Evidence Graphs for Industrial AI

Data-plane architectures for global supply chains, Industry 4.0, critical infrastructure, and B2G data exchange from a German perspective

Perspective article draft for academic and ecosystem review

Dr. Carsten Stöcker
Spherity GmbH
carsten.stoecker@spherity.com

Prepared: July 2026

Abstract

Industrial AI is shifting from isolated prediction and automation toward evidence-mediated reasoning in compliance support, risk analysis, supply-chain control, engineering assistance, market-surveillance support, and autonomous operational agents.

This article examines the data-plane requirements behind that transition. It compares plain JSON, schema-bound JSON/XML, AAS JSON/XML, AAS RDF, RDF/OWL knowledge graphs, JSON-LD, YAML-LD, CBOR-LD, SHACL validation, and W3C verifiable data mechanisms.

The central claim defines the most suitable default architecture for Industrial AI as a layered Verifiable Linked Knowledge Graph rather than a single file format: open-world linked data with persistent identifiers, typed relations, provenance, issuer-bound claims, cryptographic integrity, status information, and validation rules.

European policy developments reinforce this claim. IPCEI-AI and the EU AI Act increase the need for evidence layers that connect industrial AI services, enterprise identities, delegation credentials, risk controls, technical documentation, logs, monitoring records, and verifiable operational data.

eIDAS 2.0 and the proposed European Business Wallet extend this policy context into trust infrastructure. They offer a European route for linking legal-person identity, electronic attestations, seals, timestamps, representation rights, delegation evidence, and verifiable business data to industrial AI evidence graphs, supporting compliance, functional safety, cybersecurity, resilience, and machine-scale verification of enterprise and agent authority.

Two operational controls follow from this thesis: Know Your Data and Data Risk Scoring. They use issuer-bound provenance, proof verification, status checking, validity periods, trust-domain membership, policy rules, and cryptographic agility to reduce the risk that low-quality, manipulated, or unauthorised data is used as AI evidence.

The Asset Administration Shell remains important as an Industry 4.0 interoperability architecture, source model, and asset-centric digital-twin container. It should not, however, be treated as the universal semantic substrate for AI-first global supply chains, critical infrastructure, and B2G reporting.

The empirical evidence is necessarily indirect: public benchmarks compare plain QA/RAG with KG-RAG or GraphRAG, not AAS JSON/XML with JSON-LD. Even so, they provide relevant proxy evidence for graph-grounded AI, including more supported answers, fewer contradictions, fewer detected hallucinations, and lower token use.

Keywords. Industrial AI; Industry 4.0; Asset Administration Shell; AAS; JSON-LD; Knowledge Graph; Verifiable Credentials; Data Integrity; open-world semantics; SHACL; CBOR-LD; YAML-LD; data plane; KG-RAG; GraphRAG; Know Your Data; Data Risk Scoring; global supply chains; critical infrastructure; B2G; Digital Product Passport; eIDAS 2.0; European Business Wallet; European Digital Identity Wallet; Agent Power of Attorney; AI Service Passport.

Contents

1 Introduction: Industrial AI as an evidence problem

Industrial AI is often framed as a model-capability problem. That framing is incomplete. In industrial environments, AI capability is constrained by the quality, provenance, semantics, and retrievability of the data plane. A model can reason reliably only if the relevant entities, relations, rules, and evidence can be retrieved, checked, and explained. This is especially important in global supply chains, industrial operations, critical infrastructure, and B2G data exchange, where erroneous outputs can create safety, compliance, security, financial, or public-trust risks.

The core task extends beyond exposing more data to a model; it requires structured, trusted, and auditable evidence. This shifts the question from "What is the best data format?" to "What is the right AI data plane?" In this article, a data plane denotes the formats, identifiers, semantics, proof mechanisms, validation rules, registries, governance policies, and retrieval methods that determine how machines can use data. For Industrial AI, it must support evidence-path construction: entity identification, typed relation traversal, source selection, issuer evaluation, proof verification, status checking, rule validation, and answer explanation.

This requirement also has a security dimension. Industrial AI pipelines and multi-agent systems can be attacked through the data plane: forged supplier claims, manipulated sensor streams, poisoned retrieval corpora, or replayed credentials can all become hostile evidence. The data plane must therefore support origin verification, status checking, provenance analysis, risk scoring, and cryptographic agility, in addition to semantics and retrieval.

This article argues that the most defensible default architecture is a Verifiable Linked Knowledge Graph (VLKG). The term denotes an architectural pattern rather than a single standard. A VLKG combines Linked Data, graph semantics, open-world integration, closed-world validation where required, verifiable credentials or similar signed claim artefacts, and graph-guided retrieval. JSON-LD is the default serialization because it is web-native, compatible with JSON-based systems, and maps terms to IRIs. YAML-LD and CBOR-LD serve as companion serializations for human-authored artefacts and constrained environments. RDF/OWL provide graph semantics and open-world reasoning. SHACL provides validation. W3C verifiable data mechanisms provide issuer binding and integrity protection [1–9].

Figure 1 condenses this thesis visually. It shows the intended chain from verifiable Linked Data to evidence-path retrieval: graph semantics help AI interpret meaning; issuer, provenance, and status metadata allow checks before evidence is used; and graph-guided retrieval can select a compact evidence path for generation.

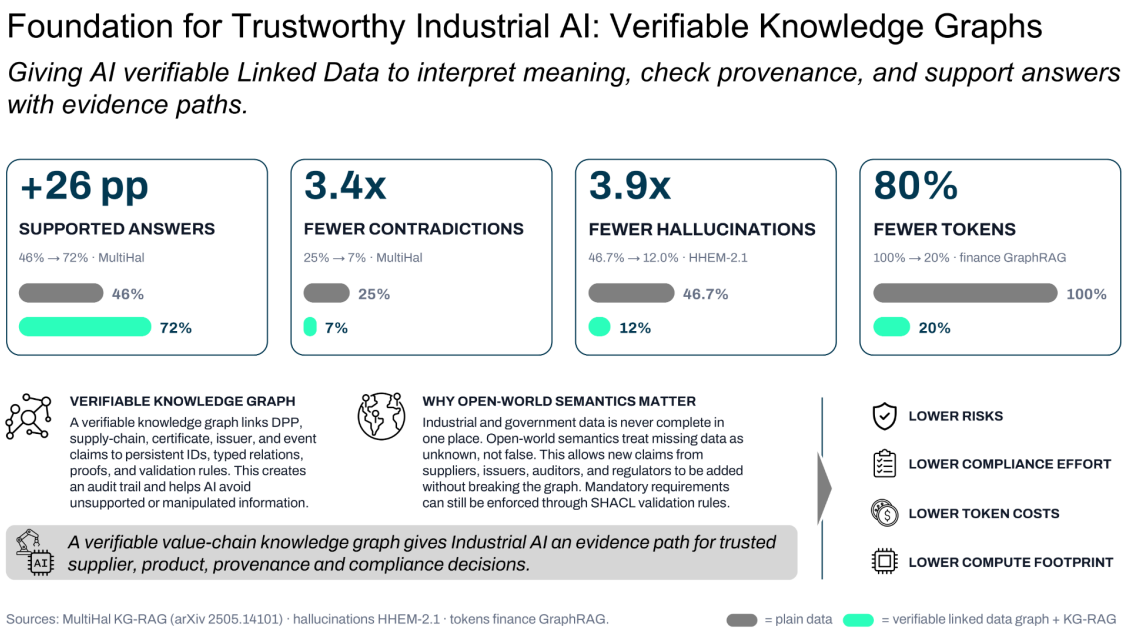

Figure 1: Verifiable Knowledge Graphs as the data foundation for trustworthy Industrial AI. The figure summarises how graph-grounded retrieval can improve AI evidence quality and efficiency compared with plain data retrieval. Verifiable Linked Data connects entities, claims, issuers, certificates, provenance, validation rules, and evidence paths. This allows Industrial AI systems to support more answers with cited evidence, reduce contradictions and hallucinations, and lower token use. The reported figures are proxy results from KG-RAG and GraphRAG studies, not direct industrial benchmark results.

The labels in Figure 1 are architectural shorthand. "Plain data" denotes plain QA/RAG or conventional retrieval baselines. "Verifiable linked data graph + KG-RAG" denotes graph-grounded retrieval over linked evidence. The figure should be read as proxy evidence for graph-grounded retrieval and not as a direct benchmark of JSON-LD, YAML-LD, CBOR-LD, or AAS as file formats.

The argument is developed from a German perspective because Germany has invested heavily in Industry 4.0 and the Asset Administration Shell. AAS is a major result of that lineage: it provides a structured digital representation of assets, submodels, APIs, registries, repositories, security concepts, and exchange formats [10–14]. It should be evaluated as a serious industrial interoperability technology. Yet LLMs, Retrieval-Augmented Generation (RAG), Knowledge-Graph RAG (KG-RAG), and agentic AI change the evaluation criteria. The primary question is no longer only whether an industrial asset can be represented, but whether AI can use distributed industrial claims as checkable evidence.

This leads to a precise position: AAS is useful as an industrial source model and adapter, while the universal AI data plane should be a verifiable knowledge graph. AAS should not be treated as the universal AI data plane.

AAS RDF can be part of a knowledge graph. AAS plus verifiable data mechanisms can be part of a VLKG. In many global supply-chain contexts, however, direct JSON-LD-based knowledge graphs are simpler, more compatible with web architecture, and better aligned with AI-oriented evidence requirements. For supply chains that involve APAC, North America, South America, Europe, public authorities, and heterogeneous supplier networks, that difference is material.

The IPCEI-AI initiative makes this architecture question more urgent. Important Projects of Common European Interest are large-scale cross-border projects that may receive State aid when

they involve significant technological or financial risk and generate common European benefits, including European ecosystems [15]. The Federal Ministry for Economic Affairs describes IPCEI Artificial Intelligence as a European project to create an industrial AI ecosystem that is accessible to companies, including SMEs [16]. Because regulated domains such as manufacturing, mobility, energy, health, logistics, and critical infrastructure participate in or are targeted by such initiatives, IPCEI-AI requires more than shared models, compute capacity, or isolated datasets. It requires a verifiable evidence graph layer that links models, training and validation data, service providers, enterprise identities, operational evidence, rights of use, conformity information, and audit trails across organisations and jurisdictions.

The EU AI Act creates a regulatory reason for the same evidence layer. Regulation (EU) 2024/1689 establishes a risk-based framework for AI systems [17]. For high-risk systems, it creates evidence obligations around lifecycle risk management, data governance, technical documentation, record-keeping and logging, transparency, human oversight, accuracy, robustness, cybersecurity, quality management, documentation retention, and post-market monitoring [17]. Reliable automation of these obligations requires connected identity, delegation, data lineage, logs, model status, and operational evidence.

