

A Semantic-Provenance Framework for Interoperability of Design-Centric Industrial Software

ISSN: 2639-0574



Xu-Lin Cai^{1,2*}, Chi Xu^{2,3}, Ming-Han Deng^{2,4}, Tai-Ping Du^{2,5} and You-Peng You²

¹Luoyang Electro-Optical Equipment Research Institute of AVIC, China

²College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, China

³COMAC Shanghai Aircraft Design & Research Institute, China

⁴AECC Chengdu Engine Co., Ltd, China

⁵Shanghai Aerospace Equipment Manufacturer Co., Ltd., China

***Corresponding author:** Xu-Lin Cai, College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, People's Republic of China

Submission: 📅 February 09, 2026

Published: 📅 February 18, 2026

Volume 7 - Issue 1

How to cite this article: Xu-Lin Cai*, Chi Xu, Ming-Han Deng, Tai-Ping Du and You-Peng You. A Semantic-Provenance Framework for Interoperability of Design-Centric Industrial Software. *Adv Civil Eng Tech.* 7(1). ACET.000654.2026.
DOI: [10.31031/ACET.2026.07.000654](https://doi.org/10.31031/ACET.2026.07.000654)

Copyright@ Xu-Lin Cai, This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract

Design-centric R&D workflows in civil engineering, mechanical engineering, and aerospace generate digital artifacts that couple geometry with constraints, solver settings, experimental traces, and provenance. Reusing these artifacts across design, simulation, and lifecycle systems requires preserving both syntactic fidelity (geometry and files) and semantic fidelity (function, intent, assumptions and lineage). This mini-review proposes a lightweight, team-owned semantic-provenance framework that combines a minimal canonical export and controlled property dictionary, a standards-aligned provenance manifest, a mediation and validation hub for cross-tool semantic enrichment and canonical part identity, standardized transport of runtime data, and automated round-trip validation to detect semantic drift. The approach aims to improve model reuse, traceability, and cross-disciplinary collaboration while remaining practical to deploy within existing engineering toolchains.

Motivation and Scope

Design-centric artifacts differ from production drawings: they capture experimental setup, solver configurations, parametric families and variant relationships that are material to reproducibility. When such artifacts are exchanged between CAD/CAE, PLM/PDM and downstream execution or monitoring systems, semantic content is often lost or transformed in ways that obscure engineering intent. The consequences are measurable: repeated manual reconciliation, mistranslated inspection rules, failed experiments and wider gaps between simulated and as-built performance. These problems are not primarily technical impossibilities but rather consequences of semantic drift, brittle point-to-point adapters and a lack of provenance [1,2]. The work surveyed here aims to combine established standards and recent research on semantic enrichment into a lightweight pattern that R&D teams can adopt without becoming integration vendors.

Elements of a Semantic-provenance Framework

The first element of the proposed framework is a minimal canonical export together with a short, project-level property dictionary. Requiring a single neutral export format (for mechanical artifacts, STEP AP242 or a well-specified XML/JSON profile) together with 10–20 mandatory property fields (role, material provenance, variant tag, acceptance criteria) reduces translator permutations and preserves the minimum engineering intent needed for downstream automation. Prior studies of AP242 demonstrate that managed model-based exchange significantly improves geometric and PMI fidelity when properly constrained by an interchange profile [1,3].

Second, each artifact should carry a machine-readable provenance manifest. W3C PROV provides an interoperable model for entities, activities and agents and is suitable for representing who performed an export, which solver and which dataset were used, and which test rig configuration applies [4]. Wrapping PROV-style manifests in an Asset Administration Shell (AAS) envelope yields predictable lifecycle endpoints and simplifies discovery for downstream consumers (PLM, MES, twin registries). AAS guidance and reading collections provide pragmatic patterns and toolkits for implementing such envelopes [5,6].

Third, a mediation and validation hub should be placed between authoring tools and consumers. The hub ingests the canonical export, runs schema and semantic validation against the property dictionary, issues canonical identifiers (UUIDs linked to human labels and barcodes), performs simple semantic enrichment (rule-based mapping or lightweight ontology alignment) and produces human-readable difference reports to assist triage. Graph-based semantic enrichment approaches for example those that construct knowledge graphs from geometric exports and infer relationships have been shown to reconstruct lost context and enable reliable reconstruction in target tools [2,7]. This hub pattern reduces the combinatorial cost of point-to-point adapters and keeps the integration logic under R&D control.

