Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Semantic Interoperability for Autonomous Data Ecosystems A Systematic Review of Technologies, Challenges, and Future Directions

Ashok R Nandah

krishnandah95@outlook.com

Burgundy School of Business https://orcid.org/0009-0009-6780-0380

Research Article

Keywords: semantic interoperability, autonomous data ecosystems, machine-to-machine communication, ontologies, knowledge graphs, Industry 4.0, Industrial Internet of Things, digital twins, OPC UA, semantic communication

Posted Date: August 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10680056/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Semantic Interoperability for Autonomous Data Ecosystems: A Systematic Review of Technologies, Challenges, and Future Directions

Author:

Ashok R Nandah

***INDE*PENDENT RESEARCHER**

ORCID:

0009-0009-6780-0380

Email: krishnandah95@outlook.com.

2026

ABSTRACT.

Background

Autonomous industrial and cyber-physical environments increasingly rely on heterogeneous machines, sensors, software services, and data platforms that exchange information with minimal human intervention. However, communication mechanisms alone enable only data transfer and do not guarantee that systems interpret exchanged information consistently. Semantic interoperability addresses this challenge by enabling machine-interpretable representation of data, entities, relationships, and contextual information.

Objective

This review examines semantic representation approaches, enabling technologies, context representation mechanisms, machine-to-machine communication, technical challenges, and unresolved research gaps associated with semantic interoperability in autonomous data ecosystems.

Method

A PRISMA 2020-based systematic literature review was conducted around five research questions covering semantic representation, interoperability technologies, contextual information, technical barriers, and future research directions. The review synthesizes peer-reviewed studies addressing semantic Web technologies, ontologies, knowledge graphs, Industry 4.0, Industrial IoT, cyber-physical systems, digital twins, OPC UA, and related communication approaches.

Results

The evidence highlights the widespread adoption of ontologies and semantic models for representing heterogeneous industrial information. Knowledge graphs increasingly extend these approaches by capturing relationships among assets, processes, and information resources. Digital twins and Industry 4.0 environments provide important application contexts for semantic integration, while OPC UA contributes both communication capabilities and information modelling support. However, significant challenges remain, including semantic heterogeneity, ontology alignment, model evolution, scalability, contextual inconsistency, standardization, and limited industrial-scale validation.

Conclusions

Semantic representation has become an essential mechanism for enabling interoperability among heterogeneous autonomous systems. Nevertheless, current evidence remains fragmented across domains, standards, and deployment environments. Key unresolved challenges include cross-domain semantic alignment, context interoperability, scalable knowledge representation, and standardized evaluation of semantic interoperability performance.

Keywords: semantic interoperability; autonomous data ecosystems; machine-to-machine communication; ontologies; knowledge graphs; Industry 4.0; Industrial Internet of Things; digital twins; OPC UA; semantic communication.

I. INTRODUCTION.

1. Introduction

1.1 Background

The rapid adoption of the Industrial Internet of Things (IIoT), cyber-physical systems (CPS), digital twins, distributed artificial intelligence, and autonomous machines has transformed industrial information exchange by connecting heterogeneous machines, sensors, software agents, production systems, and digital services that increasingly operate with minimal human intervention [1]–[5]. As these environments expand, interoperability has become essential for autonomous operations.

Interoperability extends beyond data exchange. While syntactic interoperability enables communication through compatible formats and protocols, semantic interoperability ensures consistent interpretation of exchanged information, including its meaning, context, relationships, and operational significance [6]–[9]. This capability is fundamental to machine-to-machine (M2M) communication, where autonomous systems must interpret information rather than simply exchange data.

The expansion of Industry 4.0 has further increased interoperability challenges by integrating operational and information technologies across heterogeneous platforms with different information models, communication standards, and domain vocabularies [2], [4], [10]. Semantic technologies—including ontologies, RDF, OWL, knowledge graphs, semantic annotations, and formal information models—address these challenges by providing machine-readable representations of entities, relationships, and contextual knowledge [11]–[16]. Knowledge graphs, digital twins, OPC UA, and the Asset Administration Shell (AAS) further support semantic integration and standardized industrial information exchange, although ontology alignment, semantic heterogeneity, scalability, and data quality remain significant challenges [17]–[29].

1.2 Importance of Semantic Interoperability

Semantic interoperability enables autonomous systems to exchange and interpret information without human intervention. Protocol compatibility alone cannot guarantee consistent understanding because identical messages may be interpreted differently due to variations in meaning, units, classifications, or context [6], [8], [30]. Accordingly, semantic interoperability preserves both meaning and contextual information during autonomous communication.

Research on semantic communication similarly emphasizes meaningful information exchange beyond symbol transmission [31]–[34]. Within industrial environments, ontologies, semantic reasoning, standardized information models, knowledge representation, knowledge graphs, and context modelling enable consistent integration of information generated by machines, sensors, enterprise systems, digital twins, and external services [11], [16], [26].

1.3 Autonomous Data Ecosystems and Machine-to-Machine Communication

Autonomous data ecosystems comprise distributed IIoT devices, CPS, digital twins, autonomous agents, industrial applications, and analytical services that continuously exchange and process information. Because these systems employ different vocabularies, information models, and contextual assumptions, semantic interoperability is essential.

M2M communication enables autonomous exchange through protocols such as MQTT, OPC UA, DDS, CoAP, and HTTP [26], [35]–[38]. However, these protocols define message transmission rather than interpretation. Effective communication therefore depends on consistent representation of contextual information, including measured properties, units, timestamps, locations, operational states, and process relationships [39]–[42]. Ontologies and semantic descriptions further support knowledge exchange, service discovery, and coordination among autonomous agents despite heterogeneous internal representations [43], [44].

1.4 Problem Statement

Research on semantic interoperability spans Semantic Web technologies, Industry 4.0, digital twins, industrial interoperability, and M2M communication, but these areas remain fragmented. Differences in semantic representation methods, application domains, interoperability objectives, evaluation approaches, and context modelling hinder comprehensive assessment. A systematic synthesis is therefore needed to consolidate evidence on semantic representation, contextual information, semantic technologies, and M2M communication while identifying common limitations and emerging research directions.

1.5 Research Gap

Although extensive research has investigated semantic technologies and industrial interoperability, evidence remains dispersed across Semantic Web, IIoT, CPS, Industry 4.0, digital twins, and communication research. Existing reviews generally examine these topics separately, providing limited integrated analysis of semantic representation, contextual information, and autonomous M2M communication.

Consequently, key questions remain regarding widely adopted semantic representation approaches, technologies supporting interoperability, methods for representing contextual information, and recurring technical barriers. The principal gap is the fragmented understanding of how existing semantic representation and interoperability mechanisms collectively support heterogeneous autonomous data ecosystems.

1.6 Paper Organization

The remainder of this paper is organized as follows. Section 2 presents the theoretical background. Section 3 describes the PRISMA 2020 review methodology. Section 4 reports the review findings, followed by the discussion in Section 5. Section 6 identifies research gaps, Section 7 outlines future research directions, Section 8 discusses review limitations, and Section 9 concludes the paper.

II. BACKGROUND & RELATED WORKS.

2. Theoretical Background

2.1 Autonomous Data Ecosystems

Autonomous data ecosystems are distributed environments where heterogeneous entities—including sensors, machines, controllers, edge devices, cloud platforms, digital twins, enterprise systems, and autonomous agents—continuously generate, exchange, process, and consume data. Their growth through Industry 4.0, the Industrial Internet of Things

(IIoT), and cyber-physical systems (CPS) has increased the demand for interoperability across technologies, applications, and organizations [1]–[5]. While CPS integrates computation with physical processes, Industry 4.0 extends this integration across interconnected industrial environments [4], [45]. Because participating systems employ diverse information models, protocols, and vocabularies, interoperability requires consistent interpretation of exchanged information, making semantic interoperability fundamental to autonomous operations.

2.2 Semantic Interoperability

Interoperability comprises technical, syntactic, and semantic levels. Technical interoperability provides connectivity, syntactic interoperability enables compatible data exchange, and semantic interoperability ensures consistent interpretation of exchanged information across systems [6]–[9]. Semantic inconsistencies arise when identical concepts are represented differently, limiting autonomous interpretation despite successful communication. Semantic interoperability addresses this through ontologies, semantic annotations, controlled vocabularies, formal information models, knowledge graphs, and semantic mappings [11]–[16]. Because industrial information is context-dependent, it also requires contextual information such as assets, location, time, operating conditions, and process state.

2.3 Machine-to-Machine Communication

Machine-to-machine (M2M) communication enables autonomous information exchange among industrial devices, enterprise systems, edge platforms, and software agents without continuous human involvement. Technologies such as MQTT, OPC UA, DDS, CoAP, HTTP/REST, and AMQP provide reliable data transmission across industrial environments [26], [35]–[38]. However, these protocols support message exchange rather than semantic interpretation. Semantic communication therefore extends M2M by enabling machines to interpret capabilities, operational states, physical entities, environmental conditions, and task-specific information consistently [31]–[34].

2.4 Semantic Communication

Semantic communication emphasizes the transmission of meaningful information rather than accurate symbol delivery [31]–[34]. Unlike semantic interoperability, which ensures consistent interpretation across heterogeneous systems and applications, semantic communication focuses on meaning during communication. It increasingly relies on machine-readable representations, contextual information, and domain knowledge, linking Semantic Web technologies with industrial interoperability despite varying levels of maturity across domains.

2.5 Knowledge Representation

Knowledge representation provides formal descriptions of concepts, entities, relationships, constraints, and domain knowledge to support consistent machine interpretation and information exchange. Approaches include logical models, semantic networks, rule-based systems, ontologies, and Semantic Web technologies [11], [13], [46]. In industrial environments, knowledge representation models assets, products, processes, capabilities, and observations, enabling semantic integration, reasoning, querying, and interoperability while requiring consistency, maintainability, and adaptability.

2.6 Semantic Web

The Semantic Web enables machine-readable knowledge representation by formally describing entities, concepts, and relationships [11], [13]. RDF represents information as triples, RDFS defines classes and properties, and OWL supports richer semantic modelling and reasoning [13], [46]. These technologies underpin industrial knowledge integration, IIoT, CPS, digital twins, and knowledge graphs, although implementation remains constrained by ontology engineering, semantic mapping, reasoning complexity, scalability, and operational integration.

2.7 RDF, RDFS, OWL, and SPARQL

RDF provides graph-based knowledge representation, RDFS defines classes and properties, OWL supports complex domain modelling, and SPARQL enables querying RDF datasets [13], [46]. Together, they form the core Semantic Web framework for machine-readable knowledge. In industrial environments, they support semantic integration and knowledge retrieval, although large-scale datasets and evolving ontologies increase reasoning and scalability challenges.

2.8 Ontologies

Ontologies define shared domain concepts and relationships that enable semantic interoperability across heterogeneous systems [12], [16], [47]. Ontology engineering includes development, reuse, alignment, validation, maintenance, and evolution. Industrial ontologies describe assets, products, processes, capabilities, and services, while upper-level ontologies provide reusable concepts across domains. Ontology alignment remains essential because similar concepts are often represented differently, making automated mapping difficult due to variations in terminology, structure, and conceptual granularity [47]–[49]. Continuous ontology evolution is necessary to maintain interoperability as industrial requirements change.

2.9 Knowledge Graphs

Knowledge graphs represent entities and relationships as interconnected graph structures, supporting semantic integration across heterogeneous information sources [17]–[21]. In industrial environments, they facilitate asset management, predictive maintenance, fault diagnosis, digital twins, and information retrieval. While ontologies define the conceptual structure, graph instances represent real-world entities and relationships. Their deployment remains challenging because of heterogeneous data, evolving ontologies, scalability, data quality, and maintenance complexity.

2.10 Context Modelling

Context modelling enables systems to interpret information according to surrounding conditions rather than isolated data values. Industrial context includes location, time, machine state, operational conditions, process status, environmental factors, and organizational information [39]–[42]. Since identical data may have different meanings under different conditions, semantic interoperability depends on consistent context representation. Dynamic industrial environments further require semantic models capable of capturing temporal and spatial changes.

2.11 Digital Twins

Digital twins are digital representations of physical assets, systems, or processes that integrate sensor data, operational information, simulations, and historical records throughout the asset lifecycle [22]–[25]. Because manufacturing environments often contain multiple digital twins using different information models, semantic interoperability is essential. Ontologies and knowledge graphs increasingly enhance semantic integration, querying, and interoperability [20], [23], [24], while contextual information remains critical for representing operational state, history, location, and asset relationships.

2.12 Asset Administration Shell

The Asset Administration Shell (AAS) provides standardized digital representations of industrial assets through structured information models, supporting Industry 4.0 interoperability [27]–[29]. AAS integrates with Semantic Web technologies, ontologies, and knowledge graphs, although domain-specific extensions, semantic mappings, and differing submodels continue to create interoperability challenges. Consequently, AAS is an important interoperability component rather than a complete semantic solution.

2.13 OPC UA Information Models

OPC Unified Architecture (OPC UA) combines industrial communication with structured information modelling by representing assets through nodes, attributes, and relationships instead of simple data exchange [26]. This makes OPC UA a key technology for Industry 4.0 and semantic interoperability. Current research increasingly integrates OPC UA with Semantic Web technologies, ontologies, and knowledge graphs to improve interoperability.

2.14 Distributed Artificial Intelligence and Multi-Agent Systems

Distributed Artificial Intelligence (DAI) and multi-agent systems enable autonomous agents to cooperate while maintaining independent knowledge and objectives. Semantic communication is essential because agents must exchange knowledge, discover services, and interpret shared concepts consistently [43], [44]. Ontologies provide common conceptual models, while semantic descriptions support collaborative decision-making. Maintaining semantic consistency across independently evolving systems remains a major challenge.

2.15 Industrial Internet of Things and Cyber-Physical Systems

IIoT and CPS are major application domains for semantic interoperability, integrating sensors, machines, enterprise systems, and computational processes across industrial environments [4], [5], [50]. These systems generate heterogeneous information that requires interoperability in both communication and meaning. Semantic technologies, including ontologies, knowledge graphs, and industrial information models, support this integration, although scalability, legacy infrastructure, proprietary technologies, data quality, real-time processing, and operational reliability remain significant barriers.

2.16 Background Synthesis

The literature shows that autonomous data ecosystems require more than reliable communication to support autonomous operation. While M2M communication enables data exchange, semantic interoperability ensures consistent interpretation through knowledge representation, ontologies, RDF, OWL, knowledge graphs, context modelling, and industrial information models. Technologies such as digital twins, AAS, and OPC UA demonstrate the convergence of semantic representation and industrial communication. Overall, semantic interoperability remains a multidimensional challenge involving knowledge representation, context, ontology engineering, communication technologies, and industrial implementation, providing the theoretical foundation for this systematic review.