Enterprise identity credentials identify the legal and operational actors in the AI value chain, including providers, deployers, importers, distributors, product manufacturers, notified bodies, auditors, data providers, and public authorities. Delegation credentials and Agent Power-of-Attorney credentials can express that a human, service, or software agent is authorised to perform a specific compliance or operational action for an enterprise. An AI Service Passport, used here as an operational profile rather than a formal EU instrument, can expose system identity, intended purpose, risk class, role allocation, model and service version, conformity status, technical-documentation pointers, logging endpoints, monitoring obligations, cybersecurity posture, and incident-reporting contacts. The VLKG binds these credentials and records to the evidence that AI consumes and produces. In industrial and physical AI, this serves both as an administrative compliance mechanism and as a functional-safety, cybersecurity, and resilience mechanism because it controls which data, models, agents, and operational claims may be trusted by an AI pipeline.

eIDAS 2.0 and the proposed European Business Wallet add the trust infrastructure that can make this pattern enforceable at European scale. Regulation (EU) 2024/1183 establishes a secure and interoperable European Digital Identity framework for natural and legal persons, while the proposed European Business Wallet aims to let companies and public-sector bodies identify, authenticate, exchange trusted data with legal effect, sign, seal and timestamp documents, and manage representation rights and mandates. The Council's 2026 negotiating position also treats interoperability, cybersecurity, powers of attorney, and legal mandates as central design issues [18–20]. For Industrial AI, the key point is that autonomous agents will retrieve evidence, submit reports, execute transactions, and trigger operational actions across company boundaries. A verifiable evidence graph can link a legal-person or wallet-owner identity to an enterprise role, a delegated Agent Power-of-Attorney credential, an AI Service Passport, a model or service instance, and the evidence used for a decision. The European Business Wallet therefore becomes a potential European trust anchor for real-time verification of machine counterparties and delegated authorities in regulated AI value chains.

2 Methodology

The article uses a qualitative, design-oriented methodology. Rather than presenting a new industrial benchmark, it combines five forms of analysis to assess which data-plane architecture is best suited for AI-oriented industrial systems.

1. Historical-technical analysis. The German lineage from Semantic Product Memory, Smart Factory work, Industrie 4.0, RAMI 4.0, Digitale Verwaltungsschale, and AAS is analysed as a specific trajectory in industrial standardisation.
2. Comparative standards analysis. The article compares the semantics, serializations, validation mechanisms, proof mechanisms, and operational profiles of AAS, JSON-LD, RDF/OWL, SHACL, verifiable credentials, and related web standards.
3. Evidence synthesis. KG-RAG and GraphRAG research is used as proxy evidence for graph-grounded AI. The article explicitly avoids treating those benchmarks as direct AAS-versus-JSON-LD experiments.
4. Requirements and sector analysis. Four target domains are evaluated: global supply chains, Industry 4.0 operations, critical infrastructure, and B2G data exchange.
5. Case-pattern analysis. The article evaluates representative data types such as B2G documents, supply-chain claims, customs data, Digital Product Passport data, IoT time series, and AI model or agent metadata, and analyses the reasoning strategies enabled by knowledge graphs.

The evaluation uses the following criteria: open-world integration, closed-world validation, graph-native reasoning, provenance, verifiability, status management, developer adoption, global interoperability, adoption friction, support for constrained environments, regulatory auditability, cryptographic agility, and suitability for AI grounding.

3 Historical context: from Semantic Product Memory to AAS

The AAS emerged within a German research and standardisation trajectory that predates the current LLM and agentic AI cycle. The early intellectual roots include work on digital product memories and semantic product memories. Wolfgang Wahlster described the DFKI Smart Factory work since 2005 and the SemProM project on semantic product memories since 2008 as assigning a digital twin to each factory module and each product produced [21]. SemProM also produced academic work on semantic product memories for the Internet of Things [21].

The public industrial-policy concept of Industrie 4.0 then emerged around 2011 and was consolidated in the 2013 recommendations for implementing the strategic initiative Industrie 4.0. That report presented a German strategy for securing the future of manufacturing through cyber-physical systems, connected production, and end-to-end engineering across the value chain [22]. RAMI 4.0 then provided a three-dimensional reference model to structure Industry 4.0 discussions and align participants around layers, life-cycle phases, and hierarchy levels [23].

The "Digitale Verwaltungsschale" and later the Asset Administration Shell became the vehicle for representing the digital part of an Industrie 4.0 component. Early AAS work was strongly tied to ZVEI and Plattform Industrie 4.0. IDTA documentation states that the first version of the AAS specification was produced from September 2017 to July 2018 by a joint working group with members from ZVEI and Plattform Industrie 4.0 [10]. IEC 63278-1 internationalised the concept by defining the AAS as the structure of a standardised digital representation of an

asset, giving uniform access to information and services and supporting information exchange between software applications [11].

This historical lineage explains the strengths and limits of AAS. AAS was designed for industrial interoperability, asset administration, submodel structures, and controlled industrial ecosystems, while an AI-first, open-world, claim-centric, verifiable knowledge graph data plane follows a different design logic. This makes AAS path-dependent rather than obsolete. Industrial AI evaluation should therefore ask where its design assumptions fit and where they add avoidable complexity.

4 Definitions and scope

4.1 Industrial AI

Industrial AI is used here to mean AI systems that operate on industrial, supply-chain, infrastructure, or regulatory data. This includes predictive maintenance, quality analytics, engineering assistance, compliance analysis, supply-chain risk reasoning, Digital Product Passport support, energy and grid operations, pharma and logistics compliance, procurement, market surveillance, B2G reporting, and AI agents that query industrial data spaces. The common feature is that consequential outputs must be traceable to data that can be audited.

4.2 Verifiable Linked Knowledge Graph

A Verifiable Linked Knowledge Graph is a distributed graph of typed claims about real-world entities, where each claim can be linked to identifiers, provenance, issuer metadata, proof material, status information, evidence, and validation rules. The term is an architectural synthesis, not a new formal standard. It combines standards and patterns from Linked Data, RDF/OWL, SHACL, verifiable credentials, data integrity proofs, identifiers, and graph-based retrieval.

A VLKG has four layers. First, linked data: entities and relations are identified through stable identifiers and typed graph edges. Second, open-world semantics: missing information is treated as unknown rather than automatically false. Third, verification: selected claims can be bound to issuers, timestamps, cryptographic proofs, and status mechanisms. Fourth, validation: mandatory fields, regulatory constraints, and business rules can be checked through shapes, schemas, policies, or verifier-specific logic. The architecture therefore separates meaning, evidence, and compliance checks.

4.3 AAS

The Asset Administration Shell is a standardised digital representation of an asset. In industrial systems, it can contain submodels, submodel elements, references, concept descriptions, semantic identifiers, packages, APIs, repositories, registries, and security rules. AAS should be analysed as a metamodel and industrial architecture [10–14].

4.4 QA, RAG, KG-RAG, and GraphRAG

QA means Question Answering. Plain QA is a system answering directly without structured graph grounding. RAG means Retrieval-Augmented Generation: the AI retrieves external text or documents and then generates an answer. KG-RAG retrieves from a knowledge graph, including entities, relations, paths, and structured claims. GraphRAG is a broader family of graph-based

retrieval and generation methods. The empirical benchmarks cited in this article compare plain QA or conventional RAG with KG-RAG or GraphRAG; they should be interpreted as graph-grounding benchmarks rather than direct JSON-LD-versus-AAS tests [24–28].

5 Candidate data architectures

Industrial organisations often mix several data formats and models. The relevant comparison is therefore a comparison of architectural roles rather than file formats alone. Plain JSON is useful for transport and API payloads. JSON Schema can validate structure. AAS JSON/XML provides a schema-bound industrial digital-twin payload. AAS RDF projects the AAS model into a graph. JSON-LD serializes Linked Data in JSON. YAML-LD aims to serialize Linked Data in YAML. CBOR-LD aims to serialize Linked Data compactly for constrained environments. RDF/OWL provide graph semantics. SHACL validates RDF graphs. Verifiable Credentials and Data Integrity proofs bind claims to issuers and protect authenticity and integrity [1–9].

Table 1: Candidate data architectures and their relevance for Industrial AI grounding.

Architecture	**Primary role**	**AI grounding value**	**Main limitation**
Plain JSON	General API and message format	Low. Requires external meaning and context.	Field names are local; semantics, provenance, and evidence paths are usually implicit.
JSON + JSON Schema	Structured payload with syntactic validation	Medium for tool calls and structured RAG.	Validation is useful, but graph meaning and issuer-bound evidence are not native. They require additional mechanisms.
AAS JSON/XML	Structured industrial digital-twin payload	Medium. Good source for structured RAG and AAS-aware tools.	Not graph-native unless projected to RDF or Linked Data.
AAS RDF	AAS represented as RDF graph	High for KG-RAG if deployed with graph retrieval and validation.	Adds semantic lifting and AAS-specific modelling overhead.
JSON-LD	Linked Data in JSON for APIs, credentials, and web systems	High when used as graph-compatible evidence data.	Requires vocabulary governance and context discipline.
YAML-LD	Human-readable Linked Data for specifications and configuration	Supportive. Useful for authoring and governance artefacts.	Draft status; primarily an authoring and governance format rather than an automated API payload format.
CBOR-LD	Compact Linked Data for constrained environments	Supportive. Useful for edge, offline, IoT, and barcode-related constraints.	Draft status; depends on JSON-LD ecosystem and implementation maturity.
Verifiable Linked KG	Claim-centric evidence graph with proofs and validation	Very high. Supports evidence-path retrieval and auditability.	Requires ontology, identifier, proof, status, and governance discipline.

This table leads to a key finding: AAS JSON/XML is not plain data. It is a structured industrial data model. Yet it is not equivalent to a knowledge graph. The graph layer starts with RDF/Linked Data. The verifiable layer starts when claims are bound to issuers, proofs, status mechanisms, and verifier policies.

6 AAS as an industrial digital-twin architecture

6.1 What AAS contributes

AAS contributes a disciplined asset-centric structure. It provides a metamodel for AASs, assets, submodels, submodel elements, references, and concept descriptions. Submodels constitute the content of the AAS and describe content-related or functional aspects of an asset [14]. The AAS API specifications provide normative OpenAPI files for HTTP/REST APIs [12]. AAS Security Part 4 defines an access-rule model, uses attribute-based access control, and allows access rules for registries, repositories, AASs, submodels, and individual submodel elements [13]. The AASX package format supports file-based exchange of an AAS with auxiliary documents and artefacts [10, 14].

These features are useful for manufacturing ecosystems. They help organisations structure asset data, build repositories, integrate operational technology, and communicate with AAS-aware tools. They also support the digital twin vision in Industry 4.0 and are relevant in Manufacturing-X and similar European dataspace projects.

6.2 Where AAS requires graph projection for AI

AAS JSON/XML offers a structured representation, while an AI-native evidence graph requires graph-native evidence paths. AAS can include semantic identifiers and concept descriptions, yet coherent graph semantics, cross-domain linking, provenance, issuer-bound claims, and proof verification require more than semanticIds alone. For KG-RAG and explainable AI, the relevant AAS layer is AAS RDF or another graph projection.

The AAS specification itself supports this distinction. Part 1 is a technology-neutral metamodel and provides mappings to XML, JSON, and RDF [10]. Prior AAS documentation and academic assessments have treated JSON/XML primarily as exchange representations, while RDF is the path to reasoning and advanced querying. Rongen et al. state the complementarity clearly: AAS models are easier to integrate with operational technologies in production, while RDF-based models offer more semantic expressiveness and advanced querying [29].

