

CLASSICAL AND AI-BASED EXPLAINABILITY OF ONTOLOGIES ON THE EXAMPLE OF THE DIGITAL REFERENCE – THE SEMANTIC WEB FOR SEMICONDUCTOR AND SUPPLY CHAINS CONTAINING SEMICONDUCTORS

Marta Bonik¹, Eleni Tsatsouli², Hans Ehm¹, and George Dimitrakopoulos²

¹Department of Supply Chain Innovation, Infineon Technologies AG, Munich, GERMANY

²Department of Informatics and Telematics Harokopio University of Athens, Athens, GREECE

ABSTRACT

Ontologies are essential for structuring knowledge in complex domains like semiconductor supply chains but often remain inaccessible to non-technical users. This paper presents a combined classical and AI-based approach to improve ontology explainability, using the [Digital Reference](#) (DR) as a case study. The first approach leverages classical ontology visualization tools to engage users in the learning process through interactive access and feedback. The second approach integrates Neo4j graph databases and Python with a large language model (LLM)-based architecture to enable natural language querying of ontologies. A key contribution is the implementation of a post-processing layer incorporating query syntax validation, ontology schema verification, fallback templates, and entity filtering to ensure accurate responses. The approach is evaluated through multiple natural language queries, demonstrating improved usability, robustness, and adaptability. By bridging traditional query methods and AI interfaces, this work supports broader adoption of ontology-driven systems within the Semantic Web and industry.

1 INTRODUCTION

In an era where digitalization is transforming industries, the need for structured and machine-readable knowledge representation has never been more critical. Ontologies serve as a foundational framework that enables both humans and computers to share a common understanding of data, facilitating seamless interoperability across digital systems. (Berners-Lee et al. 2001) highlight that the modern web is designed primarily for human consumption, limiting machines' ability to process and interpret information meaningfully. Ontologies bridge this gap by providing a structured and semantically rich representation of knowledge, making it accessible to both humans and artificial intelligence (AI) systems.

However, the role of ontologies extends beyond simple knowledge organization. They are a key enabler of digitalization technologies such as digital twins, which require a semantic foundation to accurately replicate real-world entities and processes. (Berners-Lee et al. 2008) emphasize that logical frameworks like RDF and OWL allow data to be structured in a way that supports automated reasoning, decision-making, and intelligent data integration. This structured representation is essential for complex domains such as semiconductor supply chains, where diverse stakeholders – from manufacturers to logistics providers – must operate within a unified knowledge framework to enhance efficiency and reduce operational uncertainties.

Despite their advantages, ontologies are often perceived as complex and difficult to understand, limiting their widespread adoption. This is particularly true in industrial applications, where domain experts may lack the technical expertise to navigate ontology-based systems effectively. Addressing this challenge requires innovative approaches to teaching and explaining ontologies, making them more accessible to non-experts while maintaining their robustness for AI-driven applications. Recent advancements in LLMs present an opportunity to bridge this gap by enhancing ontology explainability and providing interactive learning mechanisms.

This paper explores methodologies for improving the teaching and usability of ontologies, focusing on their application in semiconductor supply chains. We propose leveraging LLM-driven interfaces to facilitate ontology understanding, with a case study on the DR model. The Digital Reference is a top-level Semantic Web model for semiconductors and supply chains containing semiconductors, depicting a combination of various supply chain pillars and semiconductor production concepts. It comprises over 1,200 classes organized into 14 thematic areas, referred to as lobes. By integrating advanced ontology embeddings and prompting techniques, we aim to enhance the usability of semantic models, ensuring that both technical and non-technical users can effectively engage with ontology-based systems. Furthermore, we compare traditional and LLM-based methods to evaluate their effectiveness in different learning contexts.

By demonstrating the significance of ontological knowledge representation in semiconductor supply chains and highlighting practical approaches for improving explainability, this work contributes to the broader field of ontology education and digitalization in industry. This establishes a foundation for future research on interactive, AI-assisted ontology training and supports broader adoption of ontological frameworks in real-world applications.

This work aligns with the development of the Semantic Web by advancing explainability techniques on the example of the Digital Reference, specifically within the semiconductor and supply chain domain. Through the combination of classical and AI-based methodologies, it addresses current challenges in making complex ontologies more accessible and usable across industrial contexts.