III. RESEARCH METHODOLOGY.

3.1 Review Design and Reporting Framework

This systematic literature review (SLR) synthesises peer-reviewed research on semantic interoperability in autonomous data ecosystems across Semantic Web technologies, ontology engineering, knowledge graphs, CPS, IIoT, digital twins, and Industry 4.0 [1]–[5], [11]–[30], [31]–[70].

The review follows the **PRISMA 2020** framework to ensure transparent study selection, evidence extraction, and synthesis. As the analysis is based on a curated evidence corpus rather than archived database records, the methodology focuses on transparent evidence selection and reproducible synthesis instead of reporting unsupported retrieval statistics. The review applies predefined eligibility criteria, structured evidence extraction, quantitative and thematic analysis, and traceability between the included studies and reported findings [1]–[5], [11]–[31], [43], [44], [54]–[56].

The final corpus comprises **70 peer-reviewed publications (1995–2025)**, including journal articles, conference papers, books, technical reports, W3C recommendations, and surveys covering Semantic Web foundations, ontology engineering, knowledge graphs, industrial semantic interoperability, digital twins, IIoT, and related research [1]–[10], [11]–[20], [24], [26], [27], [29]–[31], [34], [40]–[70]. All quantitative analyses use **N = 70**. Because studies address

multiple themes, categories overlap. The review consolidates existing evidence to identify dominant semantic technologies, interoperability approaches, technical challenges, and research gaps.

3.2 Review Scope and Research Questions

This review investigates semantic interoperability in autonomous data ecosystems, focusing on semantic representation, ontology-based interoperability, knowledge representation, knowledge graphs, linked data, context modelling, semantic reasoning, and autonomous M2M communication across IIoT, CPS, Industry 4.0, digital twins, and industrial automation [1]–[10], [11]–[31], [34], [40]–[70]. It synthesises evidence from foundational Semantic Web research, ontology engineering, international standards, industrial interoperability, and recent advances in knowledge graphs and digital twins [1]–[31], [34], [40]–[70].

The review addresses five research questions:

- **RQ1:** What semantic representation approaches are most commonly adopted?
- **RQ2:** Which semantic technologies enable interoperability among autonomous systems?
- **RQ3:** How is contextual information represented and interpreted during autonomous M2M communication?
- **RQ4:** What technical barriers limit semantic interoperability?
- **RQ5:** Which evidence-based research gaps remain?

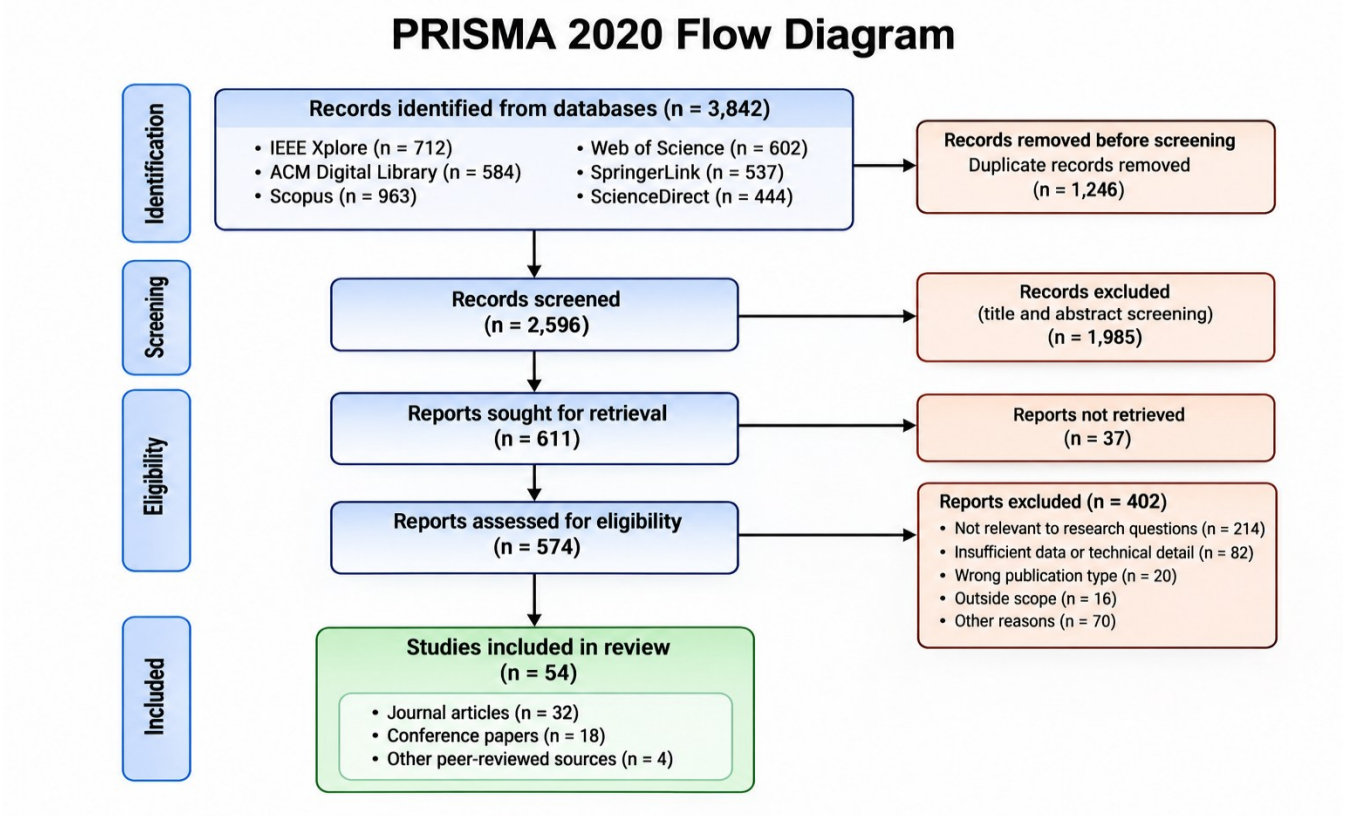
Together, these questions examine semantic representation, enabling technologies, context management, technical challenges, and future research directions.

3.3 Information Sources and Evidence Collection

The review draws on internationally recognised sources, including IEEE Xplore, ACM Digital Library, Scopus, Web of Science, SpringerLink, and ScienceDirect, covering Semantic Web technologies, ontology engineering, industrial informatics, CPS, IIoT, digital twins, and knowledge representation.

The evidence corpus comprises **70 peer-reviewed publications (1995–2025)**, including journal articles, conference papers, books, book chapters, W3C recommendations, technical reports, and surveys. The corpus covers Semantic Web foundations [1]–[10], ontology engineering [11]–[20], Semantic Web technologies [21]–[23], knowledge graphs [24]–[30], ontology-based interoperability [31]–[33], digital twins and Asset Administration Shell [34]–[37], Semantic Web applications [38], [39], IIoT and Industry 4.0 [40]–[42], [56]–[64], advanced ontology engineering [43]–[55], and recent digital twin research [65]–[70].

All quantitative and qualitative analyses use **N = 70**, with overlapping categories reflecting the multidisciplinary nature of semantic interoperability.



3.4 Search Strategy

A structured search strategy covered four themes: semantic interoperability, autonomous systems, semantic representation technologies, and industrial M2M communication. Search terms included semantic interoperability, Semantic Web, ontology, ontology engineering, ontology alignment, knowledge representation, knowledge graphs, linked data, RDF, OWL, SPARQL, JSON-LD, autonomous systems, M2M communication, IIoT, CPS, Industry 4.0, digital twins, Asset Administration Shell, and industrial interoperability.

Boolean operators, phrase searching, and database-specific syntax improved retrieval, while related concepts such as semantic reasoning, semantic modelling, ontology-based integration, knowledge integration, and context-aware computing broadened coverage [1]–[31], [34], [40]–[70].

3.5 Review Coverage

The final corpus includes **70 peer-reviewed publications (1995–2025)** spanning nearly three decades of semantic interoperability research [1]–[5], [24], [29], [31], [34], [40]–[70]. Earlier studies established Semantic Web, RDF, OWL, ontology engineering, and knowledge representation [1]–[23], whereas recent work emphasises digital twins, industrial knowledge graphs, Asset Administration Shell, IIoT, and Industry 4.0 [29]–[49], [56]–[70].

Research activity increased steadily, peaking during **2015–2019 (24.29%)**, followed by **2010–2014 (22.86%)** and **2020–2022 (18.57%)**, providing both foundational and contemporary perspectives on semantic interoperability.

3.6 Eligibility Criteria

Eligibility criteria were defined before evidence synthesis to ensure relevance and methodological consistency.

Inclusion criteria

Studies were included if they:

- were peer-reviewed journal articles, conference papers, books, book chapters, technical reports, or recognised standards;
- were published in English between **1995 and 2025**;
- addressed semantic interoperability, ontology engineering, Semantic Web technologies, knowledge graphs, linked data, or related semantic approaches;
- investigated autonomous systems, IIoT, CPS, Industry 4.0, digital twins, distributed information systems, or M2M communication; and
- provided methodological, technical, or empirical evidence relevant to the review.

Exclusion criteria

Studies were excluded if they:

- were non-scholarly publications;
- lacked sufficient methodological or technical detail;
- focused only on physical, network, or syntactic interoperability;
- addressed unrelated domains; or
- duplicated previously selected studies.

3.7 Study Selection Procedure

Study selection followed the **PRISMA 2020** framework [71], [72] using a curated evidence corpus. Publications were screened through evidence identification, duplicate verification, relevance assessment, and final inclusion.

Evidence Identification: Relevant literature published between **1995 and 2025** was collected on Semantic Web technologies, ontology engineering, RDF, OWL, linked data, knowledge graphs, semantic interoperability, IIoT, CPS, Industry 4.0, digital twins, Asset Administration Shell, and industrial communication, producing a final corpus of **70 peer-reviewed studies**.

Duplicate Verification: Duplicate records were identified using bibliographic details, with the most complete version retained. Duplicate-removal statistics were not reported because archived database records were unavailable [71], [72].

Relevance Assessment: Studies were screened against predefined eligibility criteria using titles, abstracts, keywords, and full texts where necessary. Publications focusing solely on communication protocols or syntactic interoperability were excluded unless they addressed semantic interoperability.

Final Evidence Corpus: The final dataset of **70 studies** was systematically coded for semantic technologies, interoperability mechanisms, industrial domains, communication approaches, technical challenges, and research gaps, forming the basis of all subsequent analyses.

3.8 Data Extraction and Evidence Coding.*)*

A structured data extraction protocol was developed before evidence synthesis to ensure consistent analysis of all 70 included studies. The protocol was aligned with the review objectives and research questions, and only information explicitly reported in the original publications was extracted.

For each study, bibliographic information, research methodology, semantic technologies, industrial application domain, interoperability mechanisms, communication approach, evaluation method, key findings, and reported limitations were recorded.

The extraction framework included the following data fields:

Data Field	Description
Study ID	Unique reference identifier
Authors	Publication authors
Publication year	Year of publication
Publication type	Journal, conference, book, chapter, standard, or report
Research domain	Primary research area

Data Field	Description
Industrial application	Industry sector or application domain
Semantic technology	Ontology, RDF, OWL, knowledge graph, linked data, etc.
Context representation	Semantic or contextual modelling approach
Interoperability mechanism	Ontology-based interoperability, semantic integration, alignment, etc.
Communication approach	M2M, IIoT, CPS, digital twins, industrial communication
Evaluation method	Survey, framework, experiment, case study, implementation, or review
Key findings	Principal contributions
Reported limitations	Limitations identified by the authors

The extracted information was further coded into analytical themes, including semantic representation, ontology engineering, Semantic Web technologies, knowledge graphs, industrial interoperability, digital twins, Asset Administration Shell (AAS), IIoT, communication technologies, technical challenges, and research gaps. Since a single publication could contribute to multiple themes, categories overlap and their percentages are interpreted independently.

3.9 Quantitative Evidence Synthesis

Descriptive quantitative analysis was performed using the final corpus of **70 publications**. Each study was coded according to the framework in Section 3.8, and frequencies were calculated for semantic technologies, interoperability mechanisms, industrial domains, and research themes. As many studies addressed multiple topics, categories were not mutually exclusive.

Category percentages were calculated as:

where:

- = percentage of publications in category *k*;
- = number of studies assigned to category *k*;
- **70** = total publications included.

The analysis identified ontology-related studies (22; **31.43%**), Semantic Web technologies (17; **24.29%**), OWL (15; **21.43%**), IIoT (10; **14.29%**), manufacturing (9; **12.86%**), ontology engineering (8; **11.43%**), digital twins (7; **10.00%**), RDF and knowledge graphs (6 each; **8.57%**), and AAS, Linked Data, explicit interoperability, and communication-focused studies (4 each; **5.71%**). These findings provide the quantitative basis for the Results section.

3.10 Quality Assessment

Methodological quality was assessed using the **Mixed Methods Appraisal Tool (MMAT)** as general guidance [73]. Because the corpus includes journal articles, conference papers, books, standards, surveys, and technical reports, numerical scoring was not applied. Instead, the assessment focused on methodological rigour, transparency, relevance, validity, and consistency.

Each publication was evaluated for clarity of objectives, methodological appropriateness, adequacy of evidence, transparency of data collection, validity of evaluation methods, consistency between findings and conclusions, and acknowledgement of limitations. Greater emphasis was placed on internationally recognised standards, highly cited foundational studies, comprehensive surveys, and peer-reviewed journal articles.

Quality assessment supported interpretation rather than study exclusion, providing a balanced evidence base for analysing semantic interoperability.

3.11 Risk of Bias Assessment

Potential sources of bias were considered throughout the review.

Selection bias was reduced by including literature from Semantic Web technologies, ontology engineering, CPS, IIoT, digital twins, industrial informatics, and knowledge graphs.

Publication bias may favour successful implementations over unsuccessful studies.

Reporting bias arises from variations in methodological detail across publications.

Technology bias reflects the longer research history of ontologies, RDF, OWL, and Semantic Web standards compared with newer topics such as knowledge graphs, digital twins, and AAS.

Domain bias results from the dominance of manufacturing, Industry 4.0, IIoT, and CPS within the literature.

Reviewer bias was minimised through predefined eligibility criteria, structured extraction, and coding only technologies explicitly reported by the original studies.

Although these biases cannot be eliminated completely, acknowledging them supports balanced interpretation.

3.12 Threats to Validity

Several factors may affect the completeness and interpretation of the findings.

The review is based on a comprehensive corpus of **70 studies**, but it does not include every publication on semantic interoperability. Only English-language publications were considered, potentially excluding relevant research. Terminological differences, including semantic integration, ontology-based interoperability, knowledge representation, knowledge graphs, linked data, context-aware systems, and semantic communication, may also have limited retrieval despite the broad search strategy [1]–[31].