6.3 AAS and closed-world validation

AAS is often described as closed-world because it is schema-bound and conformance-oriented. This should be stated carefully. AAS JSON/XML is better described as a metamodel with mandatory and optional elements, schemas, templates, and conformance expectations than as a formal closed-world logic. This supports closed-world validation behaviour: required fields can be checked, missing mandatory elements can fail validation, and submodel templates can constrain content. That is valuable for compliance and interoperability.

However, industrial knowledge is often incomplete and distributed. A supplier may provide material data, an OEM may provide configuration data, a certifier may provide conformity claims, a recycler may provide end-of-life events, and a regulator may provide rules. Treating all absent information as false would be incorrect. AAS JSON/XML can validate what is present against a model, but it does not naturally express the open-world epistemic stance needed for distributed industrial knowledge. This requirement points toward RDF/OWL and Linked Data [4–6].

6.4 AAS add-ons: benefits and relevance for Industrial AI

Table 2: AAS add-ons and their relevance for Industrial AI.

AAS add-on	Industrial value	AI-first relevance	Assessment
Metamodel	Common structure for assets and submodels.	Useful as source structure.	Relevant, but not enough for graph reasoning.
Submodel templates	Reusable domain data contracts.	Useful where templates are mature and adopted.	Strong in controlled ecosystems; weaker in global heterogeneous networks.
semanticId / ConceptDescription	Links elements to semantic definitions.	Reduces ambiguity if used consistently.	Relevant; still needs vocabulary governance and graph linking.
AAS JSON/XML	Exchange and API payloads.	Good for structured RAG and tool calls.	Not graph-native; should be projected to RDF/Linked Data for KG-RAG. Requires projection to RDF/Linked Data for KG-RAG.
AAS RDF	Graph representation of AAS.	Good for graph retrieval and reasoning.	Most relevant AAS layer for AI grounding.
AASX	Package exchange with auxiliary files.	Useful for document/package transfer.	Package-level integrity differs significantly from claim-level verifiability.
AAS APIs	REST-based access to AAS infrastructure.	Useful for AI agents using tools.	Useful adapter layer, not a semantic substrate.
Registries/ repositories	Find and access AAS objects.	Useful in controlled ecosystems.	Needs global identifier and trust integration.
Security Part 4	ABAC, access rules, OAuth/OIDC integration.	Important for data access control.	Access control is distinct from verifiable claims.
OPC UA / AutomationML mappings	OT and engineering integration.	Useful in factory settings.	Important for brownfield Industry 4.0; less central for global claim graphs.

AAS add-ons are relevant for industrial interoperability. Their relevance, however, remains much narrower than the AI data-plane problem. They address asset administration, operational integration, and ecosystem compatibility. Alone, they provide limited support for global, claim-centric, issuer-bound evidence-path construction.

7 Verifiable Linked Knowledge Graphs as an AI-oriented data-plane architecture

7.1 Why linked graphs fit Industrial AI

Industrial AI tasks frequently require multi-hop reasoning: product to component, component to material, material to supplier, supplier to credential, credential to issuer, issuer to registry, claim to regulation, event to status, and status to decision. These are graph questions. Knowledge graphs integrate diverse data and support querying, validation, ontologies, rules, and reasoning [30]. They have also been studied as tools for explainable machine learning because they can make model outputs more interpretable and traceable [31].

The Semantic Web and Knowledge Graph literature for Industry 4.0 already points in this direction. Yahya et al. survey Semantic Web and Knowledge Graph use for Industry 4.0 across equipment information, maintenance, resources, production, and services [32]. Semantic AAS

research similarly shows that AAS becomes more useful for reasoning when represented in RDF, validated, and queried with graph technologies [29, 33–35].

Two domain examples make the graph argument concrete. They show the same data-plane pattern in two settings: an energy-system use case, where product passports and time-series data interact, and a mobility-safety use case, where edge event data and contextual evidence must be governed before AI use.

7.1.1 Battery energy storage systems, Digital Battery Passports, and flexibility reasoning

A battery energy storage system (BESS) illustrates why Industrial AI needs linked evidence rather than isolated payloads. A BESS contains racks, battery modules, control systems, operating limits, and service histories. Each battery module can be associated with a Digital Battery Passport (DBP), understood here as a battery-specific Digital Product Passport. The DBP can expose a product knowledge graph with supply-chain, material, manufacturing, conformity, safety, lifecycle, and sustainability claims [36, 37]. The same module also produces operational time-series data, for example temperature, voltage, current, state of charge, state of health, alarms, and cycle counts. Grid forecasts and balancing-group forecasts add another layer of context.

The AI task is to retrieve the compact evidence subgraph relevant to a specific decision rather than read all DBPs and raw sensor streams into a prompt. For product safety, the AI links the battery module to its DBP, certified operating limits, issuer proofs, status checks, current thermal behaviour, and applicable safety rules. For predictive maintenance, it links degradation features to manufacturing batch, chemistry, supplier claims, maintenance events, and comparable modules. For flexibility forecasting, it links state of charge, state of health, temperature limits, contractual constraints, grid forecasts, and balancing-group forecasts. The linked data is conditional: the relevance of a claim depends on time, operating state, temperature, location, forecast horizon, and market or grid context.

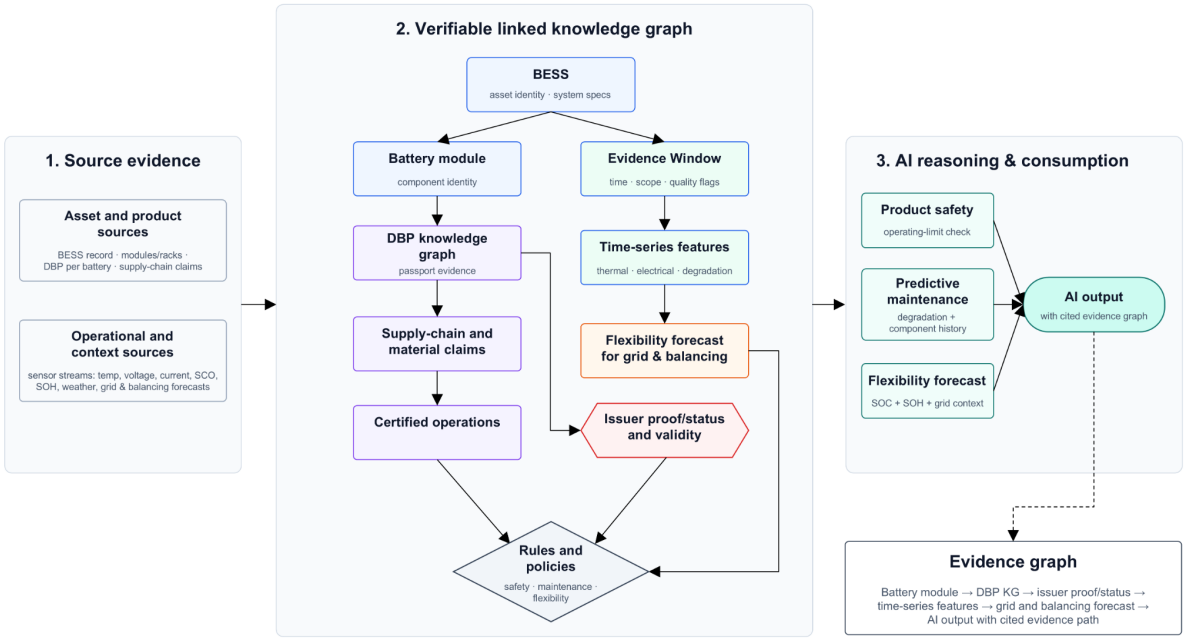

Figure 2: BESS evidence graph for Industrial AI. Asset records, Digital Battery Passport data, supply-chain claims, sensor-derived time-series features, and grid or balancing-group forecasts are linked into a verifiable knowledge graph. The graph preserves identifiers, provenance, issuer proof/status, validity, and time context so that AI systems can perform product-safety checks, predictive maintenance, and flexibility forecasting with a cited evidence path.

The design principle is separation of storage and reasoning. Raw high-frequency streams should remain in time-series systems, stream processors, or feature stores. The knowledge graph should hold stable identifiers, semantic context, derived features, data-quality flags, provenance, evidence-window links, issuer proofs, and policy constraints. KG-RAG then retrieves a smaller evidence bundle: the relevant battery, passport claims, proof status, time window, forecast assumptions, and rules. This improves explainability and can reduce prompt size because the model receives a selected evidence path rather than unfiltered documents and time-series exports.

7.1.2 Dangerous-driving event evidence chains

A second example is a dangerous-driving event evidence chain. Modern vehicles and fleets generate heterogeneous evidence: vehicle state, braking and steering signals, inertial measurements, GNSS location, camera/radar/lidar-derived features, telematics, edge classifications, environmental context, road context, backend model versions, policy constraints, and data-sharing permissions. The relevant graph is a linked event evidence chain that connects the event to the vehicle, trip, time interval, location, sensor provenance, edge-processing result, environmental data, backend analysis, and permitted uses [38–41].

The AI questions are operational and evidentiary. Was the event dangerous? Was it caused by vehicle behaviour, driver behaviour, weather, road conditions, sensor degradation, map error, traffic behaviour, or an autonomous-driving software decision? Should it trigger driver coaching, maintenance, insurance review, road-infrastructure feedback, safety-case analysis, model evaluation, or a regulatory report? A knowledge graph gives the AI a governed path through these questions. It can retrieve the event window, sensor calibration state, environmental context, software version, edge-decision confidence, backend classification, and use policy. It

can also decide that raw data should remain at the edge or in a controlled data lake while only derived features, proofs, and links are used for generation.

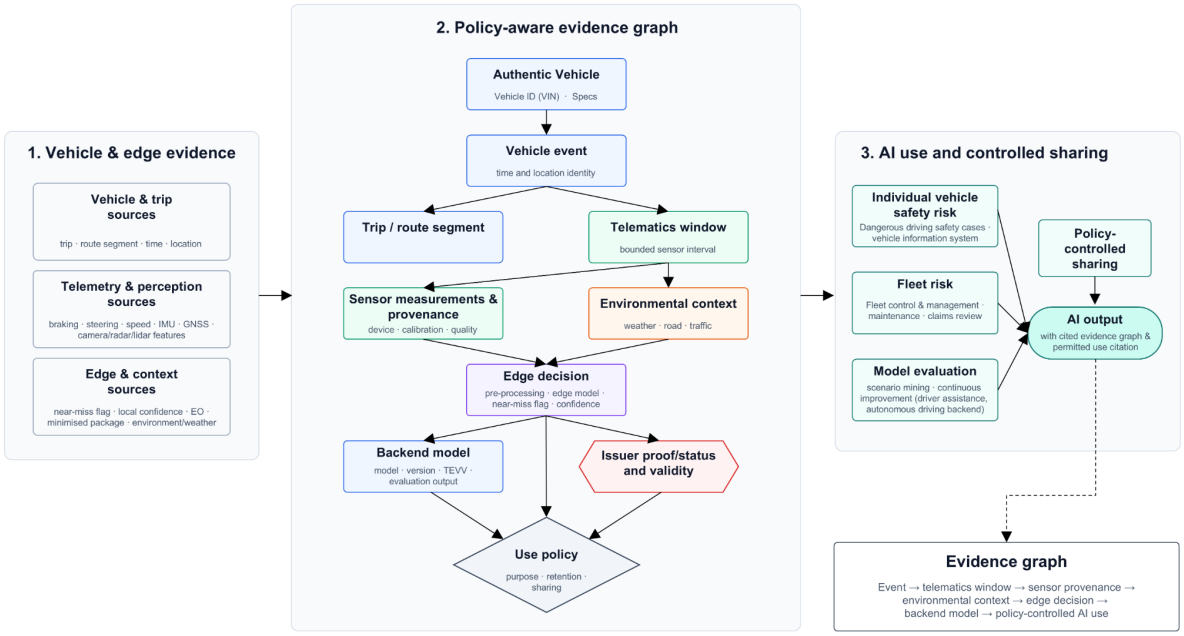

Figure 3: Dangerous-driving event evidence graph. Vehicle trip data, telematics, perception features, edge assessments, environmental context, backend model information, and use policies are fused into an event-centric evidence graph. The graph links the event to the vehicle, time window, location, sensor provenance, edge decision, backend processing, and policy constraints, enabling AI use for event reconstruction, safety-case updates, fleet-risk analysis, model evaluation, and policy-controlled sharing.