2 RELATED WORK

Ontologies provide a structured framework for knowledge representation, allowing both humans and machines to share a common semantic understanding, thereby facilitating automation, AI reasoning, and digital interoperability (Berners-Lee et al. 2001). Ontological unpacking enhances semantic interoperability by making conceptual models more explicit and interpretable (Guizzardi and Guarino 2024). These foundations are crucial for domains such as manufacturing and supply chains, where standardized data representation is essential for operational efficiency.

Moreover, digital twins rely on ontologies to structure data and maintain semantic consistency across digital replicas of real-world entities (Berners-Lee et al. 2008). By structuring information in a machine-readable way, ontologies enable interoperability across systems. The Digital Reference serves as a practical example of this approach, as it semantically connects semiconductor supply chains, ensuring real-time interoperability among stakeholders (Ramzy et al. 2021). Semantic data integration plays a key role in improving decision-making by harmonizing supply chain data across multiple sources (Ramzy 2022).

As ontologies grow in complexity, explainability becomes a challenge, particularly in AI-driven decision-making (Guizzardi and Guarino 2024). The need for transparency in large-scale semantic models has led to the integration of neuro-symbolic frameworks, which combine enterprise knowledge graphs (EKGs) with LLMs to enhance interpretability (Baldazzi et al. 2024). Moreover, aligning artificial neural networks with ontologies has shown to improve explainability by providing symbolic justifications for AI outputs, ensuring better interpretability of machine-driven decisions (de Sousa Ribeiro and Leite 2021). At the same time, domain-specific ontologies integrated with LLMs enhance transparency and trust in AI-driven systems by grounding responses in structured knowledge (Amara et al. 2024).

The semiconductor industry presents unique challenges that make it an ideal case for ontology-driven digitalization (Herding et al. 2021). The complexity of semiconductor production, characterized by global dependencies and rapid technological advancements, requires structured data representation for efficient supply chain management. Ontology-based Explainable AI (O-XAI) has been identified as a crucial tool for improving trust and decision transparency in AI-driven manufacturing, helping stakeholders interpret AI-assisted recommendations and predictions (Naqvi et al. 2024). Given these considerations, ontology-based decision support systems are critical for reducing inefficiencies and ensuring robust supply chain operations.

Recent advancements in ontology extraction have focused on integrating AI with human expertise to improve knowledge representation (Dimitrakopoulos et al. 2024). AI-assisted methods for ontology generation using LLMs, specifically ChatGPT 4 Turbo, have been developed to enhance the creation of ontologies related to EU semiconductors and supply chains containing semiconductors. Combining automated ontology extraction with expert validation ensures both accuracy and semantic richness, while human oversight remains critical for refining logical consistency and preventing ambiguity. This integration of AI and domain expertise allows for scalable yet reliable ontology development. Addressing a related challenge, the LLMs4SchemaDiscovery framework by Sadruddin et al. (Sadruddin et al. 2025) introduces Schema-Miner, a tool that automates schema extraction from unstructured data while incorporating expert feedback for relevance and accuracy. Its iterative human-in-the-loop workflow highlights the potential of AI-driven tools to improve usability and explainability in ontology systems, aligning closely with the goals of this paper.

Understanding and using ontologies effectively remains a challenge, necessitating innovative training approaches (Rahayu et al. 2022). Various strategies have been explored, including the alignment of neural networks with ontologies to improve AI interpretability (de Sousa Ribeiro and Leite 2021). Additionally, interactive, LLM-enhanced training modules have been proposed to ensure accessibility for both technical and non-technical users (Ramzy et al. 2021). Ontology-based training methods for demand fulfillment in semiconductor supply chains suggest that interactive AI systems can guide users through complex supply chain structures, helping stakeholders better navigate ontology-driven decision support systems (Herding et al. 2021).

The reviewed literature confirms that ontologies are fundamental to digitalization, AI transparency, and supply chain optimization. Their role in semantic interoperability, digital twins, and Explainable AI is well-documented. However, accessibility and usability remain key challenges. This paper builds on these foundations by proposing an LLM-enhanced approach to ontology training and explainability, specifically focused on the Digital Reference for semiconductor supply chains. The Digital Reference is a top-level Semantic Web model for semiconductors and supply chains containing semiconductors, depicting a combination of various supply chain pillars and semiconductor production concepts. It comprises over 1,200 classes organized into 14 thematic areas, referred to as lobes. By leveraging ontology embeddings, prompting techniques, and interactive models, this work aims to bridge the usability gap and enhance the adoption of ontology-driven solutions in industrial settings.