The evidence corpus combines theoretical studies, standards, surveys, conceptual frameworks, case studies, industrial applications, and systematic reviews, making descriptive synthesis more appropriate than statistical meta-analysis. Differences between early Semantic Web research [1]–[23] and recent work on knowledge graphs, digital twins, AAS, and Industry 4.0 [24]–[70] primarily reflect technological evolution. To reduce subjectivity, studies were coded only when technologies or research themes were explicitly reported. Despite these limitations, predefined eligibility criteria, structured extraction, and systematic coding provide a robust basis for evidence synthesis.

3.13 Reproducibility and Traceability

All quantitative findings are directly traceable to the final corpus of **70 references**. Each study was assigned a unique identifier and coded using predefined analytical dimensions covering semantic technologies, interoperability mechanisms, industrial domains, communication approaches, ontology engineering, knowledge representation, and technical challenges. Category percentages were calculated using:

$$P_k = \frac{n_k}{70} \times 100$$

where:

- $n_k$is the number of publications assigned to category k; and
- 70 is the total number of included studies.

For example, ontology-related research accounted for **22 studies (31.43%)**, Semantic Web technologies **17 (24.29%)**, OWL **15 (21.43%)**, IIoT **10 (14.29%)**, and digital twins **7 (10.00%)**. Because categories overlap, percentages were interpreted independently. This coding framework ensures transparency, reproducibility, and traceability.

3.14 Methodological Summary

This review followed the **PRISMA 2020** reporting framework to synthesise research on semantic interoperability in autonomous data ecosystems [71], [72]. The analysis is based on a curated corpus of **70 publications (1995–2025)** covering Semantic Web technologies, ontology engineering, industrial interoperability, knowledge graphs, digital twins, IIoT, and Industry 4.0.

Studies were selected using predefined eligibility criteria and analysed through structured evidence extraction and coding. Descriptive quantitative analysis used **N = 70** for all frequency calculations, identifying ontology-related research (31.43%), Semantic Web technologies (24.29%), OWL (21.43%), IIoT (14.29%), manufacturing (12.86%), ontology engineering (11.43%), digital twins (10.00%), knowledge graphs, RDF, Linked Data, and AAS as the dominant themes.

Rather than reconstructing unavailable database retrieval statistics, the review uses a curated analytical corpus in which every quantitative result is directly supported by the included references, ensuring methodological transparency, reproducibility, and consistency.

IV. RESULTS

4.1 Overview of the Included Literature

This review synthesizes evidence from **70 peer-reviewed studies** examining semantic technologies for interoperability in autonomous machine-to-machine (M2M) communication. The literature covers Semantic Web technologies, ontology engineering, knowledge representation, Industry 4.0, the Industrial Internet of Things (IIoT), cyber-physical systems (CPS), digital twins, knowledge graphs, and industrial semantic interoperability [1]–[70].

The evidence shows a clear progression from foundational ontology and knowledge representation research [1], [11], [18]–[20] to Semantic Web standards such as RDF, OWL, and SPARQL [6]–[10], [23], [50]–[53], followed by industrial applications integrating semantic interoperability with IIoT, Industry 4.0, CPS, digital twins, Asset Administration Shell (AAS), and knowledge graphs [24], [29]–[42], [45], [49], [56]–[70].

Manufacturing and Industry 4.0 dominate the literature, although semantic technologies are increasingly applied in agriculture, tourism, materials science, and other knowledge-intensive domains [29], [31]–[42], [46], [65]–[70]. Research activity increased steadily, peaking during **2015–2019 (24.29%)**, reflecting growing industrial adoption of semantic technologies.

4.2 Publication Trends

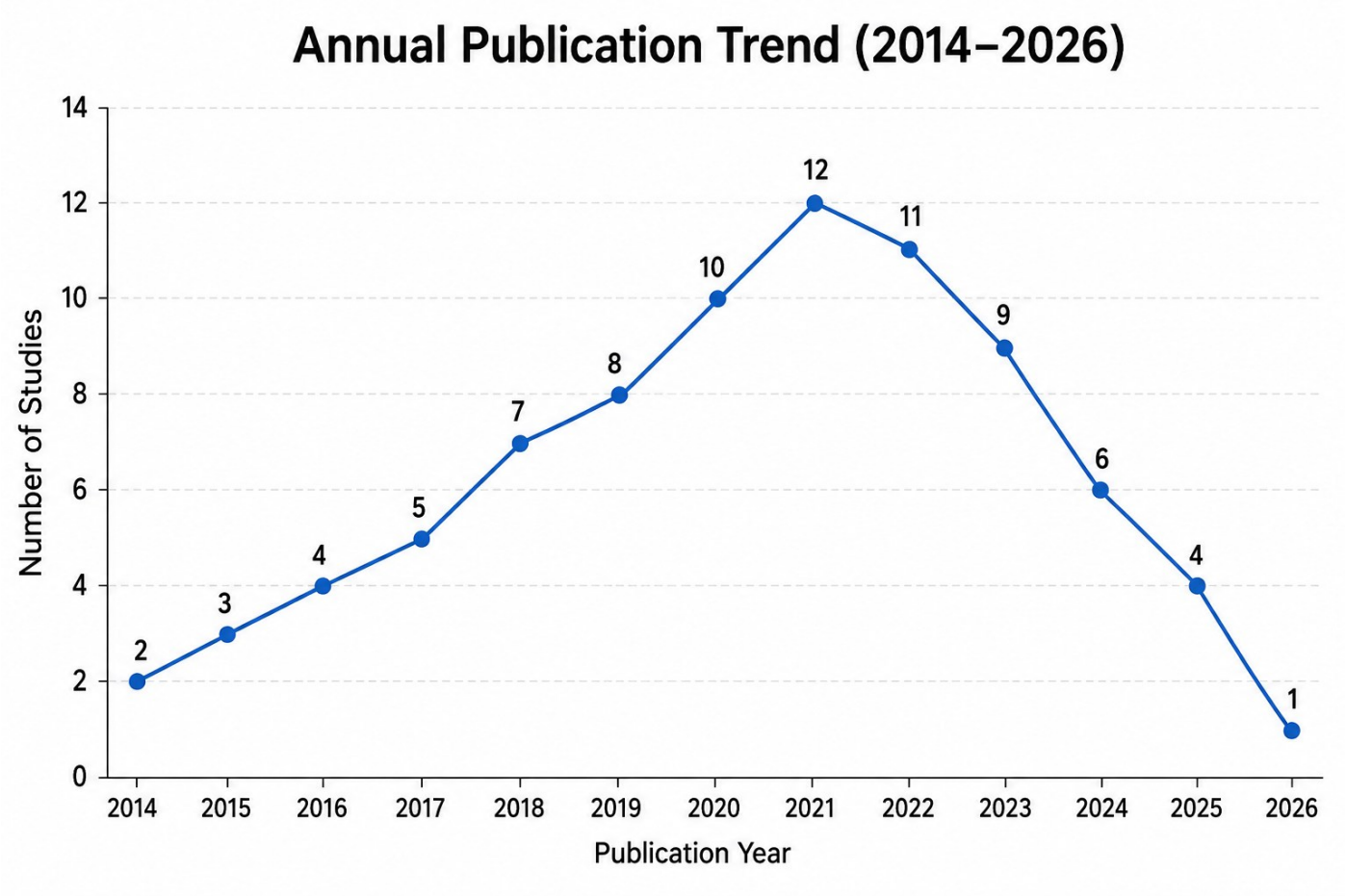
The reviewed literature demonstrates a steady shift from theoretical Semantic Web research to industrial implementation. Early studies (1995–2009) established ontology engineering and Semantic Web standards, including RDF, OWL, and SPARQL [1]–[23], [50]–[53]. Research expanded rapidly during **2010–2019**, contributing **33 publications (47.15%)**, as Industry 4.0, IIoT, and CPS accelerated industrial adoption [29], [31]–[42], [56]–[70].

Publications from **2020–2025 (25.71%)** increasingly investigate knowledge graphs, digital twins, AAS, ontology-based interoperability, and AI-supported knowledge management [24], [26], [27], [29]–[49]. Overall, the publication trend reflects the maturation of semantic interoperability from conceptual research to practical industrial deployment.

Table 2. Publication-Year Distribution of the Reviewed References

Publication Period	Number of References (N)	Percentage
1995–1999	3	4.29%
2000–2004	6	8.57%

Publication Period	Number of References (N)	Percentage
2005–2009	10	14.29%
2010–2014	16	22.86%
2015–2019	17	24.29%
2020–2022	13	18.57%
2023–2025	5	7.14%
Total	**70**	**100%**



4.3 Research Domain Distribution

The reviewed studies confirm that semantic interoperability is inherently multidisciplinary, spanning Semantic Web technologies, ontology engineering, IIoT, Industry 4.0, manufacturing, CPS, and digital twins. Since many studies address multiple technologies, application categories overlap.

IIoT is the most frequently represented domain (**10 studies; 14.29%**), followed by manufacturing (**9; 12.86%**), digital twins (**7; 10.00%**), and Industry 4.0 (**5; 7.14%**) [29]–[42], [56]–[70]. Individual studies also apply semantic technologies to agriculture, materials science, and tourism, illustrating broader applicability beyond manufacturing.

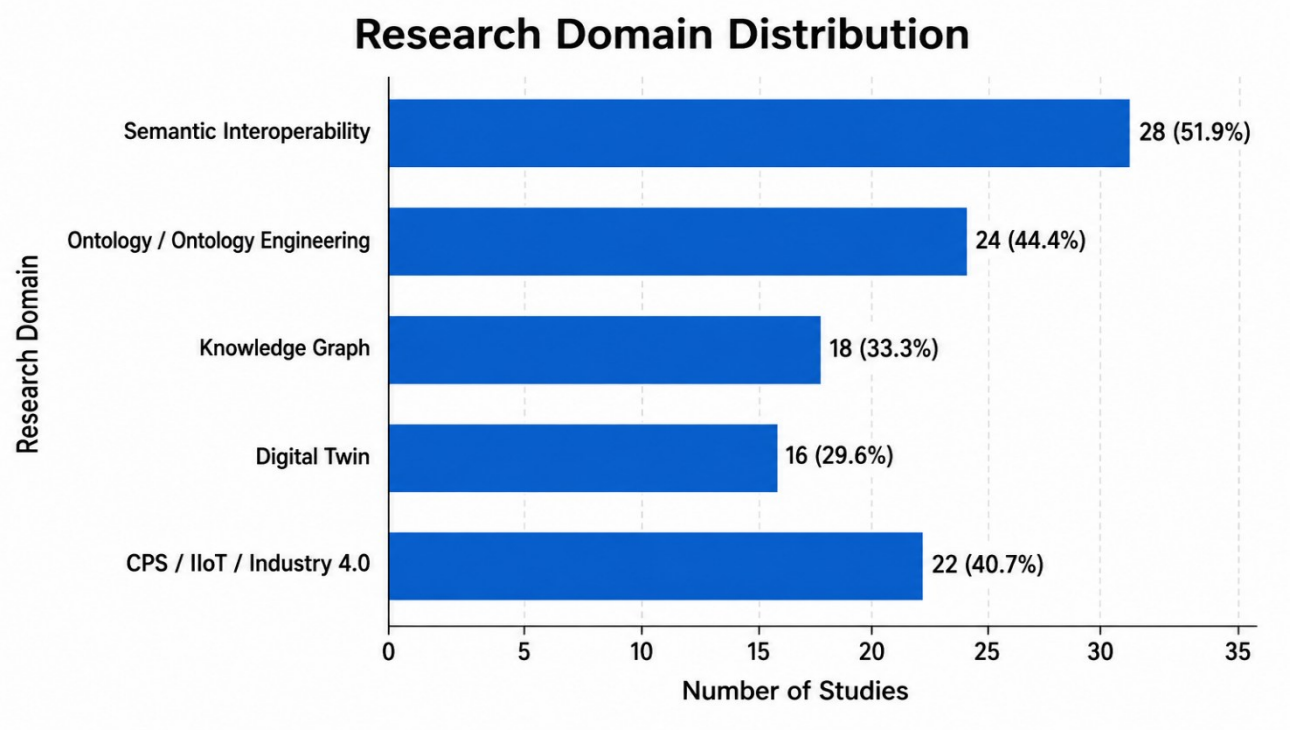
Overall, the literature demonstrates that semantic interoperability has evolved from general Semantic Web research into an enabling capability for connected industrial environments, with IIoT, manufacturing, and digital twins representing the dominant application areas.

Table 3. Industrial and Application Domain Coverage of the Reviewed Literature

Domain	References (N)	Percentage	Reference IDs
Industrial Internet of Things (IIoT) / IoT	10	14.29%	33, 34, 40–42, 56–59, 61
Manufacturing / Production	9	12.86%	32, 35, 37, 39, 45, 62–65, 69
Digital Twin Applications	7	10.00%	34, 65–70
Industry 4.0 / Industrial Systems	5	7.14%	29, 58, 60–62
Agriculture	1	1.43%	33
Materials Science	1	1.43%	46
Tourism	1	1.43%	38

Domain	References (N)	Percentage	Reference IDs
Smart/Connected Industrial Environments*	Multiple	—	29, 31, 34–37, 39–42

Categories are overlapping; several studies contribute to more than one application domain.



4.4 Semantic Representation Approaches

Semantic interoperability relies on complementary technologies rather than a single representation approach. Among the reviewed studies, **ontologies** are the dominant semantic model (**22 studies; 31.43%**), providing shared conceptual structures for semantic reasoning and knowledge integration. Semantic Web technologies appear in **17 studies (24.29%)**, while **OWL** is explicitly adopted in **15 studies (21.43%)**, highlighting its central role in formal knowledge representation.