This example also shows why the data plane must distinguish between data availability and data use. The same event may support safety engineering, fleet-risk management, claims review, model evaluation, or regulatory oversight. These uses are not equivalent. A verifiable linked graph can attach purpose, consent, retention, access, provenance, and policy constraints to the evidence path. Industrial AI can then consume the graph under policy rather than treat all retrieved telemetry as equally available context.

Both examples lead to the same architectural conclusion. The knowledge graph serves as the semantic and verifiable control layer rather than a warehouse for every raw value that tells the AI which sources exist, how they are related, which claims are valid, which evidence windows are relevant, which policies apply, and which answer can be justified. This is the core difference between document retrieval and evidence-path reasoning.

7.2 Why JSON-LD should be the default exchange serialization

JSON-LD 1.1 is a W3C Recommendation for serializing Linked Data in JSON. It is designed to integrate with systems that already use JSON and to provide a smooth upgrade path from JSON to JSON-LD [1]. This makes it a strong default for global industrial APIs because most enterprise and public-sector web systems already understand JSON. JSON-LD enables a staged path without requiring every participant to operate a triple store from the outset: ordinary JSON compatibility, stable identifiers, contexts, graph expansion, RDF processing, SHACL validation, and verifiable credential processing.

The broader web adoption of JSON-LD also matters. Google recommends JSON-LD for structured data where possible because it is easier to implement and maintain at scale [42]. The

2024 Web Almanac reports JSON-LD on 41 percent of pages in its structured-data analysis, indicating broad web deployment outside specialised industrial circles [43]. This does not make JSON-LD a complete industrial solution. It does make JSON-LD a lower-friction global entry point than AAS for many developers.

7.3 YAML-LD and CBOR-LD as companion formats

The AI-first data plane should combine JSON-LD with companion serializations. YAML-LD is useful where humans author specifications, examples, configuration files, rules, and schemas. W3C YAML-LD defines YAML conventions for serializing Linked Data based on JSON-LD syntax, semantics, and APIs [2]. CBOR-LD is useful where storage, bandwidth, offline exchange, barcode constraints, edge devices, or IoT environments matter. W3C CBOR-LD defines a CBOR-based serialization for Linked Data for storage- and bandwidth-constrained programming environments [3]. Both remain W3C Working Drafts and should be treated accordingly.

Together, YAML-LD, JSON-LD, and CBOR-LD form a coherent family: YAML-LD for human-facing governance artefacts, JSON-LD for web and API exchange, and CBOR-LD for compact industrial and constrained exchange. The same graph can be expressed in different encodings. That pattern is more suitable for global supply chains than treating a single industrial container as the universal representation.

7.4 Verifiability as an AI data-plane property

Verifiability is essential for high-stakes Industrial AI. The AI output should be able to cite evidence that a verifier can check. W3C Verifiable Credentials Data Model 2.0 describes a credential as a way to express claims made by an issuer and defines an issuer-holder-verifier ecosystem [7]. W3C Data Integrity provides mechanisms for ensuring authenticity and integrity of verifiable credentials and similar digital documents using cryptographic proofs [8]. VC-JOSE-COSE provides a standards path for securing verifiable credentials with JOSE and COSE [9].

The important caveat is that cryptographic verifiability differs from real-world truth. The W3C VCDM explicitly separates credential verification from claim truth; verifiers must apply business rules and evaluate the issuer, proof, subject, and claims before relying on them [7]. This caveat is central for Industrial AI. A signed false claim remains false. Verifiability provides issuer binding, integrity, currency, status checks, and accountability. It must be combined with trusted issuers, validation, governance, and audit.

For high-stakes Industrial AI, verifiability is also a data-plane security control. Data Integrity proofs and VC securing mechanisms allow an AI pipeline, verifier, or agent controller to check whether a data object was issued by an expected actor, whether the protected content changed, which verification method was used, and whether the credential remains current under its status mechanism [7–9, 44].

The need is practical. Industrial AI data pipelines are adversarial surfaces. Threat actors can feed AI processing pipelines or multi-agent systems with forged certificates, manipulated product declarations, poisoned retrieval documents, false environmental claims, replayed credentials, or corrupted sensor context. If such inputs are treated as ordinary text or unverified JSON, the model can reason over hostile evidence while appearing technically fluent.

A verifiable data plane changes this failure mode. Before a claim becomes evidence, the system can verify issuer, proof, proof purpose, credential status, validity period, and trust-domain membership. It can check whether the issuer belongs to a recognised ecosystem role, such as supplier, OEM, conformity assessment body, competent authority, registry operator, grid

operator, or accredited auditor. It can also reject, quarantine, downgrade, or escalate data that lacks a valid proof path.

The compliance value is broader than authenticity. With reliable key management, issuer governance, audit logging, and legal policy, verifiable credentials and integrity proofs can support non-repudiation and attributability controls. They make it possible to ask who made a claim, under which authority, at which time, with which key, under which status, and under which policy. These checks are the foundation for Know Your Data controls in Industrial AI.

In the EU AI Act context, verifiability becomes an operational control surface. Risk-management evidence, training and validation data evidence, technical documentation, logging records, human-oversight instructions, cybersecurity evidence, and post-market monitoring records should be linked to enterprise identities, delegated authorities, and system status. A verifier or AI governance service can then verify document existence, issuer identity, authority, AI-service scope, and validity period.

The eIDAS 2.0 and European Business Wallet layer provides an authoritative trust root for those controls. Legal-person identity, representation rights, electronic attestations of attributes, electronic seals, timestamps, and delegation credentials should be represented as graph nodes and edges that are checked before a model or agent uses data as evidence. Selected claims in a VLKG can be issued or anchored through Business Wallets. A software agent should present a verifiable Agent Power-of-Attorney credential stating who delegated which authority, for which purpose, within which limits, and for which validity period; an API token alone provides insufficient authority or no verifiable authority claim at all. An AI Service Passport can then bind the service, provider, deployer, model and service version, risk class, cybersecurity posture, monitoring obligations, and authorised agent roles to the same evidence graph. This supports machine-scale compliance controls, adversarial data filtering, non-repudiation, attributability, and resilience engineering in industrial AI infrastructure [18–20].

7.5 Know Your Data and Data Risk Scoring

Know Your Data is the data-plane analogue of Know Your Customer and Know Your Business. The term is used here as a control pattern. It means that an AI system should consume data only after relevant context has been checked. Before data influences a high-stakes decision, the system should know what the data is, where it came from, who issued it, how it was transformed, whether it is current, whether it conforms to the relevant schema or shape, whether it conflicts with other evidence, and whether it is authorised for the intended use.

Data Risk Scoring operationalises this principle. It assigns a policy-dependent and explainable risk signal to claims, datasets, graph paths, and retrieved evidence bundles. The score is should not be understood as a truth score. It estimates whether data is sufficiently reliable, current, authorised, and cryptographically acceptable for a given AI task. A low-risk claim can still be false, and a high-risk claim can still be true.

In a knowledge graph, risk scoring works along nodes, edges, and paths. A product claim can be linked to a supplier credential, the supplier credential to an issuer, the issuer to a trust registry, the credential to a status list, the claim to a SHACL validation result, the product to a bill of materials, and the bill of materials to time-stamped logistics and production events. Each edge adds or subtracts evidence. The system can then compute a path-level risk score and attach it to the AI answer.

Typical graph features include issuer trust, role authorisation, proof verification, revocation or suspension status, validity period, data freshness, transformation lineage, schema conformance, sensor calibration state, missing mandatory attributes, conflicts with peer claims, anomaly

indicators, privacy restrictions, export-control constraints, and human approval status. These features can be deterministic, statistical, policy-based, or learned. In critical use cases, deterministic controls should dominate and machine-learned scores should remain explainable and auditable.

This matters because Industrial AI fuses many data types. Sensor time series need context such as sensor identity, calibration, unit, sampling interval, asset relation, and quality flag. Supplier data needs legal identity, site information, authorisation, sanctions or restriction checks, and credential status. Product and DPP data need component links, material claims, conformity evidence, and lifecycle events. Environmental data needs method, geography, time interval, measurement uncertainty, and source authority. Compliance data needs regulation version, competent authority, filing status, and validity period. A single AI answer may depend on all of these data types.

The reasoning strategy is evidence selection under risk. The AI should retrieve candidate evidence paths, discard paths that fail mandatory verification, rank remaining paths by data risk, use only policy-allowed evidence in generation, and cite the resulting path. For high-risk or conflicting evidence, the system should abstain, request additional evidence, or escalate to a human reviewer. This turns graph-grounded AI from retrieval into governed evidence processing.

The security literature supports this direction. The NIST AI RMF frames AI risk management around trustworthy system behaviour and life-cycle controls [45]. NIST adversarial machine-learning work classifies poisoning, evasion, privacy, and abuse attacks across the AI life cycle [46]. Draft NIST cybersecurity guidance for AI emphasises verification of training and input-data integrity to detect and prevent poisoning or tampering [47]. OWASP also identifies training-data poisoning and AI supply-chain vulnerabilities as major LLM application risks [48]. For Industrial AI, Know Your Data and Data Risk Scoring translate these risks into data-plane controls.

7.6 Verifiable credentials and the post-quantum transition

The Verifiable Credentials Data Model is also important for the post-quantum cryptography transition. Industrial AI evidence can have long life cycles: product passports, conformity claims, inspection records, maintenance logs, customs records, infrastructure credentials, and audit trails may need to remain verifiable for years or decades. If these objects rely on quantum-vulnerable public-key cryptography without transition planning, their evidentiary value can degrade before the industrial asset or regulatory obligation expires.

NIST finalised the first three post-quantum cryptography standards in 2024: FIPS 203 for ML-KEM key encapsulation, FIPS 204 for ML-DSA digital signatures, and FIPS 205 for SLH-DSA digital signatures [49–51]. NIST also states that organisations should begin applying these standards now and should identify where vulnerable algorithms are used and plan to replace or update them [52]. CISA, NSA, and NIST have similarly recommended quantum-readiness roadmaps, especially for critical infrastructure [53].