3 METHODOLOGY

While prior research highlights the significance of ontologies in fostering semantic interoperability and enhancing decision-making, it also emphasizes challenges related to ontology complexity, usability, and explainability, especially for non-technical stakeholders. Traditional ontology interaction methods often require technical expertise, limiting broader accessibility and adoption in industrial contexts, such as semiconductor supply chains.

To address these issues, this paper evaluates two complementary methodologies: the classical interactive approach and an AI-based approach for ontology explainability. The discussion begins by briefly describing the classical approach to establish a baseline for traditional ontology learning practices. It then introduces two innovative approaches: first, a more interactive variation of the classical approach that facilitates the ontology learning process, and second, an AI-driven framework that further enhances and extends the capabilities of the interactive approach.

3.1 Classical and Interactive Approach

The classical approach for engaging with ontologies typically involves two main methods. The first method is the visual exploration of ontologies, which can be done either through a tree-based class hierarchy provided by tools such as Protégé or through graph-based representations offered by tools like WebVOWL.

These tools are designed to help users navigate the ontology structure by visualizing its relationships and components in an accessible manner. The second method is graph querying, which involves using structured query languages, such as SPARQL or Cypher, to retrieve and manipulate ontology data. This approach is particularly useful for users who need precise and targeted results from their ontology interactions.

While both methods are widely adopted, they have inherent limitations. Graph visualization can become excessively complex and overwhelming when dealing with large ontologies, such as the Digital Reference ontology. As ontologies grow in scope and complexity, these challenges become even more pronounced, making it challenging for users to navigate and derive meaningful insights from the data.

Graph querying relies on direct engagement with graph databases and structured query languages. In this method, the Digital Reference ontology is instantiated within a graph database that visualizes entities and their relationships, such as suppliers, manufacturers, and equipment, in a comprehensible graph structure.

Users interact with the ontology primarily through SPARQL or Cypher queries, which necessitate familiarity with database query languages and the specific ontology schema. There are two common ways this interaction is performed:

1. Direct scripting e.g. within the Neo4j environment: Users manually write and execute Cypher queries through the Neo4j Browser or Neo4j Desktop interface to retrieve, manipulate, and visualize ontology data directly within the database platform.
2. Python integration (e.g., Google Colab): Python scripts are used to establish a connection to the Neo4j database, allowing users to automate query execution, process results programmatically, and visualize outputs externally in a more flexible environment.

However, this approach presents key challenges. The technical barrier limits participation to users familiar with Cypher and graph databases, excluding many domain experts. Manual query formulation is also error-prone, leading to syntax mistakes or incomplete results. As the ontology grows in complexity, maintaining and scaling this process becomes increasingly inefficient.

Despite these drawbacks, classical approaches can still be effective when adopted in a user-centric manner that eliminates the need for deep technical knowledge. By focusing on intuitive design, such approaches can allow users to efficiently learn and interact with ontologies. To demonstrate this, we developed [OntoLearn](#) (OntoLearn 2025), a software application that facilitates interaction with the Digital Reference model in an accessible and interactive way. This application enables users to familiarize themselves with the detailed definitions of Digital Reference classes, their attributes, and their relationships to other classes. By engaging with the ontology in short, focused interactions, users can explore class-to-class relationships and gain a deeper understanding of the ontology structure. This intuitive interaction is achieved through graph visualization combined with attribute-specific details, providing a comprehensive yet user-friendly exploration experience.

OntoLearn utilizes the OWLready2 Python library to parse and interact with the ontology's structure, while the NetworkX library is employed to visualize ontology classes and their relationships. By integrating these tools, the application enables users to explore the Digital Reference model efficiently without requiring technical expertise in ontology querying or modeling.

The challenges outlined above underscore the need for more intuitive and adaptable solutions. While OntoLearn simplifies interaction with ontologies, a gap remains in enabling users to formulate complex queries effortlessly and accurately. To bridge this gap, we introduce an AI-based approach that leverages natural language processing and machine learning, detailed in the following section.

3.2 AI-Based Approach for Ontology Explainability

To enhance supply chain collaboration and address challenges in ontology usability within the semiconductor industry, this paper presents a structured methodology that integrates artificial intelligence and human expertise to facilitate ontology explainability. The approach is centered on the synergy between graph-

based knowledge representation, advanced language models, and expert validation. By combining the capabilities of Neo4j, Python, and LLMs, specifically the Llama 3 model, this methodology aims to enable intuitive, natural language interaction with ontologies while ensuring semantic accuracy and domain relevance.