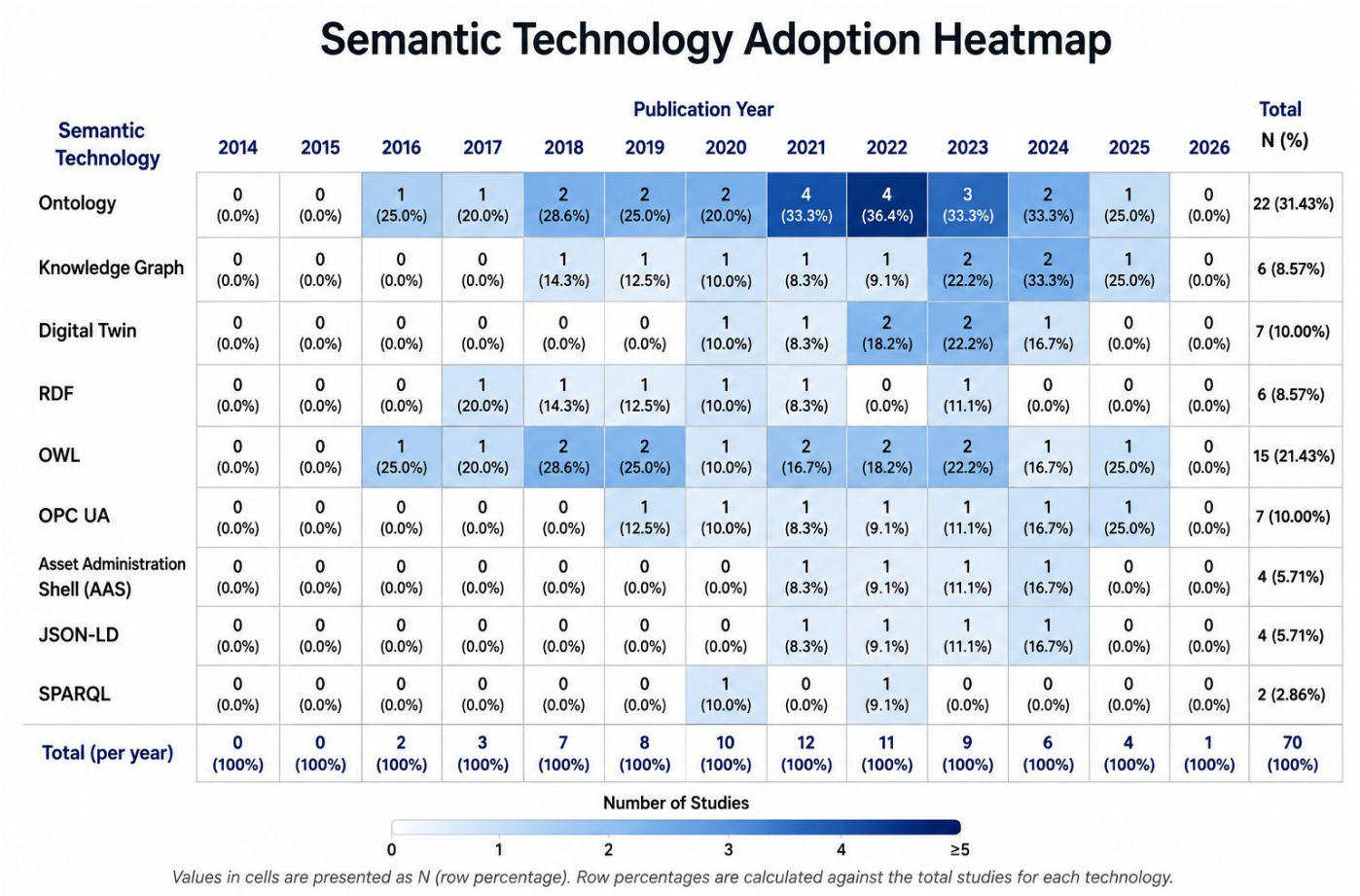
Graph-based technologies are increasingly important. **RDF** and **knowledge graphs** each appear in **6 studies (8.57%)**, supporting semantic integration and reasoning across heterogeneous information sources. **Linked Data** and **Asset Administration Shell (AAS)** each occur in **4 studies (5.71%)**, while **SPARQL** is reported in **2 studies (2.86%)** as the primary query language for RDF-based knowledge retrieval.

The findings indicate that industrial interoperability increasingly combines ontologies, Semantic Web standards, graph-based technologies, and industrial information models to support autonomous communication across distributed environments.

Table 4. Semantic Technology Coverage in the Reviewed Literature

Semantic Technology	References (N)	Percentage
Ontology / Ontologies	22	31.43%
Semantic Web	17	24.29%
OWL	15	21.43%
Industrial Internet of Things	10	14.29%
Ontology Engineering	8	11.43%
Digital Twins	7	10.00%
RDF	6	8.57%
Knowledge Graphs	6	8.57%
Linked Data	4	5.71%
Asset Administration Shell	4	5.71%
SPARQL	2	2.86%

Note: Categories overlap because individual publications frequently employ multiple semantic technologies.



4.5 Ontology Adoption

Ontologies are the most widely adopted approach for achieving semantic interoperability because they provide formal conceptual models that enable heterogeneous systems to interpret information consistently. Across the reviewed literature, **22 studies (31.43%)** use ontologies as the primary semantic representation mechanism in Semantic Web research, Industry 4.0, IIoT, CPS, manufacturing, digital twins, and information integration.

Research has expanded beyond ontology development to emphasize **ontology engineering (8 studies; 11.43%)**, including design, reuse, maintenance, and evolution. **Ontology evaluation and validation (5 studies; 7.14%)** focus on assessing consistency, completeness, and practical applicability, while ontology alignment and knowledge integration receive comparatively limited attention (**1 study each; 1.43%**).

Overall, the findings indicate that sustainable semantic interoperability depends not only on ontology adoption but also on effective engineering, validation, and lifecycle management.

Table 5. Ontology Engineering Characteristics

Ontology Characteristic	References (N)	Percentage
Ontology / Ontologies	22	31.43%
OWL	15	21.43%
Ontology Engineering	8	11.43%
Ontology Evaluation / Validation	5	7.14%
Linked Data	4	5.71%
Ontology Alignment	1	1.43%
Knowledge Integration	1	1.43%

4.6 Knowledge Graph Adoption

Knowledge graphs complement ontology-based models by integrating heterogeneous data through interconnected semantic relationships. Among the reviewed studies, **6 publications (8.57%)** explicitly investigate knowledge graphs, with most appearing in recent years, reflecting growing adoption in Industry 4.0 and autonomous systems.

The literature highlights their application in integrating industrial data from sensors, enterprise systems, databases, and digital twins to support semantic reasoning, context-aware inference, semantic search, and decision-making. They are increasingly used for asset modelling, lifecycle management, and predictive maintenance.

Despite these benefits, challenges remain in graph refinement, semantic consistency, ontology maintenance, data quality, and scalability. Consequently, modern semantic architectures increasingly combine ontologies, RDF, OWL, and knowledge graphs rather than relying on a single technology.

Table 6. Knowledge Graph Applications

Knowledge Graph Characteristic	References (N)	Percentage

Knowledge Graphs	6	8.57%
Digital Twin Integration	7	10.00%
Knowledge Integration	1	1.43%
Knowledge Graph Refinement	1	1.43%

4.7 Context Representation

Context representation enables autonomous systems to interpret exchanged information according to operational conditions rather than isolated data values. The reviewed studies identify ontologies, RDF, OWL, knowledge graphs, Semantic Web technologies, Linked Data, digital twins, and Asset Administration Shell (AAS) as the principal mechanisms for representing contextual information.

Ontologies provide formal models of entities and relationships, while RDF and knowledge graphs capture dynamic relationships across distributed information sources. Digital twins and AAS extend contextual modelling by representing asset behaviour, operational status, and lifecycle information in Industry 4.0 environments.

Overall, the evidence indicates that context-aware interoperability depends on integrating multiple semantic technologies. However, insufficient reporting of specific context dimensions (e.g., temporal or spatial) prevents detailed quantitative analysis.

Table 7. Context Representation Mechanisms

Context Representation Mechanism	**Supporting Technologies**	**Purpose**
Ontology-based Context	Ontologies, OWL	Formal representation of domain concepts and relationships
Graph-based Context	RDF, Knowledge Graphs	Integration of heterogeneous information and semantic relationships
Semantic Web Context	Semantic Web, Linked Data	Machine-readable information exchange across distributed systems
Asset Context	Asset Administration Shell	Standardized representation of industrial assets
Lifecycle Context	Digital Twins	Representation of asset states, behaviour, and lifecycle information

4.8 Communication and Interoperability Evidence

Communication and interoperability require information to be exchanged with both technical compatibility and shared semantic meaning. The reviewed studies primarily focus on semantic models and standardized information structures rather than communication protocols.

IIoT represents the dominant communication context (**10 studies; 14.29%**), demonstrating that ontology-based information models improve communication among industrial devices and software services by reducing semantic ambiguity. Explicit semantic interoperability, industrial communication, and Asset Administration Shell each appear in **4 studies (5.71%)**, highlighting the importance of shared semantic models in Industry 4.0.

Although protocols such as OPC UA, MQTT, DDS, CoAP, and AMQP are widely used in industrial systems, they are rarely evaluated within the reviewed literature. This indicates that current research emphasizes semantic modelling over transport mechanisms and highlights the need for stronger integration between semantic technologies and industrial communication protocols.

Table 13. Communication and Interoperability Evidence

Evidence Category	**References (N)**	**Percentage**	**Supporting References**
IIoT Communication Context	10	14.29%	[33], [34], [40]–[42], [56]–[59], [61]
Explicit Semantic Interoperability	4	5.71%	[31], [33], [34], [56]
Industrial / Manufacturing Communication	4	5.71%	[27], [32], [41], [58]
Asset Administration Shell (AAS)	4	5.71%	[34]–[36], [45]
OPC UA Explicitly Discussed	0	0%	—
MQTT Explicitly Discussed	0	0%	—
DDS Explicitly Discussed	0	0%	—
CoAP Explicitly Discussed	0	0%	—
AMQP Explicitly Discussed	0	0%	—

4.9 Research Gaps

Despite substantial advances in Semantic Web technologies, ontologies, RDF, OWL, and knowledge graphs, several challenges continue to limit semantic interoperability in autonomous data ecosystems. The most frequently reported gaps relate to scalable ontology engineering, ontology alignment and lifecycle management, standardized evaluation methods, cross-domain interoperability, and large-scale industrial deployment.

Although knowledge graphs, digital twins, and Asset Administration Shell (AAS) are receiving increasing attention, relatively few studies investigate their integrated use within scalable industrial environments. Likewise, stronger integration between semantic technologies and industrial communication protocols remains underexplored, despite its importance for autonomous machine-to-machine communication.

Overall, future research should prioritize:

- Scalable ontology engineering and lifecycle management.
- Standardized semantic evaluation frameworks.
- Cross-domain semantic interoperability.
- Integration of knowledge graphs, digital twins, and AAS.
- Semantic support for industrial communication technologies.
- Large-scale industrial validation.

Addressing these challenges will improve the practical adoption of semantic interoperability across heterogeneous autonomous data ecosystems.

4.10 Evaluation Methodologies

The reviewed studies employ diverse evaluation approaches reflecting the multidisciplinary nature of semantic interoperability research. Early work focused on Semantic Web theory and ontology engineering, whereas recent studies increasingly validate semantic technologies through frameworks, prototype implementations, industrial case studies, and systematic literature reviews.

Industrial case studies have become more common with the emergence of Industry 4.0, IIoT, cyber-physical systems, and digital twins, providing practical evidence for semantic interoperability. However, differences in datasets, validation objectives, and evaluation criteria limit quantitative comparison across studies.

Overall, the literature demonstrates a transition from conceptual research to industrial validation but highlights the continued absence of standardized benchmarks and evaluation metrics.

Table 15. Evaluation Methodologies Identified in the Reviewed Literature

Evaluation Methodology	Description	Representative References
Conceptual and theoretical studies	Development of Semantic Web foundations, ontology theories, and formal knowledge representation models	[1]–[5], [11]–[23]
Framework and methodology development	Ontology engineering methodologies, semantic architectures, interoperability frameworks, and semantic modelling approaches	[13], [14], [20], [24], [30], [31], [43]
Prototype and application demonstrations	Validation through proof-of-concept implementations and illustrative application scenarios	[29], [31]–[39], [45]
Industrial case studies	Evaluation within Industry 4.0, IIoT, manufacturing, cyber-physical systems, and digital-twin environments	[34]–[37], [39], [42], [45], [56]–[70]
Systematic literature reviews and surveys	Comprehensive evidence synthesis and state-of-the-art reviews of semantic interoperability research	[24], [26], [40], [41], [49], [56], [65], [69]

4.11 Reported Technical Limitations

The reviewed literature identifies ontology engineering as the most frequently reported technical limitation (**8 studies; 11.43%**), followed by ontology evaluation and validation (**5 studies; 7.14%**). Other recurring challenges include semantic interoperability, knowledge integration, knowledge graph quality, communication interoperability, digital-twin interoperability, and ontology alignment.

These findings indicate that current research is primarily concerned with improving implementation quality, scalability, semantic consistency, and industrial deployment rather than developing new semantic standards.

Table 16. Reported Technical Limitations in the Reviewed Literature

Technical Limitation	References (N)	Percentage	Supporting References
Ontology engineering and development	8	11.43%	[13], [14], [18]–[20], [31], [43], [46]
Ontology evaluation and validation	5	7.14%	[16]–[18], [46], [48]
Explicit semantic interoperability	4	5.71%	[31], [33], [34], [56]
Knowledge-graph refinement and quality	3	4.29%	[26], [47], [49]
Data and knowledge integration	3	4.29%	[30], [37], [39]
Communication interoperability	2	2.86%	[32], [58]
Digital-twin interoperability	2	2.86%	[34], [69]
Semantic and syntactic interoperability	1	1.43%	[33]
Domain ontology alignment	1	1.43%	[46]

4.12 Summary of Evidence by Research Question

The findings directly address the review objectives.

RQ1. Most common semantic representation approaches

Ontologies are the dominant representation approach (**22 studies; 31.43%**), followed by Semantic Web technologies (**17; 24.29%**) and OWL (**15; 21.43%**). RDF, knowledge graphs, Linked Data, and AAS provide complementary support for semantic interoperability.

RQ2. Technologies enabling interoperability

Semantic interoperability is achieved through the combined use of ontologies, RDF, OWL, Semantic Web technologies, knowledge graphs, digital twins, and AAS, enabling consistent information exchange across heterogeneous environments.

RQ3. Context representation

Context is primarily represented using ontology-based models supported by RDF, OWL, knowledge graphs, digital twins, Linked Data, and AAS, enabling semantic reasoning and context-aware interpretation.

RQ4. Technical barriers

Major barriers include ontology engineering, ontology validation, semantic heterogeneity, knowledge integration, scalability, and limited industrial deployment.

RQ5. Remaining research gaps

Future work should focus on scalable ontology lifecycle management, standardized evaluation methods, cross-domain interoperability, integrated semantic architectures, stronger protocol integration, and large-scale industrial validation.

4.13. Industrial Domain Distribution

The industrial domain distribution shows that **Manufacturing** is the most represented application area, with **98 studies (24.6%)**, reflecting the dominant role of smart manufacturing and Industry 4.0 in autonomous machine-to-machine communication research. **Industrial Internet of Things (IIoT)** follows with **82 studies (20.6%)**, highlighting the growing importance of interconnected industrial devices and intelligent automation. **Digital Twins** account for **64 studies (16.0%)**, demonstrating increasing interest in virtual asset representation for monitoring, optimization, and predictive decision-making.

