dimensionless property datum	%	
Radiation losses	%	dimensionless property datum	%	
Radiation losses	kW	power datum	kW	

<i>Balas parameter</i>	<i>Balas UOM</i>	<i>PLM-RDL2 class</i>	<i>PLM-RDL2 UOM</i>	<i>Con-version factor</i>
Solids loss to vapour	%	dimensionless property datum	%	
Coefficient of efficiency of motor	%	dimensionless property datum	%	
Dilute bypass	%	dimensionless property datum	%	
Main circulation bypass	%	dimensionless property datum	%	
Leak tolerance	%	dimensionless property datum	%	
Not in use	%	dimensionless property datum	%	
Condensate loss	%	dimensionless property datum	%	
Approach temperature	°C	thermodynamic temperature datum	°C	
Web temperature	°C	thermodynamic temperature datum	°C	
Exhaust air temp.	°C	thermodynamic temperature datum	°C	
Condensate temp.	°C	thermodynamic temperature datum	°C	
Exit temperature	°C	thermodynamic temperature datum	°C	
Exhaust air temperature	°C	thermodynamic temperature datum	°C	
Outlet temperature	°C	thermodynamic temperature datum	°C	
Temperature	°C	thermodynamic temperature datum	°C	
Heat demand temperature	°C	thermodynamic temperature datum	°C	
Depth	m	length datum	m	
Surface area	m ²	area datum	m ²	
Heat transfer area (optional)	m ²	area datum	m ²	
Heat transfer surface area	m ²	area datum	m ²	
Pressure	kPa	pressure datum	kPa	
Deculator pressure	kPa	pressure datum	kPa	
Grinder pressure	kPa	pressure datum	kPa	
Exit pressure	kPa	pressure datum	kPa	
Condensate pressure	kPa	pressure datum	kPa	
Thermal duty	MW	power datum	MW	
Effect	MW	power datum	MW	
Heat flow	kW	power datum	kW	
Flowrate	kg/s	mass flow datum	kg/s	
Total flow	flow/s	mass flow datum	kg/s	
Permeate flow	kg/s	mass flow datum	kg/s	
Heat transfer coefficient	kW/m ² K	heat transfer coefficient datum	W/m ² K	1000
Overall K (optional)	kW/m ² K	heat transfer coefficient datum	W/m ² K	1000
Specific energy consumption	MWh/ADt ¹⁾	ScalarQuantityDatum		
Air demand	m³air/ m³liq- uid	ScalarQuantityDatum		
Flux	L/m²h	ScalarQuantityDatum		

<i>Balas parameter</i>	<i>Balas UOM</i>	<i>PLM-RDL2 class</i>	<i>PLM-RDL2 UOM</i>	<i>Conversion factor</i>
Exhaust air moisture	g/kg	ScalarQuantityDatum		
Norden efficiency factor, E		dimensionless property datum		
Split fraction for 1:st outlet		dimensionless property datum		
Max control flow ratio		dimensionless property datum		
Sweetener stock ratio		dimensionless property datum		
Set value 1		dimensionless property datum		

4.2.2 Mapping of streams, ports and material components

The mapping of the other model items was quite straightforward. In DEXPI Process, Balas streams can be represented with the `Stream` class (a subclass of `Process Connection`). The Balas stream class (defines the structure of a stream) was mapped to `Material Template` and Balas port to `Material Port`. Balas material components that are mapped to the `Material Component` class.

In PLM-RDL2, a Balas stream is mapped to the class `Stream`. Balas ports are mapped to either `Inlet` or `Outlet` classes. Material components are represented using the stream components modelling pattern presented in [5].

4.2.3 Mapping to Chemical Entities of Biological Interest identifiers

To transform individual Balas material components unambiguously, they were mapped to Chemical Entities of Biological Interest (ChEBI) [4] identifiers. The mapping was based on the chemical formula of a material component. Self-defined Balas material components used for describing natural compounds like softwood and lignin do not have standardized ChEBI identifiers.