The VCDM, DIDs, and Data Integrity are useful here because they separate the meaning of a claim from the cryptographic method used to secure it. DID documents can express multiple verification methods and verification relationships, including authentication, assertion, and key-agreement use cases [54]. Data Integrity supports proof sets and proof chains, allowing more than one proof to protect the same secured document or to model ordered approval chains [8]. This gives a standards path for staged migration: classical proof, post-quantum proof, policy-specific proof, and migration-period proof can coexist where the ecosystem requires it.

Semantics are essential for this transition. A verifier and an AI governance layer must understand which proof was used, which verification method it refers to, which controller is responsible, which

proof purpose applies, when the proof was created, which cryptographic suite or JOSE/COSE algorithm was used, whether the key is active or deprecated, and which policy tier accepts the proof. Without semantic metadata, PQC migration is an opaque algorithm swap. With semantic metadata, it becomes a graph-managed transition.

For Industrial AI, this means that cryptographic agility should be represented in the knowledge graph itself. Claims, credentials, keys, algorithms, trust registries, revocation mechanisms, and verifier policies should be linked. A Data Risk Score can then account for cryptographic risk: for example, whether a credential still uses a deprecated algorithm, whether a PQC or hybrid proof is available, whether a signing key has been rotated, and whether the evidence satisfies the verifier policy for the risk class of the AI decision.

The practical implication is clear: the PQC transition should start before Industrial AI evidence graphs become locked into legacy cryptography. Germany and Europe should treat verifiable credentials, DID-based key discovery, multi-proof credentials, crypto-agile status mechanisms, and cryptographic bills of material as part of the Industrial AI data-plane strategy. This also concerns long-term evidentiary integrity for AI-assisted industrial decisions.

8 Open-world semantics and closed-world validation

Open-world semantics belongs in the main argument for improved reasoning in Industrial AI. Graph-guided retrieval, rather than open-world semantics by itself, explains the empirical token-use reductions. Open-world semantics improves the data foundation for reasoning, while graph-guided retrieval reduces context size and token use. Together, they support better Industrial AI.

8.1 Why open-world semantics matters

Industrial and government data is incomplete by design. No single party holds the full state of a global supply chain or critical infrastructure ecosystem. A supplier knows some facts, an OEM knows others, a certifier knows others, a regulator knows others, and a recycler or maintainer may add new facts later. In such settings, missing data should usually mean unknown.

OWL follows this open-world assumption. The W3C OWL 2 Primer contrasts database-style closed-world behaviour, where absent facts are often considered false, with OWL, where absent facts may simply be missing but possibly true [5]. This epistemic stance is well suited to global supply chains. New evidence can be added without invalidating the existing graph. Multiple parties can issue claims at different times. AI systems can reason over what is known, identify what is missing, and avoid treating absence as negation.

8.2 Why closed-world validation is still needed

Open-world integration still requires mandatory checks. Compliance often requires closed-world behaviour: a battery passport must contain required fields, a credential must contain required attributes, a safety report must include required evidence, and a public authority may reject an incomplete filing. SHACL provides the bridge by validating RDF graphs against shape conditions [6]. JSON Schema and policy engines can also validate specific payloads or workflows. The design pattern therefore combines open-world integration with closed-world validation.

This leads to the following core reasoning principle:

Industrial AI should use open-world semantics to integrate incomplete multi-party knowledge and closed-world validation to check mandatory claims, regulatory rules, and business constraints.

8.3 Token reduction is a retrieval effect

The reported token reductions in GraphRAG studies are best explained by graph-guided retrieval. A graph can select a smaller evidence subgraph: the relevant entities, relations, source claims, and paths. The LLM then receives less irrelevant text. This reduces token usage and may improve answer quality. Open-world semantics helps build an extensible graph, while graph retrieval and compact evidence selection explain token reductions. The empirical token result should be attributed to graph retrieval, graph indexing, and compact evidence selection [27, 28].

9 Evidence from KG-RAG and GraphRAG benchmarks

The available benchmark evidence supports graph-grounded AI. This distinction is important for academic precision and for avoiding overstatement. The public metrics are valuable because they show that knowledge graph integration can improve factual grounding and efficiency compared with plain QA or conventional RAG in selected tasks.

The MultiHal benchmark is a multilingual, multi-hop dataset for knowledge-graph grounded evaluation of LLM hallucinations. The authors state that knowledge graphs are useful for hallucination mitigation because they represent facts about entities and relations with minimal linguistic overhead [26]. In the aggregated NLI results over all MultiHal languages, KG-RAG reached 71.44 percent entailment and 7.31 percent contradiction, while plain QA reached 45.81 percent entailment and 24.87 percent contradiction. This corresponds to roughly +25.63 percentage points more supported answers and 3.4 times fewer contradicted answers. The same paper reports HHEM-2.1 hallucination detection on the English subset with 13, 12, and 11 percent hallucinated outputs for KG-RAG across three models, compared with 42, 45, and 53 percent for plain QA. That gives an average of about 3.9 times fewer detected hallucinations [26].

The finance GraphRAG paper by Barry et al. reports that graph-based methods in RAG achieved a 6 percent reduction in hallucinations and an 80 percent decrease in token usage compared with conventional RAG on FinanceBench [27]. The paper also reports a much larger token reduction in a specific regulatory contradiction-detection setup, but this larger number should not be understood as a general claim because it is task-specific. The safer general slide-level number is the 80 percent token decrease reported for the finance benchmark.

Table 3: Proxy evidence from KG-RAG and GraphRAG benchmarks.

Metric	**Plain QA / conventional RAG**	**KG-RAG / GraphRAG**	**Interpretation**
Supported answers	45.81% entailment	71.44% entailment	+25.63 percentage points in MultiHal NLI results [26].
Contradicted answers	24.87% contradiction	7.31% contradiction	About 3.4x fewer contradictions in MultiHal [26].
Detected hallucinations	46.7% average hallucinated outputs	12.0% average hallucinated outputs	About 3.9x fewer detected hallucinations on the English HHEM-2.1 subset [26].
Token usage	100% baseline	20% remaining	80% lower token usage in finance GraphRAG compared with conventional RAG [27].

These results are consistent with broader surveys. Agrawal et al. review knowledge-graph-based techniques for mitigating LLM hallucinations [25]. Peng et al. describe GraphRAG as using structural information across entities for more precise and comprehensive retrieval and more context-aware responses [28]. Lewis et al. showed earlier that retrieval-augmented generation can generate more factual language than a parametric-only sequence-to-sequence baseline in knowledge-intensive NLP tasks [24].

The environmental argument should be scoped in the same way. Fewer tokens can reduce inference work, cost, latency, and compute-related environmental impact. The actual footprint, however, depends on model choice, hardware, batching, data-centre overhead, and energy mix. Inference-footprint research treats model and infrastructure variables separately and warns against overly simple emissions claims [55, 56].

10 Global adoption and developer capacity

10.1 AAS developer capacity

AAS has an active specialist ecosystem with specifications, APIs, submodel templates, repositories, SDKs, AASX tooling, and German and European Industry 4.0 projects. This is valuable, but it also means that AAS adoption requires specialised knowledge: the metamodel, submodels, semanticIds, concept descriptions, AASX, registries, repositories, API profiles, and ecosystem-specific modelling conventions. For global supply-chain AI, that learning curve is material.

Outside Germany and Europe, AAS has much lower presumed developer mindshare than JSON, web APIs, schema.org, JSON-LD, or general knowledge graph technologies. AAS has been internationalised through IEC 63278 and IDTA, but its origin and early governance are German. This is an adoption-risk observation rather than a cultural criticism. Global supply chains require APAC, North American, South American, European, and public-sector actors to participate. The lower the specialised architectural burden, the faster the data plane can scale.

10.2 JSON-LD and web developer capacity

JSON-LD has a broader entry path. It starts with JSON, the dominant format for web APIs. JSON-LD adds contexts, IRIs, and graph semantics without requiring every participant to understand RDF internals at the beginning. The web deployment base is significant: Google recommends JSON-LD for structured data where possible because it is easier to implement and maintain at scale [42], the Web Almanac documents broad use of JSON-LD in web structured data [43], and Schema.org reports broad structured-data deployment across web domains and supports JSON-LD among other encodings [57].

The same web-native pattern is visible in sensor, IoT, and vehicle-data standardisation. The W3C/OGC Semantic Sensor Network ontology and SOSA model provide linked-data semantics for sensors, observations, samples, actuators, procedures, features of interest, and observed properties [38]. W3C Web of Things Thing Descriptions are encoded in JSON by default while allowing JSON-LD processing, which supports machine-understandable descriptions of physical or virtual Things and their interfaces [39]. Automotive work points in the same direction: VSSo maps the Vehicle Signal Specification catalogue into an OWL-compatible vehicle-signal ontology [58]. Because the W3C VSSo document is a discontinued draft, it should be treated as evidence of a semantic design lineage for vehicle signals rather than as a maintained W3C Recommendation-track standard [58].

JSON-LD remains demanding at enterprise scale. Context governance, vocabulary design, identifier policy, validation, proof strategy, and graph operations remain serious engineering tasks. The point is narrower: the adoption path from JSON to JSON-LD to a knowledge graph is more global and web-native than the path from generic enterprise data to AAS-first industrial modelling. For global Industrial AI, this favours JSON-LD-based VLKGs as the default.

11 Sector implications

11.1 Global supply chains

Global supply chains are multi-party and multi-jurisdictional. They include suppliers, OEMs, logistics providers, auditors, certifiers, marketplaces, recyclers, regulators, and consumers. AAS can represent asset-related information, but the AI problem is often claim-centric: who asserted which fact about which product, material, batch, supplier, certificate, shipment, or due-diligence event? A VLKG addresses this more directly through stable identifiers, graph links, provenance, issuer metadata, proofs, status, and validation rules.

GS1 Digital Link is important because it expresses GS1 identifiers such as GTIN and GLN, and qualifiers such as batch/lot and expiry date, as web identifiers [59]. This fits a web-native data plane. A product, batch, location, certificate, credential, and DPP can become linked entities in a graph. AAS can be connected as one source model where useful, while the evidence graph should remain the global abstraction.

11.2 Industry 4.0 and manufacturing operations

Within factories and manufacturing networks, AAS has strong arguments. It fits the Industry 4.0 lineage, supports asset-centric modelling, and connects to operational technology concepts. For machine builders, production lines, and brownfield integration, AAS may be the right source representation. However, AI systems that need cross-asset reasoning, regulatory evidence, supplier claims, and multi-hop graph retrieval still benefit from projecting AAS into RDF/Linked Data and adding verifiable claims.

The recommended pattern for manufacturing is therefore AAS plus knowledge graph, rather than AAS versus knowledge graph, with the graph as the AI grounding layer. AAS provides structured asset data and operational integration. The VLKG provides open-world integration, graph retrieval, evidence paths, and issuer-bound claims.