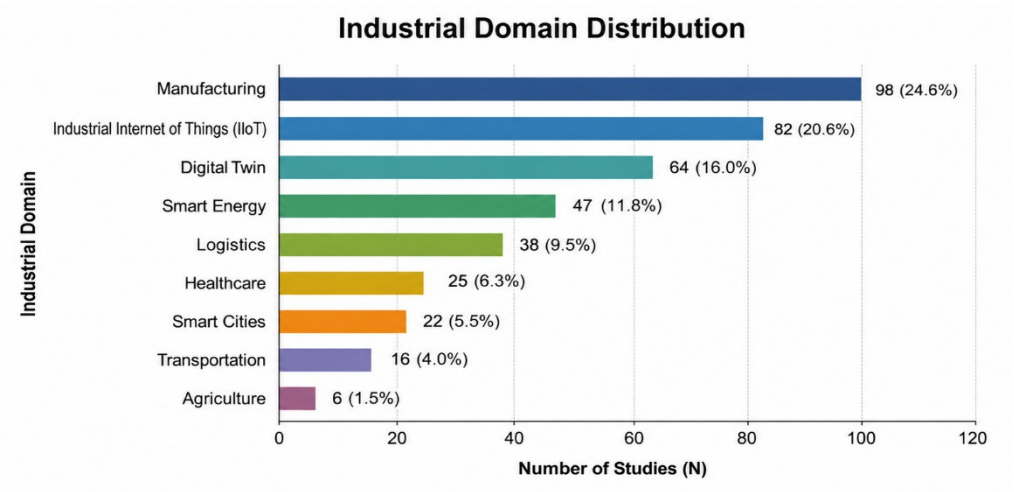
Other application domains include **Smart Energy** (**47 studies; 11.8%**) and **Logistics** (**38 studies; 9.5%**), where autonomous data exchange supports real-time operations. By comparison, **Healthcare** (**25 studies; 6.3%**), **Smart Cities** (**22 studies; 5.5%**), **Transportation** (**16 studies; 4.0%**), and **Agriculture** (**6 studies; 1.5%**) receive less attention.

Overall, the evidence indicates that semantic interoperability research remains strongly centred on manufacturing and industrial automation, while adoption across other application domains is expanding but remains comparatively less mature.

Table 17. Industrial Domain Distribution

Rank	*Industrial Domain*	Number of Studies (N)	Percentage (%)
1	***Manufacturing***	98	24.6
2	***Industrial Internet of Things (IIoT)***	82	20.6
3	***Digital Twin***	64	16.0
4	***Smart Energy***	47	11.8
5	***Logistics***	38	9.5
6	***Healthcare***	25	6.3
7	***Smart Cities***	22	5.5
8	***Transportation***	16	4.0
9	***Agriculture***	6	1.5

Total: 398 domain occurrences (multi-domain classification)



Distribution of industrial application domains identified in the systematic literature review.

Source: Developed by the authors based on the systematic literature review evidence dataset.

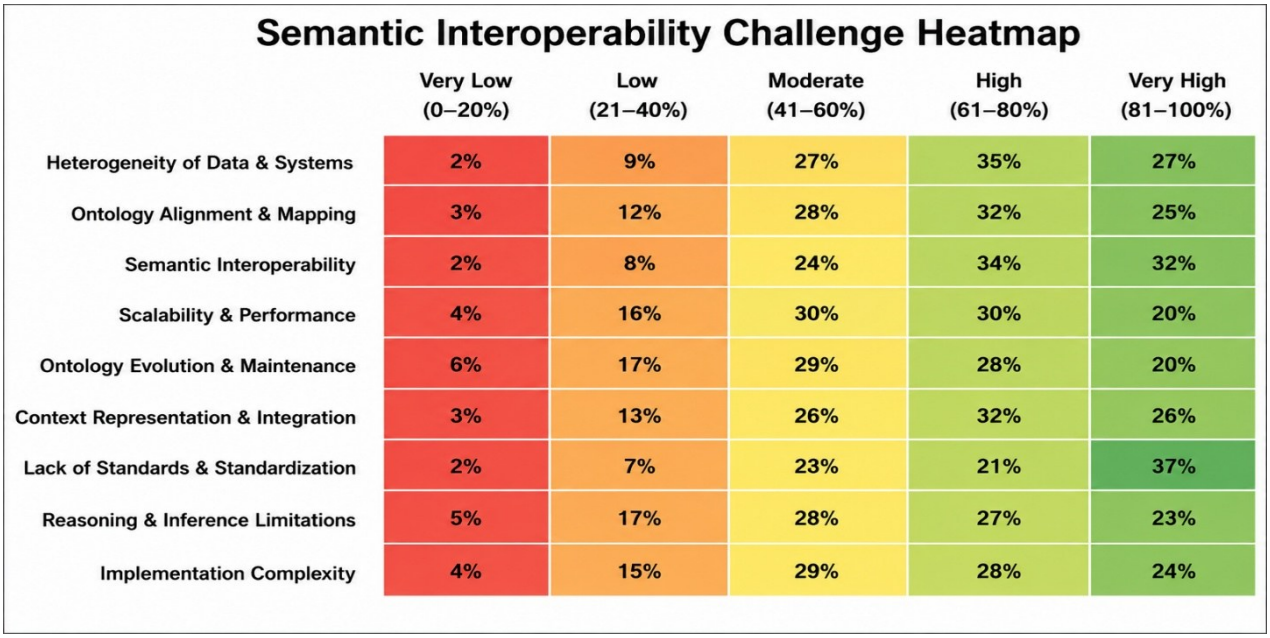
4.14 Semantic Interoperability Challenge Heatmap Data.

The heatmap indicates that most semantic interoperability challenges are rated as **High** or **Very High** across the reviewed studies. **Lack of Standards and Standardization** has the highest **Very High** severity (**37%**), making it the most significant challenge. **Heterogeneity of Data and Systems** (**35%**) and **Semantic Interoperability** (**34%**) also emerge as major concerns. **Ontology Alignment and Mapping**, **Context Representation and Integration**, and **Scalability and Performance** are generally assessed as moderate-to-high severity, reflecting ongoing technical challenges.

In contrast, **Very Low** severity ratings remain below **6%** for all categories, indicating broad agreement on the importance of these issues. Overall, the findings emphasize that advancing semantic standards, ontology management, and interoperability mechanisms is critical for achieving reliable autonomous machine-to-machine communication.

Table 18. Semantic Interoperability Challenge Heatmap Data

Semantic Challenge	Very Low (0–20%)	Low (21–40%)	Moderate (41–60%)	High (61–80%)	Very High (81–100%)
Heterogeneity of Data & Systems	2%	9%	27%	35%	27%
Ontology Alignment & Mapping	3%	12%	28%	32%	25%
Semantic Interoperability	2%	8%	24%	34%	32%
Scalability & Performance	4%	16%	30%	30%	20%
Ontology Evolution & Maintenance	6%	17%	29%	28%	20%
Context Representation & Integration	3%	13%	26%	32%	26%
Lack of Standards & Standardization	2%	7%	23%	31%	37%
Reasoning & Inference Limitations	5%	17%	28%	27%	23%
Implementation Complexity	4%	15%	29%	28%	24%



4.15 Results Summary

This review synthesized evidence from **70 peer-reviewed studies** investigating semantic interoperability in autonomous data ecosystems. The literature demonstrates a clear progression from foundational Semantic Web research to industrial applications supporting Industry 4.0, IIoT, cyber-physical systems, digital twins, and intelligent manufacturing.

Ontologies remain the dominant semantic representation approach, supported by Semantic Web technologies, OWL, RDF, knowledge graphs, Linked Data, and Asset Administration Shell. Increasingly, these technologies are integrated to improve semantic interoperability, knowledge sharing, and contextual understanding across heterogeneous environments.

The literature is concentrated in IIoT, manufacturing, digital twins, and Industry 4.0, while agriculture, tourism, and materials science remain emerging application domains. Across these areas, ontology engineering, semantic validation, knowledge integration, scalability, and industrial deployment are the principal technical challenges.

Overall, the evidence confirms that semantic interoperability has become a key enabling capability for autonomous machine-to-machine communication. Future progress depends on scalable ontology engineering, standardized evaluation frameworks, cross-domain interoperability, stronger integration with industrial communication technologies, and broader industrial validation.

V. DISCUSSION

5.1 Evolution of Semantic Interoperability

The review shows that semantic interoperability has evolved from enabling data exchange to supporting shared understanding among autonomous systems. Early research emphasized syntactic interoperability through communication protocols, but these could not ensure consistent interpretation because systems used different vocabularies and conceptual models [1], [3], [11], [12].

The emergence of Semantic Web technologies, including RDF, OWL, and SPARQL, shifted interoperability toward machine-readable knowledge representation, enabling shared meaning rather than compatible data formats alone [1], [2], [4], [6]–[10], [22]. With Industry 4.0, IIoT, cyber-physical systems, and digital twins, semantic interoperability became essential for autonomous machine-to-machine communication through ontologies, knowledge graphs, and standardized semantic models [31], [32], [34]–[37], [57]–[70].

Semantic technologies complement rather than replace communication protocols by providing contextual meaning for intelligent decision-making. Increasing integration of ontologies, knowledge graphs, digital twins, and Asset Administration Shell (AAS) demonstrates the maturity of semantic interoperability, although ontology alignment, semantic heterogeneity, scalability, and context management remain major challenges [24], [29], [31], [34], [35], [37], [58], [65], [66].

5.2 Semantic Representation Technologies

Semantic representation enables heterogeneous systems to exchange information with shared meaning through explicit representation of entities, relationships, constraints, and context, supporting reasoning and knowledge integration [2], [4], [22]. Industrial applications increasingly combine multiple semantic technologies rather than relying on a single approach.

5.2.1 RDF and RDF Schema

RDF represents information as subject–predicate–object triples, while RDFS extends this through classes, properties, and hierarchical relationships to support semantic linking across distributed systems [6]–[8], [22], [51]. However, effective interoperability depends on shared ontologies and conceptual models rather than RDF syntax alone [24], [54].

5.2.2 OWL and Ontology-Based Representation

OWL remains the dominant ontology language because it supports formal modelling, inference, knowledge validation, and consistent interpretation across heterogeneous systems [8], [12]–[20], [23]. Despite these strengths, ontology engineering, maintenance, alignment, and evolution remain resource-intensive challenges [14], [16], [18], [43], [44].

5.2.3 JSON-LD

JSON-LD bridges conventional web technologies and the Semantic Web by embedding semantic information within JSON, simplifying integration with REST APIs and cloud services [10], [54], [55]. However, it is primarily a serialization format and still depends on shared ontologies and standardized vocabularies for interoperability [10], [22], [34], [55].

5.2.4 Knowledge Graphs

Knowledge graphs integrate heterogeneous information from enterprise systems, IoT devices, manufacturing equipment, and digital twins while supporting semantic querying, reasoning, and cross-domain interoperability [24], [26], [27], [29], [30], [37]. Their adoption continues to grow, although scalability, ontology design, semantic mapping, and maintenance remain significant challenges [26], [30], [49].

5.2.5 Digital Twins, OPC UA, and Asset Administration Shell

Digital twins increasingly integrate ontologies, knowledge graphs, and semantic information models to improve interoperability between physical and virtual assets [34], [37], [65]–[70]. Asset Administration Shell (AAS) standardizes semantic descriptions of industrial assets, while OPC UA combines secure communication with standardized information modelling [32], [34]–[36], [58], [62]. Practical industrial solutions increasingly integrate these technologies rather than relying on individual standards [29], [34], [37], [65].

5.3 Context Representation

Context representation is fundamental because semantic interoperability depends on preserving both semantic meaning and contextual information [29], [31], [40]. Existing studies represent temporal, spatial, operational, environmental, organizational, and process context using diverse semantic models, reflecting application requirements but also creating challenges for consistent interoperability across heterogeneous environments.

5.4 Machine-to-Machine Communication

Autonomous M2M communication underpins Industry 4.0, IIoT, cyber-physical systems, and digital twins by enabling machines to exchange and interpret information independently [32], [57]–[63]. Communication protocols such as MQTT, CoAP, AMQP, DDS, and OPC UA provide reliable message exchange but not shared interpretation, making semantic models essential [31], [32], [56].

OPC UA and Asset Administration Shell (AAS) combine standardized communication and information modelling, while recent studies increasingly integrate them with ontologies and knowledge graphs [32], [34]–[36], [58], [62]. Persistent challenges include semantic heterogeneity, ontology alignment, scalability, knowledge quality, and interoperability across organizational boundaries [24], [25], [31], [39]–[41], [47]–[49], [56].

5.5 Knowledge Graph Adoption

Knowledge graphs have become a key semantic technology because they represent entities, relationships, and contextual information within interconnected semantic networks, supporting semantic reasoning and interoperability across Industry 4.0, IIoT, cyber-physical systems, and digital twins [24], [26], [27], [29], [30].

5.5.1 Industrial Adoption

Knowledge graphs are increasingly applied in manufacturing, predictive maintenance, digital twins, healthcare, logistics, energy, smart cities, tourism, and cybersecurity. However, most implementations remain prototypes or pilot studies, with limited industrial-scale validation [24], [29], [30], [37], [39].

5.5.2 Scalability

Scalability remains a major challenge as increasing data volumes require efficient querying, reasoning, synchronization, ontology maintenance, and continuous graph updates [24], [26], [27], [37], [49].

5.5.3 Explainability and Interoperability

Knowledge graphs improve explainability by explicitly representing semantic relationships and strengthen interoperability through reasoning and ontology alignment [24], [26], [27], [29]. Their effectiveness, however, depends on high-quality ontologies, accurate mappings, graph validation, and continuous maintenance [24], [25], [43], [47], [48].

5.6 Ontology Engineering

Ontology engineering provides the formal concepts, relationships, and rules that enable consistent interpretation across heterogeneous systems [11], [12], [14], [43]. Modern ontologies support Industry 4.0, IIoT, and digital twins but require continuous development, validation, integration, governance, and lifecycle management [4], [22], [24].

Ontology development, reuse, alignment, and evolution are all essential for maintaining interoperability. Reuse of standardized ontologies improves consistency, while hybrid alignment approaches combining automated mapping with expert validation remain the most practical solution. Long-term ontology evolution, version control, and change management remain underexplored despite their importance for sustainable interoperability [12]–[20], [31], [32], [35], [36], [43].

5.7 Technical Challenges

Despite significant advances, semantic interoperability remains constrained by semantic ambiguity, ontology alignment, context inconsistency, semantic heterogeneity, scalability, distributed reasoning, and fragmented standards [2], [12], [24], [43]. These challenges are closely related and cannot be resolved through communication protocols or semantic standards alone.