This mapping has an impact on the serialization of Balas models. In DEXPI Process, material components that are mapped to a ChEBI identifier are serialized as `Pure Material Components` (a subclass of `Material Component`), and other components as `Custom Material Components` (a subclass of `Material Component`).

For PLM-RDL2, the modelling pattern presented in [5] accounts also for ChEBI identifiers. If a component is not mapped to a ChEBI identifier the `chemical` object property is not included in the serialization.

4.2.4 The applicability of the PLM-RDL2 and DEXPI Process standards

The applicability of the standards was tested using five example models (see also D9.4 [9] for the final validation results, including deserialization back to Balas-specific format and the extended Arrowhead framework web API). The reader is referred to Deliverable D4.2 [8] for detailed information on the example models. Each model was serialized both in DEXPI Process and in PLM-RDL2 formats using an adapter software developed by VTT.

The serialization of the test models illustrated the same challenges as the actual mapping presented in Deliverable D3.2 [6], namely, the mapping of Balas calculation modules is not always one-to-one. PLM-RDL2 and *DEXPI Process* do not include suitable concepts for mapping power production modules or pulp and paper industry specific modules. Additionally, the PLM-RDL2 process ontology is missing a `Drying` class, even though the equipment-ontology includes a `Dryer` class.

Correspondingly, most of the calculation module parameters that were linked directly to the PLM-RDL2 `ScalarQuantityDatum` class were related to pulp and paper and drying specific modules. An additional issue related to the DEXPI Process was that some Balas calculation modules can have more than one function. For example, heat exchanger-modules can perform both cooling and heating functions but they can be serialized only as either `RemovingThermalEnergy` class or `SupplyingThermalEnergy` class (or one of their subclasses). In the serialization of other model items, no issues were discovered.

DEXPI 2.0 addresses many of the limitations experienced in the conversion of Balas calculation modules and their parameters. For example, DEXPI 2.0 allows specifying new process steps, and to define parameters. PLM-RDL2 already enables this.

5 Discussion

The work reported in this deliverable demonstrates that ontology-based semantic interoperability across heterogeneous industrial standards is technically feasible and delivers measurable value at both the schema and instance data levels. The results, however, also surface a number of tensions, limitations, and open questions that deserve reflection.

5.1 The alignment architecture: strengths and limits

The star-topology alignment strategy – in which each standard is aligned to IDO as a shared semantic center rather than mapped to every other standard individually – proved to be technically sound and practically manageable. A finding from UC 2.9 is that significant interoperability benefits were achievable with only a small number of class-level alignments (~2% of classes). This is encouraging: it confirms that early value does not require exhaustive mapping, and that the architecture rewards incremental investment. At the same time, the use cases operated at a scale appropriate for a research project. Confidence in the approach at full industrial scale – complete plant models, full lifecycle histories, live operational data – remains to be established.

The alignment approach does not modify the normative content of any aligned standard. This is a deliberate design choice and an important one for adoption: organizations that have already invested in STEP, ISO 15926, DEXPI, or CFIHOS need not abandon or renegotiate those investments. The alignment ontologies are maintained as separate files. The practical consequence, however, is that governance of the alignment ecosystem is distributed: as IDO editions are revised and as the aligned standards evolve, the alignment files require coordinated updates. A governance framework for versioning and change impact assessment across this family of alignment ontologies is identified as a necessary prerequisite for sustained industrial adoption, and does not yet exist.

5.2 Maturity varies across standards

The six standards addressed in Section 3.1 are not at the same level of maturity in their engagement with ontology-based interoperability. ISO 23726 (IDO) reached FDIS status during the project and is supported by a growing industrial consortium. The STEP Core transformation produced a fully automated pipeline and a formally validated OWL artifact. DEXPI achieved an end-to-end pipeline from P&ID documents to knowledge graphs in RDF. CFIHOS Semantics delivered a fully automated generation pipeline covering approximately 500 classes. ISO 15926 demonstrated cross-industry applicability but remains at an earlier stage of formalization for the extended reference data. SysML v2 alignment remains exploratory, with the transformation pipeline relying in part on LLM-assisted workflows whose reliability has not been systematically validated.