11.3 Critical infrastructure

Critical infrastructure raises the trust and audit requirements. NIS2 establishes a unified legal framework for cybersecurity in 18 critical sectors across the EU [60]. AI systems in critical infrastructure must retrieve information while also respecting access control, provenance, reliability, and auditability. AAS submodel standards are useful where AAS infrastructure exists. However, verifiable knowledge graphs provide a more general evidence model across assets, operators, vendors, inspections, incident reports, credentials, policies, and regulatory statements.

For critical infrastructure, accountability depends on auditable AI-assisted decisions rather than the visual representation of a digital twin. A VLKG can represent the path from a risk claim to a sensor, asset, operator, certificate, policy, maintenance event, issuer, and proof. That is closer to the accountability needs of AI in critical infrastructure.

11.4 B2G and public-sector reporting

B2G data exchange is moving toward machine-readable, reusable, and auditable reporting. The EU Data Act is applicable from 12 September 2025 and strengthens the legal context for access to and use of data [61]. Digital Product Passports under the Ecodesign for Sustainable Products Regulation are designed to store and share product sustainability and environmental data for consumers, businesses, and public authorities [36]. These trends favour structured, linked, and verifiable data.

B2G reporting also favours open-world integration because a public authority rarely receives all evidence from one system. It receives claims from economic operators, registers, testing bodies, auditors, and other authorities. At the same time, public law requires closed-world validation for mandatory submissions. Again, the pattern is open-world linked data for integration, SHACL or comparable validation for mandatory checks, and verifiable credentials or data integrity proofs for issuer binding and integrity.

12 Use-case patterns: data types and knowledge-graph reasoning strategies

The argument for a Verifiable Linked Knowledge Graph becomes stronger when tested against concrete industrial data types. These data types differ in granularity, lifecycle, legal status, and velocity. The graph should serve as the semantic and verifiable control plane that links operational records, documents, events, certificates, models, policies, and AI skills into evidence paths.

The main reasoning pattern is the same across the cases: identify the entity, classify the claim, connect it to a source, bind it to an issuer or authority, evaluate status and validity, apply mandatory validation where needed, and return an answer with a compact evidence path. This pattern is difficult to achieve with plain document retrieval and only partially addressed by schema-bound payloads.

12.1 B2G documents and administrative evidence

B2G data includes company certificates, public-procurement declarations, tax confirmations, regulatory filings, permits, inspection reports, and evidence exchanged between public authorities. European public procurement already illustrates the direction of travel: the eProcurement Ontology provides a formal semantic foundation for linked open procurement data, while the Once-Only Technical System is designed for cross-border evidence exchange between competent authorities [62, 63].

In a knowledge graph, a procurement or reporting document should be decomposed into issuer-bound claims. The AI system can then reason over eligibility, exclusion criteria, register status, document currency, and provenance. The preferred AI strategy is claim extraction, entity resolution, rule retrieval, SHACL or policy validation, and evidence-path generation. The output should cite the source claim, the public register or authority, and the rule that was applied.

12.2 Supply-chain data and provenance chains

Supply-chain data is multi-party, incomplete, and often contested. It includes supplier declarations, purchase orders, bills of material, batch identifiers, logistics events, certificates, audit

reports, and chain-of-custody statements. A Verifiable Linked Knowledge Graph can link these artefacts through persistent identifiers for products, organisations, sites, shipments, batches, and certificates. PROV-O is relevant here because it provides a W3C ontology for representing provenance across systems and contexts [40].

The reasoning strategy is path-based and temporal. The AI system should ask which supplier issued which claim, which component or batch it concerns, which certificate supports it, whether the certificate was valid at the relevant time, and how the claim propagates to the product or shipment. Open-world semantics is important because a missing supplier claim should usually be treated as unknown, not false. Closed-world validation is still needed when a regulation or customer policy requires mandatory evidence.

12.3 Global trade and customs data

Global trade and customs data includes declarations, invoices, transport documents, certificates of origin, tariff classifications, export-control statements, and risk indicators. The WCO Data Model is the central international reference for harmonised customs data and is now available in the Version 4 series, with Version 4.2.0 released in 2025 [64].

For Industrial AI, customs data should be linked to products, shipments, economic operators, jurisdictions, certificates, and regulatory requirements. The AI strategy is hybrid: deterministic rule checking for mandatory fields and legal classifications, graph traversal for supply-chain and document lineage, and anomaly detection or risk scoring over graph features. AAS adds little as a primary model here because the dominant entities extend much beyond assets to legal claims, documents, authorities, shipments, and jurisdictions.

12.4 Digital Product Passport data

Digital Product Passport data is naturally graph-shaped. It spans products, components, materials, economic operators, conformity claims, environmental indicators, repair events, recycling events, and regulatory requirements. CIRPASS and CIRPASS-2 focus on cross-sectoral DPP data models and DPP pilots across sectors such as textiles, electronics, tyres, and construction [37].

The key reasoning strategies are component-to-product aggregation, requirement coverage analysis, conformity-path explanation, lifecycle-event integration, and conflict detection. An AI system should be able to answer why a product claim is supported, which component DPPs it depends on, which issuer made each claim, which evidence is missing, and whether a mandatory field fails validation. This is an open-world integration problem with closed-world compliance checks. It is also a strong use case for verifiable credentials and Data Integrity proofs.

12.5 IoT data and sensor time series

IoT and sensor time-series data can remain outside the knowledge graph as raw high-frequency samples. The graph should at least describe the semantic context of the stream: sensor, actuator, observed property, feature of interest, unit, calibration, sampling procedure, asset, location, time interval, quality flag, and provenance. The W3C/OGC SSN and SOSA ontology describes sensors, observations, samples, actuators, and the observed properties; the W3C Web of Things Thing Description provides a JSON-based and JSON-LD-compatible way to describe the metadata and interfaces of Things [38, 39]. OWL-Time is relevant for temporal relations and intervals [41].

The AI strategy is therefore two-level. Numerical analytics should remain in time-series databases, stream processors, or feature stores. The knowledge graph should bind derived events, aggregates, alerts, sensor metadata, calibration records, and maintenance context to assets and claims. KG-RAG can then retrieve the relevant event window, sensor provenance, asset context, and rule set instead of sending large raw streams to the model. This strengthens the token-efficiency argument: graph-guided retrieval reduces context size, while open-world semantics explains why new sensor evidence can be integrated without treating absent data as false.

12.6 AI model, agent, and skill metadata

Industrial AI systems increasingly depend on model metadata, evaluation reports, tool descriptions, agent instructions, prompts, and operational skills. Model cards and datasheets for datasets show why AI artefacts need explicit documentation of intended use, performance, data provenance, limitations, and governance context [65, 66]. For agentic systems, human-authored Markdown skill files, policy files, or tool manuals can be treated as operational artefacts and should be linked through metadata.

A practical approach is to add YAML-LD front matter or sidecar metadata to agent skills, prompt libraries, tool definitions, and model cards. YAML-LD is useful here because the artefacts are written and reviewed by humans, while the metadata can still resolve into the same linked-data graph. The AI reasoning strategy is model and tool selection under constraints: retrieve the model or skill that is permitted for a domain, check its version, evaluate its test evidence, verify its approval status, and cite the policy or evaluation record that justifies use. This turns AI operations into a governable knowledge graph rather than a loose collection of Markdown files and configuration snippets.

12.7 Dataset catalogues and data products

Industrial AI also depends on data products: feature tables, training datasets, simulation outputs, regulatory datasets, master-data exports, and dataspace offers. DCAT Version 3 provides an RDF vocabulary for interoperability between data catalogues and for describing datasets, data services, and related resources [67]. In an AI-first data plane, catalog metadata should link datasets to provenance, usage rights, quality indicators, access policies, and downstream models.

The reasoning strategy is suitability analysis. Before an AI system uses a dataset, it should determine whether the dataset is current, authorised, complete enough, licensed for the purpose, and suitable for the decision context. This is particularly important in critical infrastructure and B2G use cases, where the question extends beyond data retrieval to whether the data is appropriate for use.

12.8 Implication for AAS

These use cases show why AAS should be integrated where it adds value. AAS submodel templates can become source profiles or SHACL shapes. AAS repositories can become graph sources. AAS APIs can become tools for AI agents. AAS RDF can participate directly in KG-RAG. However, the common denominator across B2G documents, customs data, DPP data, sensor data, and AI model metadata is the verifiable evidence graph rather than an AAS container defined in a local Industry 4.0 ecosystem.

The architectural conclusion is therefore sharper: AAS is useful where an asset-centric Industry 4.0 model is already present, adopted and required. For AI-first global supply chains and public-

sector evidence exchange, the default should be a Verifiable Linked Knowledge Graph with JSON-LD as the web/API serialization, YAML-LD for human-authored governance artefacts, CBOR-LD for compact and constrained exchange, SHACL for mandatory validation, and verifiable data mechanisms for issuer binding and integrity.

13 Reference architecture

The recommended architecture is knowledge-graph-first with AAS compatibility. It lifts existing systems into a shared evidence layer, which is required for AI applications in higher stakes B2B, B2G or G2G use cases.

For EU AI Act, IPCEI-AI, eIDAS 2.0, and European Business Wallet deployments, this evidence layer should include enterprise identity credentials, Agent PoA credentials, and an AI Service Passport evidence bundle. The AI Service Passport should be graph-linked rather than a static PDF. It should connect the AI service identifier, provider and deployer identities, delegated agent authorities, model and data lineage, conformity evidence, operational logs, human-oversight responsibilities, cybersecurity controls, status information, and post-market monitoring evidence.

Where European Digital Identity Wallet and European Business Wallet infrastructure is available, the reference architecture should treat wallet-derived credentials as organisational roots of trust for regulated Industrial AI. The EBW can anchor legal-person identity, business-register identifiers, official documents, electronic seals, timestamps, mandates, and Agent PoA credentials. The VLKG should reference these credentials and status checks as graph evidence while remaining separate from wallet provision. This separation lets the same evidence graph verify who is acting, under which authority, for which AI service, with which proof, and under which policy.

Reference architecture
Source systems / AAS / ERP / MES / PLM / IoT / registries / public registers
→ JSON-LD, AAS-to-RDF, or other graph projection
→ Verifiable Linked Knowledge Graph
→ SHACL and policy validation for mandatory claims
→ VC/Data Integrity or VC-JOSE-COSE proof layer
→ eIDAS / European Business Wallet enterprise identity, Agent PoA credentials, and AI Service Passport evidence bundle
→ KG-RAG / GraphRAG / agentic workflows
→ AI output with evidence path, proof check, status check, and audit trail

This architecture separates concerns. AAS remains an industrial adapter and source model. JSON-LD is the default web/API serialization. YAML-LD supports human-authored specifications and examples. CBOR-LD supports constrained environments. RDF/OWL supplies graph semantics and open-world reasoning. SHACL supplies mandatory validation. VCs, Data Integrity, and VC-JOSE-COSE supply issuer binding and integrity. GraphRAG supplies compact evidence retrieval. The AI layer retrieves from an evidence graph instead of inferring trust from free text.

The architecture should expose two additional control services. First, a Know Your Data and Data Risk Scoring service evaluates retrieved evidence paths before generation. Second, a cryptographic agility service records algorithms, verification methods, proof suites, key status, and PQC migration state in the graph. These services allow an AI agent to evaluate both what evidence exists and whether the evidence is trusted, current, policy-compliant, and cryptographically acceptable for the intended decision.