The methodology leverages the complementary strengths of AI technologies and human input. LLMs excel at processing large datasets, identifying patterns, and generating structured outputs, while human oversight provides contextual understanding and ensures semantic relevance. This interaction promotes efficiency and adaptability, supporting the continuous refinement of ontology structures.

While LLMs enable natural language interaction with ontologies, translating queries into formal query languages like Cypher and SPARQL involves several challenges. Ambiguity in natural language often requires disambiguation of terms or relationships, especially in complex ontologies. Additionally, incorrect entity mapping or syntax issues can lead to invalid queries, reducing system reliability (D'Abromo et al. 2025). To address these issues, the proposed system uses schema-aware querying by comparing user input to a list of DR entities, ensuring control over which IRIs are used for query generation. Explicit prompts are employed to guide the system's behavior and produce the desired outputs, mitigating ambiguity. These strategies are integral to bridging the gap between natural language and formal query systems.

The proposed methodology is designed not only for users without technical expertise, but also for individuals working with complex ontologies who face challenges in exploring and interacting with them effectively. The framework supports scalability and replicability, providing clearly defined steps, including ontology modeling, natural language query processing, and iterative refinement, to ensure consistent application across various contexts.

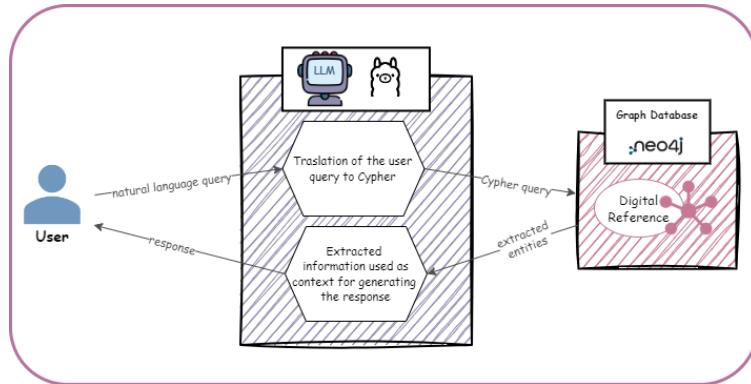


Figure 1: System Architecture.

As shown in Figure 1, the system architecture consists of the following steps: the user submits a query, which is translated by the LLM-based layer into a formal Cypher representation; relevant entities are extracted from the ontology model; the extracted information supports response generation in the LLM layer, and the final response is returned to the user. Below, we describe each phase of the methodology, illustrating how the integration of AI-driven tools and human validation contributes to the development of accessible and transparent ontology systems tailored to semiconductor supply chains.

Step 1: Ontology Modeling in Neo4j:

In the initial phase, the Digital Reference ontology is formalized within the Neo4j graph database. The ontology encapsulates key entities and relationships pertinent to semiconductor supply chains, such as Supplier, Manufacturer, and Equipment. Each entity is modeled as a labeled node, and domain-specific relationships (e.g., SUPPLIES, USES) are represented as directed edges. Relevant properties are

assigned to both nodes and relationships to reflect real-world attributes and ensure semantic richness. This graph-based structure facilitates machine readability, efficient querying, and human interpretability.

Step 2: Python Integration:

The second phase involves establishing a Python environment to interface with the Neo4j database. Python scripts are developed to manage the connection, handle user input, execute Cypher queries, and process query results. This integration enables seamless interoperability between the ontology graph and external systems. Additionally, Python serves as the platform through which AI-driven natural language processing is incorporated.

Step 3: Natural Language Query Interpretation via LLM:

The third phase focuses on enabling natural language interaction with the ontology. The Llama 3 model is utilized to interpret user queries formulated in everyday language. To further enhance the handling of complex relationship queries, such as *“What is the relation between Workshop and Current WIP?”*, we introduce a specialized condition leveraging the shortest path algorithm in Neo4j, along with a relation limit constraint. This ensures that the model retrieves the most relevant and concise connections without overwhelming users with unnecessary complexity.

By limiting the maximum number of relationships (e.g., up to 6), the system guarantees semantically meaningful and interpretable results, supporting efficient ontology exploration and allowing users without technical expertise to engage effectively with the ontology.

Additionally, to improve the system’s ability to explain the specific structure of the Digital Reference, a system prompt is included to explicitly define its lobes. This ensures accurate explanation and categorization of ontology entities during natural language interaction.