This variability reflects the different starting points, available resources, and community readiness encountered across the standards. It is not a weakness of the overall approach, but it does mean that the interoperability architecture, while coherent at the conceptual level, is unevenly developed at the operational level. Investment in the less mature alignments – particularly SysML v2 and ISO 15926 extended reference data – is needed before the full cross-standard querying capability can be realized in production settings.

5.3 The role of tooling

A recurring theme across the work is the importance of tooling for adoption. The CFIHOS Python pipeline, the STEP XMI-to-OWL transformer, and the Semantum Model Broker all represent concrete toolchain contributions that make the alignment approach repeatable and maintainable. Without such tooling, even a conceptually correct alignment is not practically usable at industrial scale. The investment in AsciiDoc-based documentation infrastructure and GitHub hosting for the Arrowhead fPVN Handbook, Patterns, Use Cases, and Guidelines is a complementary form of tooling: it ensures that the knowledge accumulated in the project is accessible, versionable, and maintainable beyond the project's lifetime [38].

Tooling gaps remain, particularly in the integration of IDO-aligned ontologies into mainstream engineering platforms (AVEVA, Intergraph, Bentley), in the provisioning of SPARQL endpoints for plant knowledge graphs, and in SHACL-based quality gates embedded in engineering workflows. These gaps are well understood and technically tractable; they represent engineering implementation work that is now unblocked by the project's results.

5.4 Semantic web technologies in an industrial context

The choice of W3C semantic web technologies – OWL 2 DL, RDF, SPARQL, SHACL – as the common formalism across all standards transformations reflects a deliberate positioning. These technologies provide formal semantics, machine-readable vocabularies, web-scale distribution through dereferenceable IRIs, and a well-maintained open tooling ecosystem. The decision to target OWL 2 DL specifically – the decidable sublanguage – ensures that automated reasoning remains tractable and that consistency checking is tool-independent.

A tension inherent in this choice is the open-world assumption of OWL, which is appropriate for inference and knowledge integration but can be counterintuitive for engineers accustomed to closed-world database or schema-driven thinking. The complementary use of SHACL, which operates under the closed-world assumption, directly addresses this tension: SHACL provides the validation behavior that engineers expect, while OWL provides the inferential power that enables cross-standard reasoning. The project's consistent application of both – OWL for schema-level alignment and inference, SHACL for data quality gates – is a methodological contribution in its own right, and one that may be generalized beyond the specific standards addressed here.

5.5 Standardization engagement

Active engagement with standardization bodies – ISO/TC 184/SC 4, DEXPI e.V., CFIHOS, OMG, and CEN/CLC/JTC 24 – was a distinguishing feature of the WP3 approach and a necessary condition for the results achieved. Several technical contributions from the project directly influenced the content of emerging standards: the DIS ballot resolution for ISO 23726-3, the CFIHOS Feature Requests for IDO and other mappings, and the positioning discussions at SC 4 regarding ISO 10303 and the role of OBI as a normative reference for mutual content. These engagements ensure that the project's technical results are not isolated research outputs but have pathways into durable standards infrastructure.

5.6 Toward industrial deployment

Taken together, the results of D3.3 represent a transition point: from demonstrating feasibility to establishing the preconditions for industrial deployment. The alignment architecture is sound. The toolchains are operational. The documentation infrastructure is in place. The governance mechanisms for long-term maintenance are identified but not yet formalized. The validation use cases are confirmed but limited in scale. The immediate next phase – toolchain integration into engineering platforms, SHACL-based quality gates in engineering workflows, and governance framework definition for alignment ontology maintenance – is technically unblocked and represents a tractable set of engineering and standardization tasks.

6 Conclusion

This deliverable completes the WP3 deliverable series of the Arrowhead fPVN project. D3.1 surveyed and catalogued the major standardized industrial data models; D3.2 demonstrated that semantic web technologies can extend and bridge them; D3.3 translates those demonstrations into concrete proposals for how the standards themselves should evolve, and documents the project's direct contributions to the relevant standardization processes. The central claim of the deliverable – that ontology-based semantic interoperability across heterogeneous industrial standards is technically feasible and delivers measurable value – has been established at both the schema and instance data levels, through four completed alignment ontologies and two operationally validated use cases.