For IPCEI-AI and EU AI Act deployments, these control services make the architecture suitable for automated compliance controls, functional safety, and cyber-resilience in industrial and physical AI systems. They allow a compliance agent to verify the claim together with the enterprise authority, delegated agent authority, service status, evidence validity, and policy context under which the claim may be used.

The architecture can also absorb AAS add-ons where useful. AAS APIs can be tools for AI agents. AAS submodels can become graph shapes or application profiles. AAS RDF can be loaded directly into graph stores. AASX packages can be ingested as document containers. AAS security concepts can align with OAuth/OIDC and dataspace policies. The universal AI substrate, however, remains the evidence graph.

14 Comparative assessment

Table 4: Comparative assessment of AAS and Verifiable Linked Knowledge Graph options.

Criterion	**AAS JSON/XML**	**AAS RDF / Semantic AAS**	**Direct VLKG with JSON-LD**
AI grounding	Medium: structured source data, but not graph-native.	High: graph paths possible.	High, when governed and used for evidence-path retrieval.
Open-world integration	Limited: schema-bound payloads.	Good if RDF/OWL used.	Native design principle.
Closed-world validation	Good through schemas and templates.	Good through SHACL and shapes.	Good through SHACL, JSON Schema, policies, verifier rules.
Verifiability	Possible through packages/security, but not claim-native by default.	Possible if proof layer is added.	Core property through VCs/Data Integrity/JOSE/COSE.
Global developer familiarity	Specialist, strongest in German/EU manufacturing.	Specialist semantic-web plus AAS knowledge.	Higher entry path via JSON and web Linked Data.
Operational technology fit	Strong.	Moderate to strong with adapters.	Requires adapters to OT systems.
Global supply-chain fit	Moderate; asset-centric.	High if graph linked.	Very high; claim-centric and web-native.
Critical-infrastructure fit	Good inside AAS ecosystems.	High with access/proof governance.	High, especially for audit and multi-party evidence.
B2G fit	Good where AAS is mandated.	High if RDF/SHACL used.	Very high for linked, verifiable public reporting.
Implementation complexity	High for non-AAS developers.	High because both AAS and RDF skills are needed.	Medium to high, but modular and web-native.
Recommended role	Industrial source model and adapter.	Bridge between AAS and AI grounding.	Preferred default for multi-party AI evidence workflows.

The assessment supports a balanced decision. AAS is strongest where asset-centric industrial interoperability is the dominant requirement. A direct VLKG is strongest where AI grounding, global interoperability, verifiable claims, public reporting, and multi-party evidence are dominant. In many global Industrial AI strategies, the second set of requirements is more urgent.

15 Recommendations for Germany and Europe

1. Adopt a knowledge-graph-first strategy. If justified by ecosystem adoption, the AAS can remain an important adapter and source model, but the primary AI data plane should be a Verifiable Linked Knowledge Graph.
2. Use JSON-LD as the default web and API serialization for industrial claims. It provides a practical upgrade path from existing JSON systems to graph-compatible data.
3. Use YAML-LD for human-authored specifications, rules, examples, agent skills, and configuration artefacts. This supports governance and developer readability while preserving linked-data semantics.
4. Use CBOR-LD for constrained, edge, offline, barcode, IoT, and bandwidth-sensitive use cases where compact linked data is needed.
5. Use open-world semantic vocabularies for integration and closed-world validation for compliance. Missing data should mean unknown, rather than false, but mandatory claims must be checked with SHACL, schemas, and policy rules.
6. Use verifiable data mechanisms for issuer binding, integrity, status, and auditability. Claims should be tied to issuers and proofs where they are used as AI evidence.
7. Introduce Know Your Data and Data Risk Scoring as core Industrial AI controls. AI systems should evaluate issuer trust, proof status, validation results, data freshness, provenance, conflicts, policy permissions, and cryptographic risk before using evidence in high-stakes outputs.
8. Start the post-quantum transition for industrial evidence now. Germany and Europe should include DID-based key discovery, multi-proof credentials, crypto-agile verifier policies, and PQC migration metadata in the Industrial AI data-plane agenda.
9. Treat AAS submodel templates as useful source profiles, not as the universal semantic model. Where an AAS profile is required, project it into RDF/JSON-LD and integrate it into the VLKG.
10. Develop industrial KG-RAG benchmarks. Germany and Europe should fund direct benchmarks comparing AAS JSON/XML, AAS RDF, JSON-LD knowledge graphs, and verifiable graphs for AI grounding, token use, compliance accuracy, and auditability.
11. Align dataspace, IPCEI-AI, EU AI Act, eIDAS 2.0, and European Business Wallet implementation initiatives with W3C semantic and verifiable data standards. Gaia-X already points toward verifiable credentials and linked data representations for trusted statements [68], and Catena-X Knowledge Agents use W3C Semantic Web standards such as OWL, SPARQL, SHACL, and RDF [69]. IPCEI-AI and EU AI Act implementation should also use enterprise identity credentials, Agent PoA credentials, AI Service Passports, and graph-linked evidence to automate risk management, data governance, technical documentation, logging, human oversight, cybersecurity, and post-market monitoring controls [15–17]. eIDAS 2.0 and the proposed European Business Wallet should be treated as the trust infrastructure for enterprise identity, delegation, electronic attestations, seals, timestamps, and cross-border verification in these evidence graphs [18–20].
12. Prioritise globally portable evidence infrastructure. Germany and Europe should export the principle of verifiable, linked, open-world industrial evidence rather than making the AAS container model the default for every global data-sharing context.

16 Limitations and research agenda

The first limitation is empirical. There is no strong public benchmark that directly compares AAS JSON/XML, AAS RDF, direct JSON-LD knowledge graphs, and Verifiable Linked Knowledge Graphs for Industrial AI. Existing KG-RAG and GraphRAG results are proxy evidence. They support graph-grounded AI while leaving the superiority of any one industrial standard unresolved.

The second limitation is implementation maturity. JSON-LD, RDF, SHACL, VCs, Data Integrity, and DIDs are mature or maturing standards, but deployments still require governance, trusted issuer registries, key management, status mechanisms, privacy design, and operational monitoring. A VLKG requires disciplined implementation. This implementation effort can be justified by the expected benefits at the application level. As laid out in this paper, a verifiable evidence graph can reduce token use in AI-supported processes because relevant evidence can be retrieved, filtered, and presented in a structured form instead of being repeatedly supplied as large unstructured context. It can also support more reliable AI model output because claims, sources, constraints, and verification status are explicit and machine-checkable. This is especially relevant for functional safety, cyber-security, and reliable compliance automation, where unsupported model output is not sufficient. In these contexts, the VLKG can act as a controlled evidence layer for AI-based reasoning, decision support, and audit preparation. A full system-level cost-benefit justification, including implementation, operating, assurance, and liability costs, still needs to be developed and is left to future work.

The third limitation is organisational. Industrial data is political. Suppliers may resist disclosure, OEMs may seek control, regulators may mandate specific reporting models, and public authorities may require legal certainty. Technical formats contribute to governance by making good governance machine-readable and enforceable.

Future research should address six questions. First, what benchmark tasks best measure evidence-path reasoning in industrial AI? Second, how should signed claims, provenance, SHACL validation, and graph retrieval be evaluated together? Third, how can AAS submodels be mapped into JSON-LD/RDF without losing operational meaning? Fourth, how should token efficiency, latency, cost, and environmental impact be measured for industrial KG-RAG systems?

Fifth, research should examine Know Your Data and Data Risk Scoring as graph-native AI controls, including whether risk-aware evidence selection improves factuality, security, auditability, and abstention under adversarial data-plane conditions. Sixth, research should examine PQC migration for verifiable industrial evidence, including multi-proof credentials, DID-based key rotation, hybrid proof policies, and long-term validation of product, infrastructure, and compliance records.

17 Discussion and conclusion

The decisive question for Industrial AI has shifted from whether industrial assets can be described to whether AI systems can construct, verify, and explain evidence paths across distributed industrial claims. AAS answers the first question well in asset-centric Industry 4.0 settings. A Verifiable Linked Knowledge Graph answers the second more directly.

AAS should therefore be repositioned as an industrial source model, interoperability adapter, and digital-twin container rather than the natural universal semantic substrate for global AI-first data strategies. When AAS is used, it should be lifted into RDF/Linked Data and connected to proofs, status mechanisms, validation, and graph retrieval. Where no ecosystem requirement

mandates AAS, a direct JSON-LD-based VLKG is often the better and much less complex default.

From a German perspective, this shift matters. Germany has made major contributions to Industry 4.0, but should avoid converting a historically bounded architecture into unnecessary global complexity. The opportunity is to move from asset administration to evidence infrastructure. For global supply chains, critical infrastructure, B2G reporting, and Industrial AI, the priority is to get the data plane right: open-world linked data, persistent identifiers, typed relations, provenance, issuer-bound claims, cryptographic integrity, status checks, validation rules, and graph-guided retrieval.

The best data format for Industrial AI is therefore a layered architecture rather. JSON-LD is the default web serialization. YAML-LD and CBOR-LD extend the same graph into human-authored and constrained environments. RDF/OWL, SHACL, and W3C verifiable data mechanisms provide semantics, validation, and trust. AAS remains useful where it fits. The AI-first data plane should be the Verifiable Linked Knowledge Graph.

The practical design rule is simple: use AAS where the industrial ecosystem already requires it; use JSON-LD and RDF/OWL where AI needs graph-native meaning; use SHACL where mandatory evidence must be checked; and use verifiable data, eIDAS 2.0 trust services, and the European Business Wallet where an AI answer depends on issuer-bound claims. This layered position avoids a false choice between AAS and knowledge graphs and gives Germany and Europe an AI-first path that can scale beyond their industrial policy ecosystem. In regulated industrial and physical AI, the European Business Wallet can provide a distinctive European trust layer by connecting enterprise identity, delegated authority, and verifiable operational evidence to a legally backed digital-identity and trust-services framework [18–20].