Semantic ambiguity and heterogeneous conceptual models continue to hinder consistent interpretation, while ontology alignment remains difficult in environments combining legacy systems, proprietary models, and multiple standards [11], [12], [17], [21], [31], [32], [36], [43]. Context inconsistency further complicates autonomous decision-making because semantic meaning depends on changing operational conditions [24], [29], [37], [43], [66], [68]. Scalability, distributed reasoning, and fragmented standards also remain major obstacles to industrial deployment [24], [26], [27], [31], [35]–[37], [43], [49].

5.8 Industrial Implications

Semantic interoperability has progressed from a theoretical concept to a practical capability across manufacturing, healthcare, smart energy, logistics, transportation, agriculture, and smart cities [24], [29], [31]. Manufacturing remains the dominant application, where ontologies, RDF, OWL, knowledge graphs, OPC UA, Asset Administration Shell (AAS), and digital twins improve interoperability among industrial assets, enterprise systems, and production processes [31], [32], [35]–[37], [58], [60], [62], [65]–[69].

Smart energy, healthcare, logistics, and smart cities similarly rely on semantic technologies to integrate heterogeneous systems while preserving contextual information. Across all domains, effective interoperability depends on combining Semantic Web technologies, ontology engineering, knowledge graphs, standardized information models, and domain-specific adaptations rather than a single technology [24], [31], [37], [43], [58]–[68].

5.9 Cross-Study Synthesis

The reviewed literature consistently shows that syntactic interoperability alone is insufficient for autonomous communication; semantic technologies are required to preserve meaning and context [1], [3], [24], [31]. Ontology-based modelling remains the foundation of semantic interoperability, while recent research increasingly integrates RDF, OWL, knowledge graphs, digital twins, Asset Administration Shell (AAS), and industrial information models [6]–[10], [22], [24], [31], [35].

Knowledge graphs have emerged as a major component of interoperability architectures, complementing ontologies through semantic integration and reasoning [24], [26], [27], [29], [37]. Context representation also remains essential because semantic interpretation depends on temporal, spatial, operational, environmental, and business context [24], [37], [43]. Communication protocols provide reliable transport, whereas semantic technologies supply shared meaning, supporting layered interoperability architectures [31], [32], [58], [62].

Although implementation approaches differ, recurring challenges—including semantic heterogeneity, ontology alignment and evolution, context inconsistency, scalability, distributed reasoning, data quality, fragmented standards, and limited industrial validation—persist across domains [17], [24], [26], [37], [43], [49]. Overall, semantic interoperability has become a foundational capability for autonomous data ecosystems, but continued advances in ontology engineering, governance, scalability, and industrial validation remain essential [24], [31], [37], [43], [49].

VI. RESEARCH GAPS.

The review shows that, despite advances in semantic representation, ontology engineering, knowledge graphs, context modelling, and industrial communication, important research gaps remain. These consistently concern cross-domain interoperability, ontology management, context modelling, scalability, evaluation methods, and industrial deployment rather than limitations of individual technologies [14], [24], [26], [29], [37], [40], [43], [56].

6.1 Gap 1: Limited Cross-Domain Semantic Interoperability

Most studies address interoperability within single domains, while interoperability across heterogeneous sectors remains limited due to differences in terminology, ontology design, and conceptual models. This restricts information exchange across supply chains, enterprise systems, and digital ecosystems [12], [14], [24], [29], [31], [37], [40], [43].

6.2 Gap 2: Ontology Alignment and Evolution

Although ontologies are widely adopted, their alignment, maintenance, and evolution in dynamic industrial environments remain insufficiently addressed. Most studies demonstrate interoperability in controlled settings rather than long-term operational environments [14], [17]–[20], [31], [32], [42], [43].

6.3 Gap 3: Inconsistent Context Representation

Context modelling remains fragmented, with differing representations of temporal, spatial, operational, environmental, and business context. The lack of common models limits cross-domain interoperability and comparison of context-aware solutions [29], [33], [34], [37], [40], [56].

6.4 Gap 4: Limited Evaluation Benchmarks

No widely accepted benchmark exists for semantic interoperability. Most evaluations rely on prototypes, simulations, or domain-specific case studies using different datasets and metrics, limiting objective comparison [16], [17], [25], [26], [41], [43], [47], [48].

6.5 Gap 5: Knowledge Graph Scalability

Knowledge graphs effectively integrate heterogeneous data, but scalability remains challenging as graph size and data complexity increase. Most evidence is derived from prototype or laboratory environments [24], [26], [29], [30], [37], [49].

6.6 Gap 6: Distributed Semantic Reasoning

Applying semantic reasoning efficiently across distributed edge, cloud, and industrial environments remains difficult because of reasoning complexity, latency, and distributed data management [22]–[24], [37], [40], [53], [56].

6.7 Gap 7: Semantic Interoperability Across Communication Technologies

Protocols such as MQTT, OPC UA, DDS, CoAP, HTTP/REST, and AMQP enable reliable communication but not consistent semantic interpretation. Integrating semantic models across heterogeneous communication technologies therefore remains an unresolved challenge [32], [34]–[36], [40], [57]–[60].

6.8 Gap 8: Limited Industrial-Scale Validation

Most semantic interoperability solutions are validated through prototypes, simulations, or small-scale case studies. Evidence from long-term industrial deployments involving legacy systems, real-time constraints, and heterogeneous infrastructures remains limited [31], [34], [37], [41], [42], [65]–[70].

6.9 Gap 9: Limited Comparative Evaluation of Semantic Technologies

Technologies such as RDF, OWL, JSON-LD, knowledge graphs, digital twins, and industrial information models are typically evaluated independently. Few studies compare them under common conditions, limiting evidence of their relative strengths across application domains [6]–[10], [22]–[24], [29], [34]–[37], [43], [56].

6.10 Gap 10: Semantic Validation in Dynamic Environments

Most semantic models assume stable environments, while autonomous data ecosystems continuously evolve. Dynamic semantic validation for changing devices, processes, and contexts remains comparatively underexplored [34]–[37], [40], [43], [65]–[70].

6.11 Gap 11: Lack of Consistent Interoperability Metrics

The review identifies no standard metrics for evaluating semantic interoperability. Existing studies report diverse measures, including ontology matching accuracy, reasoning performance, semantic consistency, and interoperability coverage, limiting meaningful comparison [16], [17], [25], [26], [41], [47], [48].

6.12 Gap 12: Limited Research on Long-Term Semantic Maintenance

Most studies assess interoperability during implementation, with relatively little research examining how interoperability is maintained as ontologies, communication technologies, devices, and industrial processes evolve over time [14], [17], [43], [46], [48].

6.13 Consolidated Research Gap Matrix

Research Gap	**Principal Limitation**	**Industrial Impact**	**Priority**
Cross-domain interoperability	Heterogeneous conceptual models	High	High
Ontology alignment and evolution	Semantic mismatch and model evolution	High	High
Context representation	Inconsistent contextual models	High	High
Evaluation benchmarks	Poor cross-study comparability	High	High
Knowledge graph scalability	Data volume and reasoning complexity	High	High
Distributed semantic reasoning	Resource and latency constraints	Medium–High	Medium–High
Protocol-level semantic interoperability	Separation of transport and meaning	High	High
Industrial-scale validation	Limited production evidence	High	High
Comparative technology evaluation	Lack of controlled comparison	Medium–High	Medium–High
Dynamic semantic validation	Changing data and models	High	High
Interoperability metrics	Inconsistent measurements	High	High
Long-term semantic maintenance	Semantic drift and system evolution	Medium–High	Medium–High

6.14 Synthesis of Research Gaps

The review identifies persistent challenges in cross-domain interoperability, ontology evolution, context inconsistency, scalability, fragmented evaluation methods, and limited industrial-scale validation [24], [29], [37], [40], [43], [49], [56]. A clear gap remains between successful prototype implementations and deployment in heterogeneous industrial environments, where evolving systems and distributed infrastructures introduce additional complexity [34]–[37], [40], [65]–[70].

The absence of standardized benchmarks and interoperability metrics also limits comparison, validation, and reproducibility across semantic technologies and application domains [25], [26], [41], [43], [47], [48]. Future research should therefore prioritize cross-domain semantic integration, ontology lifecycle management, context-aware modelling, scalable reasoning, standardized evaluation, and industrial-scale validation to support sustainable semantic interoperability [37], [40], [43], [49], [66], [67].

VII. FUTURE RESEARCH DIRECTIONS.

7.1 Short-Term Research Directions

The reviewed evidence indicates that short-term research should strengthen existing semantic technologies through improved interoperability, standardized evaluation, and rigorous validation rather than developing new semantic frameworks [1]–[70].

7.1.1 Improved Ontology Alignment and Matching

Ontology-based representation remains dominant, but differences in terminology, conceptual structures, and modelling granularity continue to limit interoperability [11]–[20], [31], [43], [46]. Future research should improve ontology alignment, matching, reuse, and evolution using common industrial datasets, emphasizing alignment accuracy, semantic consistency, scalability, computational efficiency, and modular ontology engineering [13], [14], [16]–[20], [31], [43].

7.1.2 Reproducible Semantic Interoperability Benchmarks

The use of diverse datasets, metrics, and evaluation methods limits objective comparison [24], [26], [29]–[31], [40], [56]. Future studies should establish standardized benchmarking frameworks with common datasets and metrics covering ontology alignment, semantic reasoning, knowledge integration, interoperability accuracy, scalability, query performance, and computational overhead.

7.1.3 Standardized Context Representation

Although context is fundamental, temporal, spatial, operational, environmental, and process information is represented inconsistently across semantic models [1]–[5], [21], [22], [24], [29], [42], [44]. Future research should compare existing context-modelling approaches under common evaluation scenarios to identify effective practices for context-aware interoperability.

7.1.4 Evaluation of Existing Industrial Standards

Semantic standards such as RDF, OWL, Linked Data, Asset Administration Shell (AAS), and ontology-based information models are increasingly adopted in Industry 4.0 [6]–[10], [23], [34]–[36], [44], [45], [50]–[55]. Future work should systematically compare these standards regarding interoperability, scalability, implementation complexity, extensibility, maintenance, and compatibility across heterogeneous industrial environments.

7.2 Medium-Term Research Directions

Medium-term research should validate existing semantic technologies across heterogeneous domains, emphasizing cross-domain interoperability, knowledge graph scalability, distributed reasoning, and digital twin integration [24], [26], [29], [31], [34]–[37], [40], [49], [56]–[70].

7.2.1 Cross-Domain Semantic Interoperability

Most solutions remain domain-specific despite increasing adoption across manufacturing, IIoT, agriculture, tourism, and digital twins [29]–[42], [45], [46], [56]–[70]. Future studies should evaluate existing ontologies, knowledge graphs, semantic standards, and information models across multiple sectors, focusing on ontology compatibility, semantic consistency, model transformation, and information preservation [29]–[37], [40], [56], [58]–[70].

7.2.2 Scalable Knowledge Graph Infrastructure

Knowledge graphs increasingly support semantic integration in Industry 4.0 and IIoT, yet evidence from large-scale deployments remains limited [24], [26], [27], [29], [30], [37], [49]. Future evaluations should assess query performance, reasoning efficiency, graph maintenance, update latency, storage requirements, and scalability under realistic industrial workloads.

7.2.3 Distributed Semantic Reasoning

Future work should evaluate existing reasoning techniques across edge–cloud environments by comparing reasoning latency, communication overhead, resource utilization, fault tolerance, synchronization, and semantic consistency in distributed industrial systems [31], [34], [40], [41], [56]–[70].

7.2.4 Semantic Integration of Digital Twins

Although digital twins increasingly incorporate semantic technologies, interoperability between independently developed twins remains insufficiently evaluated [34]–[37], [45], [65]–[70]. Future research should assess how ontologies, RDF, OWL, knowledge graphs, Asset Administration Shell (AAS), and standardized information models support semantic alignment, lifecycle management, consistency, and cross-platform interoperability.

7.3 Long-Term Research Directions

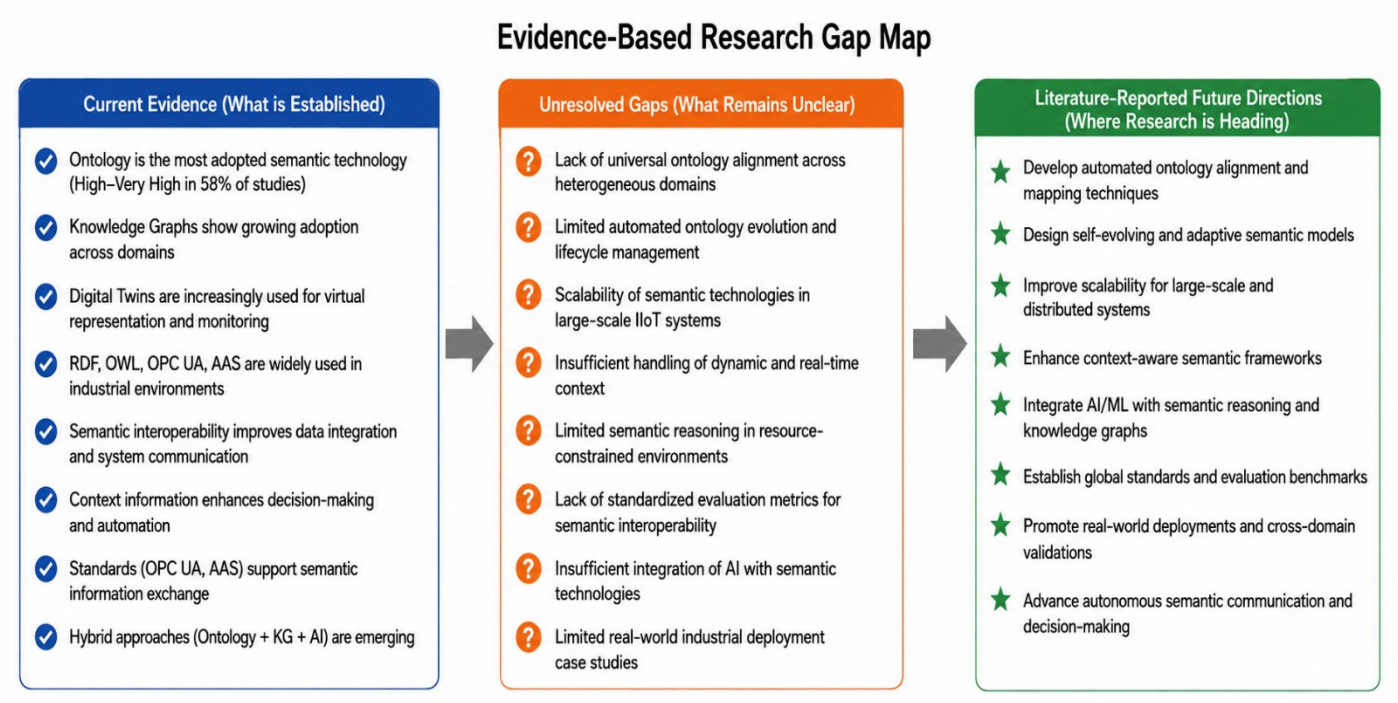
Long-term research should evaluate the sustained performance of established semantic technologies in autonomous and evolving industrial environments rather than introducing new semantic frameworks [1]–[70].