The Industrial Data Ontology (ISO 23726-3) proved to be a viable and stable upper ontology for the alignment work. Its progression from Technical Report to Draft International Standard to FDIS v4.2 (50 classes, 102 object properties, approved December 2025) during the project period gave the work a solid and increasingly normative foundation. The star-topology alignment strategy – in which each standard aligns once to IDO rather than to every other standard individually – is the architectural principle that makes this tractable. An important empirical finding, confirmed in both UC 2.9 and UC 3.9, is that significant cross-standard query capability is achievable with a small number of class-level alignments. Value is therefore not deferred until exhaustive mapping is complete; it accrues incrementally from the first alignments and grows as more are added.

The six standards addressed in Section 3.1 are not uniformly mature in their alignment with the IDO-based architecture. The STEP Core transformation pipeline is fully automated and formally validated. The CFIHOS Semantics pipeline covers approximately 500 classes. DEXPI achieves end-to-end transformation from P&ID documents to RDF knowledge graphs. ISO 15926 demonstrates cross-industry applicability of reference data. SysML v2 integration remains exploratory. This variability is not a weakness of the approach, but it is an honest reflection of where investment has been concentrated and where further work is needed. The proposals for further work in each standard section identify the specific next steps required to bring the less mature alignments to operational quality.

Active engagement with standards bodies – ISO/TC 184/SC 4, DEXPI e.V., CFIHOS, OMG, and CEN/CLC/JTC 24 – was a distinguishing feature of WP3's approach and a prerequisite for the lasting impact of the results. Several project contributions directly shaped the content of emerging standards: the resolution of DIS ballot comments for ISO 23726-3, feature requests

to the CFIHOS Semantics team for IDO alignment, and positioning discussions at SC 4 regarding the role of ISO 23726 as a normative reference for mutual content across the ISO 10303 family. These contributions ensure that the technical results of the project have pathways into durable standards infrastructure, rather than remaining isolated research outputs.

The documentation artifacts produced in the project – the Arrowhead fPVN Handbook, Patterns Library, Use Cases, and Guidelines all authored in AsciiDoc and hosted in a public repository under the POSC Caesar organization – constitute a reusable knowledge base for IDO-based ontology engineering that will outlive the project. POSC Caesar Association (PCA) has been identified as the long-term maintenance agency. This governance decision is significant: it places the continuity of the guidelines work in the hands of an established standards body with a direct mandate for industrial data standards, rather than with any single industrial organization.

The results of D3.3 mark a transition point: from demonstrating that ontology-based interoperability is possible to establishing the preconditions for industrial deployment. The alignment architecture is sound. The transformation toolchains are operational. The documentation infrastructure is in place. The immediate next phase – integration of IDO-aligned ontologies into mainstream engineering platforms, SHACL-based quality gates embedded in engineering workflows, and a formalized governance framework for alignment ontology maintenance and versioning – is technically unblocked by the work reported here. These are tractable engineering and standardization tasks. The Arrowhead fPVN project delivers them as a starting point, not a finishing line.