Fourth, runtime and experiment data (sensor traces, rig logs) must be transported and discoverable using standard patterns. Use of OPC UA for structured machine/edge data in combination with REST/MQTT gateways and a twin registry provides practical, vendor-neutral paths for runtime interrogation and for linking sensor records to canonical artifact IDs. Liaison agreements and joint initiatives show active industry convergence on these patterns [8].

Finally, automated round-trip validation (export → ingest → re-export → compare) is essential to detect semantic loss early. Building SMART's validation efforts for IFC and parallel validation tooling for AP242 illustrate that automated conformity checks substantially reduce exchange regressions; applying the same principle to design artifacts yields an early signal for mapping failures [3,9].

Pilot Roadmap and Metrics

A focused pilot accelerates learning and demonstrates value. Recommended scope is a single repeatable artifact family (for example, a parametric bracket, a test coupon family or a modular subassembly) with clearly measurable KPIs such as reduction in handover corrections, time from design freeze to validated build artifact, and mismatch rate between simulation and experiment. The pilot phases are: (1) define the property dictionary and canonical ID scheme; (2) produce neutral exports and ingest them into the mediation hub; (3) attach PROV-backed AAS manifests; (4) run semantic-enrichment passes; (5) execute round-trip validation and quantify KPI deltas. A successful pilot commonly yields a 30–

50% reduction in manual handover work for the scoped artifact class [5,6].

Discussion and Deployment Cautions

The principal operational risks are over-engineering the schema, misallocating ownership, and vendor resistance. To mitigate these, the initial dictionary should be minimal; ownership of the hub must be assigned to a single R&D team that controls validation gates; and the hub should be marketed internally as a non-intrusive mediator rather than a tool replacement. Open-source AAS implementations and community validation tooling lower implementation cost and reduce the need for bespoke heavy engineering. As semantic enrichment tools mature, they should be applied incrementally and only where evidence shows recurring interoperability failures.

Conclusion

Design-centric R&D teams can preserve engineering intent and enable reproducible, automatable handovers by combining a minimal canonical export and property dictionary, a PROV-based provenance manifest wrapped in an AAS envelope, a mediation/validation hub with semantic enrichment, and automated round-trip validation. These elements form a pragmatic, team-owned pattern that balances fidelity with implementation cost and that is compatible with existing standards and emerging industry practice. A narrow pilot is the recommended path to build confidence and to quantify ROI before scaling.

References

1. Pratt MJ (2001) Introduction to ISO 10303—the STEP standard for product data exchange. *Journal of Computing and Information Science in Engineering* 1(1): 102–103.
2. Wang Z, Ying H, Sacks R, Borrmann A (2023) CBIM: A graph-based approach to enhance interoperability using semantic enrichment. *arXiv preprint arXiv:2304.11672*.
3. Feeney AB, Frechette SP, Srinivasan V (2014) A portrait of an ISO STEP tolerancing standard as an enabler of smart manufacturing systems. *Journal of Computing and Information Science in Engineering* 14(2): 021001.
4. Moreau L, Groth P, Cheney J, Lebo T, Miles S (2013) The PROV data model and abstract syntax notation. W3C Recommendation.
5. Industrial Digital Twin Association (IDTA) (2024) Asset Administration Shell Reading Guide. IDTA, Germany.
6. Platform Industrie 4.0 (2020) Details of the Asset Administration Shell – Part 1: The exchange of information between partners in the value chain of Industrie 4.0. Federal Ministry for Economic Affairs and Energy, Germany.
7. Pauwels P, Zhang S, Lee YC (2017) Semantic web technologies in AEC industry: A literature overview. *Automation in Construction* 73: 145–165.
8. OPC Foundation (2023) Digital Twin Consortium and OPC Foundation announce liaison agreement. OPC Foundation Press Release.
9. Building SMART International (2021) IFC validation service and model checking guidelines. Building SMART Technical Documentation.