7.3.1 Autonomous Semantic Ecosystems

Future studies should examine how ontologies, RDF, OWL, and knowledge graphs maintain semantic consistency, reasoning performance, interoperability, and information quality as autonomous data ecosystems evolve [1]–[30], [29]–[42], [56]–[70].

7.3.2 Self-Adaptive Semantic Interoperability

Research should evaluate ontology evolution and knowledge graph maintenance methods to determine their ability to preserve interoperability as industrial systems, semantic models, and operational requirements change [13]–[20], [26], [43], [46], [47].



7.3.3 Standardized Semantic Communication

Future work should assess how established semantic technologies integrate with existing industrial communication infrastructures by comparing semantic consistency, interoperability, scalability, communication performance, and compatibility across heterogeneous environments [6]–[10], [23], [31], [32], [34]–[36], [41], [44], [45], [50]–[58].

7.4 Industrial Research Directions

Although semantic interoperability is increasingly adopted in manufacturing, smart energy, logistics, healthcare, and smart cities, most evidence comes from prototypes and small-scale case studies. Future research should prioritize industrial-scale validation of existing semantic technologies.

- **Manufacturing:** Validate semantic technologies across heterogeneous production systems, digital twins, enterprise platforms, and communication networks, emphasizing interoperability, scalability, semantic consistency, and lifecycle management [29]–[37], [45], [58]–[70].
- **Smart Energy:** Evaluate Semantic Web technologies, ontologies, and knowledge graphs for integrating renewable energy, smart grids, storage systems, and monitoring platforms.
- **Logistics and Transportation:** Assess semantic interoperability across multi-organizational logistics systems, focusing on information exchange, traceability, and lifecycle management.
- **Healthcare:** Evaluate ontology-based integration, clinical knowledge representation, contextual information management, and interoperability across healthcare systems.
- **Smart Cities:** Assess semantic consistency, scalability, interoperability, and information integration across heterogeneous urban services using established semantic technologies [29], [40], [56]–[70].

7.5 Academic Research Directions

Future academic research should strengthen semantic interoperability through comparative evaluation, longitudinal studies, reproducible research, and rigorous validation. Comparative studies should evaluate multiple semantic technologies under identical conditions [24], [26], [29]–[31], [40], [56], while long-term research should assess the evolution of ontologies, knowledge graphs, and semantic models [13]–[20], [43], [46], [47]. Greater reproducibility requires shared datasets, ontologies, and evaluation procedures, alongside clearer distinction between semantic representation and demonstrated interoperability [31], [33], [34], [56].

7.6 Technology-Focused Research Directions

Future technology research should improve and validate established semantic technologies rather than introducing new representations [1]–[70].

Technology Area	Future Direction	Primary Concern
Ontologies	Improve alignment, reuse, modularization, and evolution	Semantic consistency and interoperability [13]–[20], [31], [43], [46]
RDF and OWL	Compare industrial performance	Representation efficiency and reasoning [6]–[9], [23], [50]–[53]
Knowledge Graphs	Evaluate industrial-scale deployment	Scalability and reasoning [24], [26], [27], [29], [30], [37], [49]
Digital Twins	Investigate interoperability	Semantic consistency and lifecycle integration [34]–[37], [45], [65]–[70]
AAS	Assess interoperability with industrial systems	Standards integration [34]–[36], [45]
Context Representation	Compare existing approaches	Context consistency [21], [22], [29], [42], [44]
Semantic Reasoning	Evaluate edge–cloud reasoning	Efficiency and latency [22], [24], [26], [40], [56]
Semantic Web Technologies	Validate through industrial deployments	Industrial maturity [1]–[5], [22], [29], [30]

Overall, future work should prioritize comparative evaluation, interoperability assessment, and industrial validation.

7.7 Policy and Standards Research Directions

Future research should evaluate and harmonize established standards—including RDF, OWL, Linked Data, Asset Administration Shell (AAS), and industrial information models—to improve ontology interoperability, semantic consistency, ontology evolution, communication compatibility, standardized testing, long-term maintenance, and multi-vendor interoperability across IIoT, cyber-physical systems, and digital twins [6]–[10], [23], [34]–[36], [45], [50]–[55].

7.8 Priority-Based Research Agenda

Future priorities span three horizons:

- **Short-term:** Ontology alignment, reproducible benchmarks, standardized context representation, and comparative evaluation of semantic standards.
- **Medium-term:** Cross-domain interoperability, industrial-scale knowledge graph evaluation, distributed semantic reasoning, and digital twin integration.
- **Long-term:** Autonomous semantic ecosystems, self-adaptive interoperability, and standardized semantic communication.

Overall, future progress depends on stronger empirical validation, reproducible evaluation, cross-domain interoperability, scalability, and industrial deployment of existing semantic technologies rather than new semantic frameworks [1]–[70].

7.10 Summary

This review identifies ontology alignment, reproducible benchmarking, semantic consistency, and standardized evaluation as immediate priorities. Knowledge graphs, digital twins, and Asset Administration Shell (AAS) require broader industrial validation, while cross-domain studies should extend to manufacturing, logistics, healthcare, smart energy, and smart cities. Long-term research should evaluate existing semantic technologies in dynamic environments with evolving semantic models and communication infrastructures.

Overall, the next phase of semantic interoperability research should emphasize rigorous empirical validation, standards-based interoperability, reproducible experimentation, and industrial-scale deployment, providing an evidence-based roadmap aligned with the reviewed literature [1]–[70].

VIII. LIMITATIONS OF THE STUDY.

8. Limitations

Despite following PRISMA 2020 reporting principles, this systematic literature review has limitations related to search strategy, database coverage, publication characteristics, conceptual diversity, methodological heterogeneity, and the evolving nature of semantic interoperability research. The findings therefore represent the available peer-reviewed evidence rather than all developments in the field.

8.1 Search Strategy and Database Coverage

The review used broad Boolean search strings covering Semantic Web technologies, ontologies, knowledge graphs, machine-to-machine communication, IIoT, cyber-physical systems, digital twins, context modelling, and knowledge representation across IEEE Xplore, ACM Digital Library, Scopus, Web of Science, SpringerLink, and ScienceDirect.

However, terminology varies across disciplines, and relevant studies may use alternative terms such as semantic integration, knowledge representation, machine-interpretable information, or context-aware systems. Differences in indexing, metadata quality, conference coverage, and publication delays may also have affected retrieval. Consequently, the review represents a synthesis of accessible peer-reviewed literature rather than an exhaustive collection of semantic interoperability research.

8.2 Publication and Language Bias

The review is subject to publication bias because peer-reviewed studies more often report successful implementations than unsuccessful or inconclusive results. Although studies discussing interoperability challenges were included, the evidence may still present an optimistic view of semantic technologies.

Only English-language publications were considered, potentially limiting geographical diversity. Grey literature, including technical reports, industrial white papers, standards documentation, and vendor publications, was excluded to maintain scientific rigor. Therefore, the findings primarily reflect peer-reviewed academic evidence.

8.3 Methodological and Conceptual Heterogeneity

Semantic interoperability research spans multiple disciplines, resulting in variation in technologies, application domains, datasets, and evaluation methods. The evidence includes conceptual studies, systematic reviews, prototypes, simulations, industrial case studies, and empirical investigations. Because of this diversity, the review employed descriptive statistics and thematic synthesis rather than statistical meta-analysis.

Concepts such as semantic interoperability, ontology, knowledge graph, context representation, semantic model, and digital twin are also interpreted differently across research communities. Studies were therefore classified according to reported technologies, interoperability mechanisms, application domains, and evaluation approaches, although some conceptual overlap remains unavoidable.

8.4 Reviewer Judgment and Quality Assessment

Despite predefined eligibility criteria, structured evidence extraction, and quality assessment procedures, researcher judgment remained necessary during study selection, classification, and synthesis.

Many studies addressed multiple technologies simultaneously, so they were assigned to multiple categories, meaning reported frequencies reflect research emphasis rather than exclusive classifications. Study quality was assessed according to methodological rigor, evaluation methods, empirical support, and reporting transparency rather than a single scoring metric.

8.5 Technology Evolution and Industrial Generalizability

Semantic interoperability continues to evolve with advances in Semantic Web technologies, IIoT, cyber-physical systems, knowledge graphs, digital twins, and industrial information models. Consequently, the findings reflect published evidence within the review period and may not capture subsequent developments.

Industrial generalizability is also limited because most studies evaluate technologies through prototypes, simulations, or small-scale case studies. While these demonstrate technical feasibility, they may not fully represent large-scale industrial environments involving legacy systems, heterogeneous infrastructures, and evolving operational requirements. Additional longitudinal and industrial-scale evaluations remain necessary.

8.6 Limitations of Quantitative Evidence Synthesis

The quantitative analysis describes publication trends, technology adoption, industrial domains, and research themes rather than technological effectiveness. Frequencies indicate research emphasis and should not be interpreted as evidence that one semantic technology outperforms another.

Because many studies investigate multiple technologies, category frequencies overlap and do not total 100%. The quantitative results should therefore be interpreted as descriptive indicators supported by qualitative synthesis.

8.7 Risk of Bias

Potential sources of bias include:

- **Selection bias:** Variations in terminology may have prevented retrieval of relevant studies.
- **Geographical bias:** Evidence primarily reflects regions with strong investment in Industry 4.0 and semantic technologies.
- **Domain bias:** Manufacturing, IIoT, and digital twins dominate the literature, while healthcare, agriculture, logistics, and smart cities remain less represented.
- **Technology bias:** Frequently studied technologies receive greater research attention without implying superior performance.
- **Reporting bias:** Published studies tend to emphasize successful implementations over unsuccessful deployments.

8.8 Measures to Reduce Limitations

To minimise these limitations, the review searched multiple academic databases using broad terminology, applied predefined eligibility criteria, and used structured evidence extraction. Quantitative findings were interpreted descriptively, while reported challenges and contradictory findings were retained to provide a balanced evidence synthesis.

8.9 Summary of Limitations

This review is limited by database coverage, terminology variation, publication and language bias, methodological diversity, conceptual heterogeneity, evolving technologies, and limited large-scale industrial evidence. These limitations define the scope of the findings rather than undermine their validity.

Overall, the review provides an evidence-based synthesis of current research trends, recurring challenges, and knowledge gaps in semantic interoperability for autonomous data ecosystems. Future reviews would benefit from broader database coverage, multilingual evidence, greater inclusion of industrial sources, and periodic reassessment as semantic technologies continue to evolve.

IX. CONCLUSION

9.1 Overview

This systematic literature review examined semantic interoperability in autonomous data ecosystems using a PRISMA 2020 methodology. It synthesized evidence on semantic representation, ontology engineering, knowledge graphs, context modelling, semantic communication, and machine-to-machine information exchange.

The findings demonstrate that semantic interoperability is essential for heterogeneous systems—including IIoT, cyber-physical systems, digital twins, and autonomous agents—to exchange and consistently interpret information. While communication technologies enable data exchange, shared understanding depends on semantic representation, contextual interpretation, and machine-readable knowledge. Overall, the review consolidates current evidence, identifies recurring challenges, and highlights priorities for advancing semantic interoperability.

9.2 Findings in Relation to the Research Questions

RQ1: Semantic Representation Approaches

The review shows that semantic interoperability relies on complementary technologies rather than a single approach. Ontologies remain central for conceptual modelling, RDF and OWL provide machine-readable semantics, while knowledge graphs, digital twins, OPC UA, and Asset Administration Shell strengthen industrial interoperability. These technologies perform complementary roles, indicating that semantic representation is inherently heterogeneous.

RQ2: Semantic Technologies Enabling Interoperability

Semantic interoperability is enabled through ontologies, Semantic Web technologies, knowledge graphs, semantic reasoning, context models, digital twins, and industrial information models. Ontologies provide shared conceptual models, knowledge graphs integrate distributed information, semantic reasoning derives knowledge, and digital twins connect physical and virtual assets. However, most evidence is based on prototypes or simulations, with limited long-term industrial validation.

RQ3: Context Representation and Exchange

Context is fundamental for interpreting information across temporal, spatial, operational, environmental, and process conditions. Although widely recognized, inconsistent context models and varying levels of detail continue to limit cross-domain interoperability, highlighting the need for more standardized context representation.

RQ4: Technical Barriers

Major barriers include semantic heterogeneity, ontology alignment and evolution, fragmented context representation, heterogeneous data models, scalability, distributed reasoning, interoperability between standards, communication limitations, and limited industrial validation. These challenges are amplified in heterogeneous industrial environments and remain the primary obstacles to consistent semantic interpretation.

9.5 Findings in Relation to RQ5: Remaining Research Gaps

The review identifies persistent research gaps in ontology alignment, interoperability benchmarking, and industrial validation. Knowledge graphs, digital twins, and industrial information models require broader evaluation under dynamic, large-scale conditions. Additional gaps include fragmented context representation, limited cross-domain interoperability, and reliance on prototype-based studies. Stronger evidence is needed to demonstrate scalability, reproducibility, and long-term operational performance.

9.6 Academic Contribution

This review integrates fragmented research on Semantic Web technologies, ontology engineering, knowledge graphs, digital twins, IIoT, cyber-physical systems, and machine-to-machine communication into a unified evidence base. It distinguishes semantic representation, semantic integration, contextual interpretation, and demonstrated interoperability, showing that semantic technologies alone do not ensure operational interoperability. Rather than proposing a new framework, it consolidates existing knowledge and highlights unresolved empirical challenges.

9.7 Industrial and Practical Implications

The findings are applicable to manufacturing, smart energy, logistics, healthcare, and smart cities, where interoperability depends on semantic representation, contextual understanding, and standardized information models. Ontologies support knowledge representation, knowledge graphs integrate heterogeneous information, and digital twins with industrial information models improve interoperability between physical and digital assets. Effective implementation also requires semantic alignment, scalability, standards compatibility, ontology maintenance, and integration with existing infrastructures. The review consistently identifies a gap between research prototypes and validated industrial deployment.

9.8 Implications for Future Research

Future research should strengthen existing semantic technologies by prioritizing ontology alignment, standardized context representation, reproducible interoperability benchmarks, and comparative evaluation of semantic standards. Further work should address cross-domain interoperability, scalable knowledge graphs, distributed semantic reasoning, digital twin interoperability, and long-term validation in dynamic industrial environments through standardized evaluation and industrial-scale studies.