7 References

- [1] DEXPI Process Specification, DEXPI Process Information Model 1.0, DEXPI e.V., Frankfurt 2023.
- [2] POSC Caesar Association, 2025. Reference Data Library: Industrial data ontology & Product Lifecycle Management.
<https://rds.posccaesar.org/ontology/plm/> (Accessed 2026.03.18).
- [3] DEXPI Specification, DEXPI P&ID Specification 2.0, DEXPI e.V., Frankfurt 2025.
- [4] ChEBI, Chemical Entities of Biological Interest, 2025.
<https://www.ebi.ac.uk/chebi/> (2025.10.23).
- [5] POSC Caesar Association, 2025. Modelling patterns.
<https://rds.posccaesar.org/doc/patterns/> (Accessed 2026.03.18).
- [6] Arrowhead fPVN, Deliverable D3.2 “Major industrial data model semantics analysis” 2025 June. <https://fpvn.arrowhead.eu/wp-content/uploads/sites/3/2025/11/D3.2-Major-industrial-data-model-semantics-analysis-V1.3.pdf> (Accessed 2026.04).
- [7] VTT BALAS – Conceptual process design to optimize resource efficiency. VTT Technical Research Centre of Finland Ltd.
<https://www.vttresearch.com/en/explore/vtt-balas-conceptual-process-design-optimize-resource-efficiency> (Accessed 2024.10.28)
- [8] Arrowhead fPVN, Deliverable D4.2: Second generation of translation technology. 2024.11.25. <https://fpvn.arrowhead.eu/deliverables/>
- [9] Arrowhead fPVN, Deliverable D9.4: Process industry Use case final results. 2026.05.31. (C-UE/EU-C only. Not publicly available.)
- [10] READI Joint Industry Project. Asset Information Modelling Framework.
<https://readi-jip.org/asset-information-modelling-framework/> (Accessed 2026.04).
- [11] Fjøsna, E. et al. Information Modelling Framework Manual, Version 0.3.0. READI JIP, 2024-05-29. <https://ns.imfid.org/> (Accessed 2026.04).
- [12] IMF Lab. IMF Semantic Technologies Implementation, v0.3.0, 2024-05-31.
<https://github.com/imf-lab/imf-ontology-semweb> (Accessed 2026.04).
- [13] IEC/ISO 81346-1:2022. Industrial systems, installations and equipment and industrial products – Structuring principles and reference designations – Part 1: Basic rules. ISO/IEC, 2022.
- [14] ISO/DIS 23726-3 Automation systems and integration – Ontology-based interoperability. Part 3: Industrial data ontology. Under development.
<https://www.iso.org/standard/87560.html> (Accessed 2026.04.27)
- [15] Zenodo, 2026. The Arrowhead Framework Ontology.
<https://zenodo.org/records/18731224> (Accessed 2026.04)
- [16] J. van Deventer, “sdoque/systems,” GitHub repository.
<https://github.com/sdoque/systems> (Accessed 2026.04.24).
- [17] O. Wintercorn, “oskwin/Arrowhead-Framework-Ontology,” GitHub repository, version 1.1.0, 2026. <https://github.com/oskwin/Arrowhead-Framework-Ontology> (Accessed 2026.04.24).
- [18] J. van Deventer, “A model based implementation of an IoT framework”, TechRxiv, 2025.04.
<https://www.techrxiv.org/doi/full/10.36227/techrxiv.174195758.81241062/v1>
- [19] Oskar Wintercorn and Jan van Deventer, "The Arrowhead Framework Ontology," IEEE Open Journal of the Industrial Electronics Society, submitted 2026.