References

[1] W3C. "JSON-LD 1.1 - A JSON-based Serialization for Linked Data." W3C Recommendation, 2020. https://www.w3.org/TR/json-ld11/

[2] W3C. "YAML-LD 1.0." W3C Working Draft, 2026. https://www.w3.org/TR/yaml-ld-10/

[3] W3C. "CBOR-LD 1.0." W3C Working Draft, 2026. https://www.w3.org/TR/cbor-ld-10/

[4] W3C. "RDF 1.1 Concepts and Abstract Syntax." W3C Recommendation, 2014. https://www.w3.org/TR/rdf11-concepts/

[5] W3C. "OWL 2 Web Ontology Language Primer." W3C Recommendation, 2012. https://www.w3.org/TR/owl2-primer/

[6] W3C. "Shapes Constraint Language (SHACL)." W3C Recommendation, 2017. https://www.w3.org/TR/shacl/

[7] W3C. "Verifiable Credentials Data Model v2.0." W3C Recommendation, 2025. https://www.w3.org/TR/vc-data-model-2.0/

[8] W3C. "Verifiable Credential Data Integrity 1.0." W3C Recommendation, 2025. https://www.w3.org/TR/vc-data-integrity/

[9] W3C. "Securing Verifiable Credentials using JOSE and COSE." W3C Recommendation, 2025. https://www.w3.org/TR/vc-jose-cose/

[10] Industrial Digital Twin Association. "Specification of the Asset Administration Shell Part 1: Metamodel." IDTA-01001, Version 3.0/3.2 series. https://industrialdigitaltwin.io/aas-specifications/IDTA-01001/v3.2/part-1-metamodel.pdf

[11] IEC. "IEC 63278-1:2023 - Asset Administration Shell for industrial applications - Part 1: Asset Administration Shell structure." https://webstore.iec.ch/en/publication/65628

[12] Industrial Digital Twin Association. "IDTA-01002 - APIs of the Asset Administration Shell." GitHub repository with normative OpenAPI files. https://github.com/admin-shell-io/aas-specs-api

[13] Industrial Digital Twin Association. "Specification of the Asset Administration Shell Part 4: Security - IDTA Number 01004." https://industrialdigitaltwin.org/en/content-hub/aasspecifications/specification-of-the-asset-administration-shell-part-4-security-idta-number-01004

[14] Industrial Digital Twin Association. "AAS Submodel Templates." https://industrialdigitaltwin.org/en/content-hub/submodels

[15] European Commission. "Important Projects of Common European Interest (IPCEI)." Competition Policy, State Aid. https://competition-policy.ec.europa.eu/state-aid/ipcei_en

[16] Federal Ministry for Economic Affairs and Energy (BMWE). "IPCEI Artificial Intelligence: Sector-specific AI for a competitive Europe." https://www.bundeswirtschaftsministerium.de/Redaktion/EN/Dossier/ipcei-ai.html

[17] European Parliament and Council. "Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)." Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

[18] European Parliament and Council. "Regulation (EU) 2024/1183 of 11 April 2024 amending Regulation (EU) No 910/2014 as regards establishing the European Digital Identity Framework." Official Journal of the European Union, 2024. https://eur-lex.europa.eu/eli/reg/2024/1183/oj

[19] European Commission. "Proposal for a Regulation on the establishment of European Business Wallets." COM(2025) 838 final, 2025. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-establishment-european-business-wallets ; European Commission. "European Business Wallets." Shaping Europe's Digital Future, 2025. https://digital-strategy.ec.europa.eu/en/policies/business-wallets

[20] Council of the European Union. "European business wallets: Council adopts negotiating position." Press release, 9 June 2026. https://www.consilium.europa.eu/en/press/press-releases/2026/06/09/european-business-wallets-council-adopts-negotiating-position/

[21] Wahlster, W. "Prof. Wolfgang Wahlster on Industrie 4.0." RICAIP, 2018. Includes discussion of DFKI Smart Factory work since 2005 and SemProM since 2008. https://ricaip.eu/prof-wolfgang-wahlster-on-industrie-4-0/

[22] Kagermann, H., Wahlster, W., and Helbig, J. "Recommendations for implementing the strategic initiative INDUSTRIE 4.0." acatech, 2013. https://en.acatech.de/publication/recommendations-for-implementing-the-strategic-initiative-industrie-4-0-final-report-of-the-industrie-4-0-working-group/

[23] VDI/VDE and ZVEI. "Reference Architecture Model Industrie 4.0 (RAMI 4.0)." Status Report, 2015. https://www.zvei.org/fileadmin/user_upload/Presse_und_Medien/Publikationen/2016/januar/GMA_Status_Report___Reference_Archtitecture_Model_Industrie_4.0___RAMI_4.0_/GMA-Status-Report-RAMI-40-July-2015.pdf

[24] Lewis, P. et al. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks." NeurIPS 2020. https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html

[25] Agrawal, G. et al. "Can Knowledge Graphs Reduce Hallucinations in LLMs? A Survey." NAACL 2024. https://aclanthology.org/2024.naacl-long.219/

[26] Lavrinovics, E., Biswas, R., Hose, K., and Bjerva, J. "MultiHal: Multilingual Dataset for Knowledge-Graph Grounded Evaluation of LLM Hallucinations." arXiv:2505.14101, 2025. https://arxiv.org/abs/2505.14101

[27] Barry, M. et al. "GraphRAG: Leveraging Graph-Based Efficiency to Minimize Hallucinations in LLM-Driven RAG for Finance Data." GenAIK 2025. https://aclanthology.org/2025.genaik-1.6/

[28] Peng, B. et al. "Graph Retrieval-Augmented Generation: A Survey." arXiv:2408.08921, 2024. https://arxiv.org/abs/2408.08921

[29] Rongen, S. et al. "Modelling with AAS and RDF in Industry 4.0." Computers in Industry, 2023. https://www.sciencedirect.com/science/article/pii/S016636152300060X

[30] Hogan, A. et al. "Knowledge Graphs." ACM Computing Surveys, 2021. https://dl.acm.org/doi/fullHtml/10.1145/3447772

[31] Tiddi, I. and Schlobach, S. "Knowledge Graphs as Tools for Explainable Machine Learning: A Survey." Artificial Intelligence 302, 2022. https://www.sciencedirect.com/science/article/pii/S0004370221001788

[32] Yahya, M. et al. "Semantic Web and Knowledge Graphs for Industry 4.0." Applied Sciences 11(11):5110, 2021. https://www.mdpi.com/2076-3417/11/11/5110

[33] Bader, S. R. and Maleshkova, M. "The Semantic Asset Administration Shell." ISWC 2019 Satellite Tracks. https://arxiv.org/abs/1909.00690

[34] Beden, S. et al. "Semantic Asset Administration Shells in Industry 4.0: A Survey." 2021. https://cronfa.swan.ac.uk/Record/cronfa56661/Download/56661___19655___5c66fc9820b04e74b43744d07b06af94.pdf

[35] Rimaz, M. H. et al. "Semantic Asset Administration Shell for Circular Economy." 2024. https://www.dfki.de/fileadmin/user_upload/import/15613_paper1.pdf

[36] European Commission. "Commission launches consultation on the Digital Product Passport." 2025. https://single-market-economy.ec.europa.eu/news/commission-launches-consultation-digital-product-passport-2025-04-09_en

[37] CIRPASS. "Digital Product Passport." https://cirpassproject.eu/ ; CIRPASS-2. https://cirpass2.eu/

[38] W3C and OGC. "Semantic Sensor Network Ontology." W3C Recommendation, 2017. https://www.w3.org/TR/vocab-ssn/

[39] W3C. "Web of Things (WoT) Thing Description 1.1." W3C Recommendation, 2023. https://www.w3.org/TR/wot-thing-description11/

[40] W3C. "PROV-O: The PROV Ontology." W3C Recommendation, 2013. https://www.w3.org/TR/prov-o/

[41] W3C. "Time Ontology in OWL." W3C Recommendation, 2022. https://www.w3.org/TR/owl-time/

[42] Google Search Central. "Introduction to structured data markup in Google Search." https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data

[43] HTTP Archive. "Structured data." The 2024 Web Almanac. https://almanac.httparchive.org/en/2024/structured-data

[44] W3C. "Bitstring Status List v1.0." W3C Recommendation, 2025. https://www.w3.org/TR/vc-bitstring-status-list/

[45] NIST. "Artificial Intelligence Risk Management Framework (AI RMF 1.0)." NIST AI 100-1, 2023. https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

[46] Vassilev, A. et al. "Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations." NIST AI 100-2e2025, 2025. https://csrc.nist.gov/pubs/ai/100/2/e2025/final

[47] NIST. "Cybersecurity Framework Profile for Artificial Intelligence." NIST IR 8596 Initial Public Draft, 2025. https://nvlpubs.nist.gov/nistpubs/ir/2025/NIST.IR.8596.iprd.pdf

[48] OWASP. "OWASP Top 10 for Large Language Model Applications." 2025 project documentation. https://owasp.org/www-project-top-10-for-large-language-model-applications/

[49] NIST. "FIPS 203: Module-Lattice-Based Key-Encapsulation Mechanism Standard." 2024. https://csrc.nist.gov/pubs/fips/203/final

[50] NIST. "FIPS 204: Module-Lattice-Based Digital Signature Standard." 2024. https://csrc.nist.gov/pubs/fips/204/final

[51] NIST. "FIPS 205: Stateless Hash-Based Digital Signature Standard." 2024. https://csrc.nist.gov/pubs/fips/205/final

[52] NIST NCCoE. "Migration to Post-Quantum Cryptography." https://www.nccoe.nist.gov/applied-cryptography/migration-to-pqc

[53] CISA, NSA, and NIST. "Quantum-Readiness: Migration to Post-Quantum Cryptography." 2023. https://www.cisa.gov/resources-tools/resources/quantum-readiness-migration-post-quantum-cryptography

[54] W3C. "Decentralized Identifiers (DIDs) v1.1." W3C Candidate Recommendation Snapshot, 2026. https://www.w3.org/TR/did-1.1/

[55] Jegham, N. et al. "How Hungry is AI? Benchmarking Energy, Water, and Carbon Footprint of LLM Inference." arXiv:2505.09598, 2025. https://arxiv.org/abs/2505.09598

[56] Wilhelm, P., Wittkopp, T., and Kao, O. "Beyond Test-Time Compute Strategies: Advocating Energy-per-Token in LLM Inference." EuroMLSys 2025. https://euromlsys.eu/pdf/euromlsys25-27.pdf

[57] Schema.org. "Schema.org." Structured data vocabulary adoption statement. https://schema.org/

[58] W3C. "VSSo: Vehicle Signal Specification Ontology." W3C Discontinued Draft, 2024. https://www.w3.org/TR/vsso/

[59] GS1. "GS1 Digital Link URI Syntax." GS1 Standard. https://ref.gs1.org/standards/digital-link/uri-syntax/

[60] European Commission. "NIS2 Directive: securing network and information systems." https://digital-strategy.ec.europa.eu/en/policies/nis2-directive

[61] European Commission. "Data Act." Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/data-act

[62] Publications Office of the European Union. "eProcurement Ontology." https://ted.europa.eu/en/simap/eprocurement-ontology

[63] European Commission. "Once-Only Technical System High Level Architecture." 2023. https://ec.europa.eu/digital-building-blocks/sites/spaces/TDD/pages/706382128/1.+Once-Only+Technical+System+High+Level+Architecture+December+2023

[64] World Customs Organization. "WCO Data Model." Version 4 series and Version 4.2.0 release information. https://www.wcoomd.org/DataModel

[65] Mitchell, M. et al. "Model Cards for Model Reporting." Proceedings of FAT* 2019. https://dl.acm.org/doi/10.1145/3287560.3287596

[66] Gebru, T. et al. "Datasheets for Datasets." Communications of the ACM, 2021. https://dl.acm.org/doi/10.1145/3458723

[67] W3C. "Data Catalog Vocabulary (DCAT) - Version 3." W3C Recommendation, 2024. https://www.w3.org/TR/vocab-dcat-3/

[68] Gaia-X. "Gaia-X Trust Framework." 22.10 Release. https://docs.gaia-x.eu/policy-rules-committee/trust-framework/22.10/

[69] Catena-X. "CX-0084 Federated Queries in Data Spaces." https://catenax-ev.github.io/docs/next/standards/CX-0084-FederatedQueriesInDataSpaces