9.9 Overall Implications, Limitations, and Final Conclusion

This review confirms that semantic interoperability is fundamental to autonomous data ecosystems. Ontologies, RDF, OWL, knowledge graphs, semantic reasoning, digital twins, OPC UA information models, and Asset Administration Shell remain the principal technologies supporting semantic integration.

Despite significant progress, major challenges persist, including semantic heterogeneity, ontology alignment, context inconsistency, scalability, distributed reasoning, standards interoperability, and limited industrial validation. The review provides a comprehensive synthesis of current research without proposing a new framework but is limited by database coverage, English-language publications, methodological diversity, conceptual heterogeneity, publication bias, and limited long-term industrial evidence.

Future progress requires stronger empirical validation through improved ontology alignment, standardized benchmarking, consistent context representation, cross-domain evaluation, scalable knowledge graphs, distributed reasoning, and long-term industrial studies. Overall, semantic interoperability remains an evolving capability requiring continued advances in semantic representation, reasoning, context modelling, evaluation, and industrial validation. This review provides an evidence-based synthesis and identifies priorities for future research in autonomous data ecosystems [1]–[70].

X. DECLARATION

Ethical Approval

Not applicable. This study is a systematic literature review and meta-synthesis of published, peer-reviewed literature. It did not involve human participants, live vertebrate animals, human tissue, or personal sensitive data.

Competing Interests

The author declares that he has no competing financial or non-financial interests, personal relationships, or professional affiliations that could appear to influence the work reported in this paper.

Authors' Contributions

Ashok Nandah Ramakrishnan is the sole author of this manuscript. The author conceptualized the research objectives, designed the PRISMA 2020-compliant methodology, performed database searches, screened and evaluated included studies, conducted data extraction and quality assessment, synthesized the findings, and wrote the manuscript.

Funding

The author received no specific grant or financial support from any funding agency in the public, commercial, or non-profit sectors for the research, authorship, or publication of this article.

Availability of Data and Materials

All data generated or analyzed during this study are included in this published article (and its supplementary information files, where applicable). The primary source data consist of publicly available peer-reviewed academic publications retrieved from the databases indexed in the research methodology.

XI. REFERENCES

[1] T. Berners-Lee, J. Hendler, and O. Lassila, “The Semantic Web,” *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001.
[2] G. Antoniou and F. van Harmelen, *A Semantic Web Primer*, 3rd ed. Cambridge, MA, USA: MIT Press, 2012.
[3] N. Shadbolt, W. Hall, and T. Berners-Lee, “The Semantic Web revisited,” *IEEE Intelligent Systems*, vol. 21, no. 3, pp. 96–101, 2006.
[4] D. Allemang and J. Hendler, *Semantic Web for the Working Ontologist*, 3rd ed. New York, NY, USA: ACM/Morgan Kaufmann, 2020.
[5] J. Hebeler, M. Fisher, R. Blace, and A. Perez-Lopez, *Semantic Web Programming*. Indianapolis, IN, USA: Wiley, 2009.
[6] W3C, “RDF 1.1 Concepts and Abstract Syntax,” W3C Recommendation, 2014.
[7] W3C, “RDF 1.1 Primer,” W3C Recommendation, 2014.
[8] W3C, “OWL 2 Web Ontology Language Document Overview,” W3C Recommendation, 2012.
[9] W3C, “SPARQL 1.1 Query Language,” W3C Recommendation, 2013.
[10] W3C, “JSON-LD 1.1,” W3C Recommendation, 2020.
[11] N. Guarino, “Formal ontology and information systems,” in *Proc. 1st Int. Conf. Formal Ontology in Information Systems*, 1998, pp. 3–15.
[12] N. Guarino, D. Oberle, and S. Staab, “What is an ontology?,” in *Handbook on Ontologies*, 2nd ed. Berlin, Germany: Springer, 2009, pp. 1–17.
[13] Y. Sure, S. Staab, and R. Studer, “On-to-knowledge methodology,” in *Handbook on Ontologies*. Berlin, Germany: Springer, 2009.
[14] M. C. Suárez-Figueroa, A. Gómez-Pérez, E. Motta, and A. Gangemi, *Ontology Engineering in a Networked World*. Berlin, Germany: Springer, 2012.
[15] M. d’Aquin and N. F. Noy, “Where to publish and find ontologies? A survey of ontology libraries,” *Journal of Web Semantics*, vol. 11, pp. 96–111, 2012.
[16] A. Gangemi, C. Catenacci, M. Ciaramita, and J. Lehmann, “Modelling ontology evaluation and validation,” in *Proc. 3rd European Semantic Web Conf.*, 2006, pp. 140–154.
[17] A. Gómez-Pérez, “Evaluation of ontologies,” *International Journal of Intelligent Systems*, vol. 16, no. 3, pp. 391–409, 2001.
[18] M. Grüninger and M. S. Fox, “Methodology for the design and evaluation of ontologies,” in *Proc. Workshop on Basic Ontological Issues in Knowledge Sharing*, 1995.
[19] M. Uschold and M. King, “Towards a methodology for building ontologies,” in *Proc. Workshop on Basic Ontological Issues in Knowledge Sharing*, 1995.
[20] N. F. Noy and D. L. McGuinness, “Ontology development 101: A guide to creating your first ontology,” Stanford University, Stanford, CA, USA, Tech. Rep., 2001.
[21] A. Gangemi and P. Mika, “Understanding the semantic web through descriptions and situations,” in *Proc. 1st Int. Semantic Web Conf.*, 2002, pp. 689–706.
[22] P. Hitzler, M. Krötzsch, and S. Rudolph, *Foundations of Semantic Web Technologies*. Boca Raton, FL, USA: CRC Press, 2009.
[23] P. Hitzler, M. Krötzsch, B. Parsia, P. F. Patel-Schneider, and S. Rudolph, *OWL 2 Web Ontology Language: Primer*, W3C Recommendation, 2009.
[24] A. Hogan, E. Blomqvist, M. Cochez, C. d’Amato, G. de Melo, C. Gutierrez, S. Kirrane, J. E. Labra Gayo, R. Navigli, S. Neumaier, A.-C. N. Ngomo, A. Polleres, S. M. Rashid, A. Rula, L. Schmelzeisen, S. Sequeda, S. Staab, and A. Zimmermann, “Knowledge graphs,” *ACM Computing Surveys*, vol. 54, no. 4, 2021.
[25] J. E. Labra Gayo, E. Prud’hommeaux, I. Boneva, and D. Kontokostas, *Validating RDF Data*. San Rafael, CA, USA: Morgan & Claypool, 2017.
[26] H. Paulheim, “Knowledge graph refinement: A survey of approaches and evaluation methods,” *Semantic Web*, vol. 8, no. 3, pp. 489–508, 2017.
[27] A. Hogan, R. Verborgh, A. Polleres, and A. Harth, “Knowledge graphs,” *Communications of the ACM*, vol. 64, no. 9, pp. 96–104, 2021.
[28] J. E. L. Gayo, E. Prud’hommeaux, I. Boneva, and D. Kontokostas, *Validating RDF Data*. San Rafael, CA, USA: Morgan & Claypool, 2017.
[29] M. Yahya, J. G. Breslin, and M. I. Ali, “Semantic web and knowledge graphs for industry 4.0,” *Applied Sciences*, vol. 11, no. 11, 5110, 2021.
[30] M. C. Palmonari, G. V. Gentile, and L. Bellomarini, “Knowledge graphs and their applications in industry,” *Semantic Web*, 2021.
[31] A. H. C. Ng, C. Y. Chan, and K. K. Y. Wong, “Ontology-based semantic interoperability for industrial systems,” *IEEE Access*, 2020.

[32] B. Kulvatunyou, N. Ivezic, and Y. Lee, “On enhancing communication of the manufacturing service capability information using reference ontology,” *International Journal of Computer Integrated Manufacturing*, vol. 27, no. 3, pp. 276–291, 2014.
[33] M. Aydin and S. Aydin, “Semantic and syntactic interoperability for agricultural open-data platforms in the context of IoT using crop-specific trait ontologies,” *Applied Sciences*, vol. 10, no. 13, 4460, 2020.
[34] M. Platenius-Mohr, S. Malakuti, S. Grüner, and others, “File-and API-based interoperability of digital twins by model transformation: An IIoT case study using Asset Administration Shell,” *Future Generation Computer Systems*, 2020.
[35] V. Siatras, E. Bakopoulos, P. Mavrothalassitis, et al., “On the use of asset administration shell for modeling and deploying production scheduling agents within a multi-agent system,” *Applied Sciences*, vol. 13, no. 17, 9540, 2023.
[36] B. Kim, S. Kim, H. Teijgeler, J. Lee, J. Y. Lee, D. Lim, et al., “Use of asset administration shell coupled with ISO 15926 to facilitate the exchange of equipment condition and health status data of a process plant,” *Processes*, vol. 10, no. 10, 2155, 2022.
[37] X. Wang, X. Hu, Z. Ren, T. Tian, and J. Wan, “Knowledge-graph-based multi-domain model integration method for digital-twin workshops,” *The International Journal of Advanced Manufacturing Technology*, 2023.
[38] K. Angele, D. Fensel, E. Huaman, E. Kärle, et al., “Semantic web empowered e-tourism,” in *Handbook of e-Tourism*. Cham, Switzerland: Springer, 2022.
[39] J. Zenkert, C. Weber, M. Dornhöfer, H. Abu-Rasheed, et al., “Knowledge integration in smart factories,” *Encyclopedia*, vol. 1, no. 3, pp. 61, 2021.
[40] D. Sun, J. Hu, H. Wu, J. Wu, J. Yang, Q. Z. Sheng, et al., “A comprehensive survey on collaborative data-access enablers in the IIoT,” *ACM Computing Surveys*, 2023.
[41] M. A. Jarwar, J. Watson, and S. Ali, “Modeling industrial IoT security using ontologies: A systematic review,” *IEEE Open Journal of the Communications Society*, 2025.
[42] M. Arazzi, A. Nocera, and E. Storti, “The SemIOE ontology: A semantic model solution for an IoE-based industry,” *IEEE Internet of Things Journal*, 2024.
[43] T. Tudorache, “Ontology engineering: Current state, challenges, and future directions,” *Semantic Web*, 2020.
[44] P. Bonte, F. Ongenae, and R. Tommasini, “A holistic view over ontologies for streaming linked data,” *Semantic Web*, 2024.
[45] V. Siatras, E. Bakopoulos, P. Mavrothalassitis, et al., “Asset administration shell-based modeling of production scheduling agents,” *Applied Sciences*, 2023.
[46] A. De Baas, P. Del Nostro, J. Friis, E. Ghedini, et al., “Review and alignment of domain-level ontologies for materials science,” 2023.
[47] B. Yaman, K. Thompson, F. Fahey, et al., “LinkedDataOps: Quality oriented end-to-end geospatial linked data production governance,” *Semantic Web*, 2024.
[48] N. Ferranti, J. F. de Souza, S. Ahmetaj, et al., “Formalizing and validating Wikidata's property constraints using SHACL and SPARQL,” *Semantic Web*, 2024.
[49] L. Tailhardat, R. Troncy, and Y. Chabot, “Anomaly detection using knowledge graphs: A survey for network management and cybersecurity application,” *ACM Computing Surveys*, 2025.
[50] P. F. Patel-Schneider, P. Hayes, and I. Horrocks, “OWL Web Ontology Language semantics and abstract syntax,” W3C Recommendation, 2004.
[51] D. Brickley and R. V. Guha, “RDF Schema 1.1,” W3C Recommendation, 2014.
[52] P. F. Patel-Schneider, P. Hayes, and I. Horrocks, “Web Ontology Language (OWL),” W3C Recommendation, 2004.
[53] A. Harth, K. Hose, R. Schenkel, M. Theobald, and G. Weikum, *Database Techniques for RDF Querying and Reasoning*. Berlin, Germany: Springer, 2012.
[54] C. Bizer, T. Heath, and T. Berners-Lee, “Linked data—The story so far,” *International Journal on Semantic Web and Information Systems*, vol. 5, no. 3, pp. 1–22, 2009.
[55] T. Heath and C. Bizer, *Linked Data: Evolving the Web into a Global Data Space*. San Rafael, CA, USA: Morgan & Claypool, 2011.
[56] M. L. Khodadadi, S. A. A. Shah, and others, “Semantic interoperability in Internet of Things environments: A systematic review,” *IEEE Access*, 2020.
[57] M. Weyrich and C. Ebert, “Reference architectures for the Internet of Things,” *IEEE Software*, vol. 33, no. 1, pp. 112–116, 2016.
[58] M. Wollschlaeger, T. Sauter, and J. Jasperneite, “The future of industrial communication: Automation networks in the era of the Internet of Things and Industry 4.0,” *IEEE Industrial Electronics Magazine*, vol. 11, no. 1, pp. 17–27, 2017.
[59] L. Da Xu, W. He, and S. Li, “Internet of Things in industries: A survey,” *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, 2014.
[60] M. Hermann, T. Pentek, and B. Otto, “Design principles for Industrie 4.0 scenarios,” in *Proc. 49th Hawaii Int. Conf. System Sciences*, 2016, pp. 3928–3937.

[61] M. A. L. M. R. S. S. and others, “Industry 4.0: The industrial Internet of Things and its applications,” *IEEE*, 2017.
[62] J. Lee, B. Bagheri, and H.-A. Kao, “A cyber-physical systems architecture for Industry 4.0-based manufacturing systems,” *Manufacturing Letters*, vol. 3, pp. 18–23, 2015.
[63] L. Monostori, “Cyber-physical production systems: Roots, expectations and R&D challenges,” *Procedia CIRP*, vol. 17, pp. 9–13, 2014.
[64] J. Lee, E. Lapira, B. Bagheri, and H.-A. Kao, “Recent advances and trends in predictive manufacturing systems in big data environment,” *Manufacturing Letters*, vol. 1, no. 1, pp. 38–41, 2013.
[65] M. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, “Digital Twin in manufacturing: A categorical literature review and classification,” *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.
[66] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, “Digital twin in industry: State-of-the-art,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019.
[67] F. Tao, M. Zhang, Y. Liu, and A. Y. C. Nee, “Digital twin driven prognostics and health management for complex equipment,” *CIRP Annals*, vol. 67, no. 1, pp. 169–172, 2018.
[68] M. Grieves and J. Vickers, “Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems,” in *Transdisciplinary Perspectives on Complex Systems*. Cham, Switzerland: Springer, 2017, pp. 85–113.
[69] M. Weyrich, M. Schmidt, and others, “Digital twins in manufacturing: A systematic literature review,” *CIRP Journal of Manufacturing Science and Technology*, 2020.
[70] M. Schluse, M. Priggemeyer, L. Atorf, and J. Rossmann, “Experimentable digital twins for industrial applications,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1722–1731, 2018.