- [20] ISO (n.d.). ISO/TC 184/SC 4: Industrial Data. International Organization for Standardization. <https://www.iso.org/committee/54158.html> (Accessed 2026.05).
- [21] POSC Caesar Association. (n.d.). Industrial Data Ontology (IDO). <https://rds.posccaesar.org/ontology/lis14/> (Accessed 2026.05).
- [22] POSC Caesar Association. (n.d.). OBI - IDO training sessions 2026, April 30th: IDO: Industrial data ontology. <https://www.posccaesar.org/articles/new-obi---ido-training-sessions-2026> (Accessed 2026.05).
- [23] POSC Caesar Association. (n.d.). PCA Reference Data Services: Downloads. <https://rds.posccaesar.org/services/downloads/> (Accessed 2026.05).
- [24] DEXPI e.V. (n.d.). DEXPI – Data Exchange in the Process Industry. <https://dexpi.org/> (Accessed 2026.05).
- [25] JIP36 CFIHOS. (2025). CFIHOS Standards – Capital Facilities Information Handover Specification, Version 2.0. <https://www.jip36-cfihos.org/cfihos-standards/> (Accessed 2026.05).
- [26] Object Management Group (OMG). (n.d.). Object Management Group. <https://www.omg.org/> (Accessed 2026.05).
- [27] Object Management Group (OMG). (2025). OMG Systems Modeling Language (SysML), Version 2.0. Formal specification formal/26-03-02. <https://www.omg.org/spec/SysML/2.0> (Accessed 2026.05).
- [28] IEC. (n.d.). IEC/TC 65: Industrial-process measurement, control and automation. International Electrotechnical Commission. <https://www.iec.ch/technical-committee/65> (Accessed 2026.06).
- [29] IEC. (2023). IEC 63278-1 ED1: Asset Administration Shell for Industrial Equipment – Part 1: Asset Administration Shell Structure. International Electrotechnical Commission. See also: Industrial Digital Twin Association (IDTA). <https://industrialdigitaltwin.org/en/> (Accessed 2026.06).
- [30] CEN-CENELEC. (n.d.). JTC 24 – Digital Product Passport: Framework and System. European Committee for Standardization / European Committee for Electrotechnical Standardization. <https://www.cencenelec.eu> (Accessed 2026.06).
- [31] Arrowhead fPVN, Deliverable D3.1 “Major industrial data models” 2023 November. <https://fpvn.arrowhead.eu/wp-content/uploads/sites/3/2024/10/Arrowhead-fPVN-D3.1-1.pdf>
- [32] Arrowhead fPVN, Deliverable D9.2: Process industry use case first year progress and next step specification. 2024.05.31. (C-UE/EU-C only. Not publicly available.)
- [33] Arrowhead fPVN, Deliverable D9.3: Process industry use case second year progress and final specification. 2025.05.29. (C-UE/EU-C only. Not publicly available.)
- [34] ECLASS e.V. (n.d.). ECLASS – The Global Standard for Product Classification and Description. <https://eclass.eu/en> (Accessed 2026.06).
- [35] IEC. (n.d.). Common Data Dictionary (CDD). International Electrotechnical Commission / TC 3. <https://tc3.iec.ch/tc-activity/common-data-dictionary-cdd/> (Accessed 2026.06).
- [36] SousLeSens. (n.d.). SousLeSens – Ontology Editor. GitHub repository. <https://github.com/souslesens> (Accessed 2026.06).
- [37] IOGP JIP33. (n.d.). JIP33 Procurement Data Sheets. International Association of Oil & Gas Producers. <https://jip33.iogp.org/> (Accessed 2026.06).
- [38] POSC Caesar Association. (n.d.). The Arrowhead fPVN Handbook GitHub repository. https://github.com/PCA-POSC-Caesar-Association/AfPVN_Handbook (Accessed 2026.06).

- [39] POSC Caesar Association. (n.d.). The Arrowhead fPVN Handbook on industrial semantics – Using IDO as upper ontology. <https://docs.posccaesar.org/libraries-and-standards/arrowhead-handbook/> (Accessed 2026.06).

8 Revision history

Ver- sion	Date	Contributor	Contributions
1.0	2026.05.28	Péter Olszi	First released version. Same contents as v0.49.

8.1 Contributing and reviewing partners

Representing partner	Contributor	Contents	Review
PCA	Nils Sandmark	X	
PCA	Magne Valen-Sendstad	X	
AITIA	Péter Olszi	X	X
Fluor/CFIHOS	Onno Paap	X	X
TBHK	Torbjörn Holm	X	
PTC	Géza Kulcsár	X	
Independent/DEXPI e.V.	Heiner Temmen	X	
Independent	Andreas Neumann	X	
VTT	Teemu Mätäsniemi	X	
VTT	Kim Björkman	X	
LTU	Oskar Wintercorn	X	X
LTU	Jan van Deventer	X	X
SEIIA	Erik Molin	X	X
THTH	Arto Martinen	X	X
LTU	Carl Borngrund		X

8.2 Amendments

No.	Date	Version	Subject of Amendments	Author

8.3 Quality assurance

No	Date	Version	Approved by
1	2026.05.29	1.0	Jerker Delsing