Step 4: Query Execution and Result Handling:

Once the Cypher query is generated, it is executed within the Neo4j database via the Python interface. The query retrieves relevant data, identifying entities connected to the specified node based on the ontology’s structure. The results are captured and processed in a structured format, ensuring coherence and consistency.

Step 5: Natural Language Response Generation:

In the final phase, the structured data retrieved from Neo4j is converted back into natural language responses. The processed output provides users with clear, semantically accurate information, closing the interactive loop. This ensures that ontology content remains accessible and comprehensible to both technical and non-technical stakeholders.

By following this methodology, the gap between complex ontological structures and user-friendly interaction is bridged, fostering greater adoption and usability of ontology-driven systems in the semiconductor supply chain domain. Nevertheless, like all AI-driven systems, the use of LLMs introduces certain challenges related to query accuracy and consistency. To address these, we introduce an additional post-processing mechanism, described in the following subsection.

3.3 Post-processing and Error Handling

While the integration of a Llama-based architecture enables intuitive natural language querying, AI-generated Cypher queries often encounter practical challenges. Common issues include syntactic errors, incorrect use of ontology labels, and incomplete or semantically inconsistent relationship patterns, particularly when processing complex or ambiguous user inputs.

To address these challenges and enhance the reliability of the interaction, the methodology incorporates a structured post-processing and validation layer, ensuring robust query execution. Specifically, the following mechanisms are implemented:

1. **Query Syntax Validation:** Each AI-generated Cypher query undergoes syntax checking via Neo4j's parser to ensure correctness. Any detected parsing errors are handled gracefully, preventing execution failures and maintaining system stability.
2. **Ontology Schema Verification:** The generated queries are cross-referenced against the DR ontology schema. Entity labels and relationship types are verified using a predefined mapping, ensuring that only valid and semantically consistent elements are queried. Queries containing unknown or incorrect entities are automatically flagged and handled appropriately.
3. **Fallback Templates for Critical Queries:** For essential query types, such as retrieving entity definitions or exploring specific relationships, predefined Cypher templates are used as reliable backups. These templates guarantee meaningful responses even in cases where the AI-generated query does not meet the required standards.

Additionally, the system filters out non-informative or meaningless entity identifiers (e.g., hash-like IDs), further enhancing the clarity and usability of the results. Through this layered validation process, the methodology ensures the generation of syntactically valid, semantically aligned, and user-comprehensible outputs, supporting both technical accuracy and accessibility. Moreover, by systematically handling errors and incorporating fallback mechanisms, the approach offers a practical and reproducible solution for improving ontology explainability in industrial applications.

4 IMPLEMENTATION: CASE STUDY

In this study, the proposed methodologies were applied to the Digital Reference ontology, which captures key entities and processes within semiconductor supply chains. Given the complexity of the DR, manually exploring and querying its structure is time-consuming and prone to oversight. Thus, the classical interactive learning tool involves users in learning sessions where they can familiarize themselves with DR classes and fostering a deeper understanding of the ontology's structure and content. Meanwhile, the AI-driven approach streamlines the explainability process, providing an intuitive interface for both technical and non-technical users.

4.1 Ontology Learning Tool

OntoLearn, the ontology learning tool implemented as part of the Digital Reference ecosystem, provides users with an intuitive and engaging way to familiarize themselves with complex ontology models while contributing to their refinement. Accessible through a web-based interface, the tool is seamlessly integrated with the DR, leveraging the rich ontology data to enable exploration, learning, and feedback without requiring additional software installation.

Learning sessions can be customized, starting with a minimum of five questions for a quick and focused experience. Users can choose to extend the session by selecting a larger number of questions, enabling a deeper and more comprehensive exploration of the ontology content if desired.

During a learning session, a randomized selection of ontology classes is presented to the user. Each selected class is accompanied by a definition extracted directly from the interpretation, providing users with a clear understanding of its context and meaning. Additionally, the tool incorporates a dynamic graph visualization feature that displays the selected class, its attributes, and all directly connected classes, effectively showing all the triples in which the class participates, as shown in Figure 2. In this example, the Semiconductor Company class and its relationships are displayed, illustrating the tool's ability to present entities and their direct connections interactively.

One of the tool's standout features is its built-in mechanism for user feedback. While interacting with a class, users may notice errors, inconsistencies, or unclear information. To address this, the tool includes a "Suggest Improvement" button that allows users to propose corrections or report issues directly to the Digital Reference administrative team. This functionality not only helps keep the ontology accurate and up-to-date but also encourages users to actively contribute to its improvement, fostering a collaborative

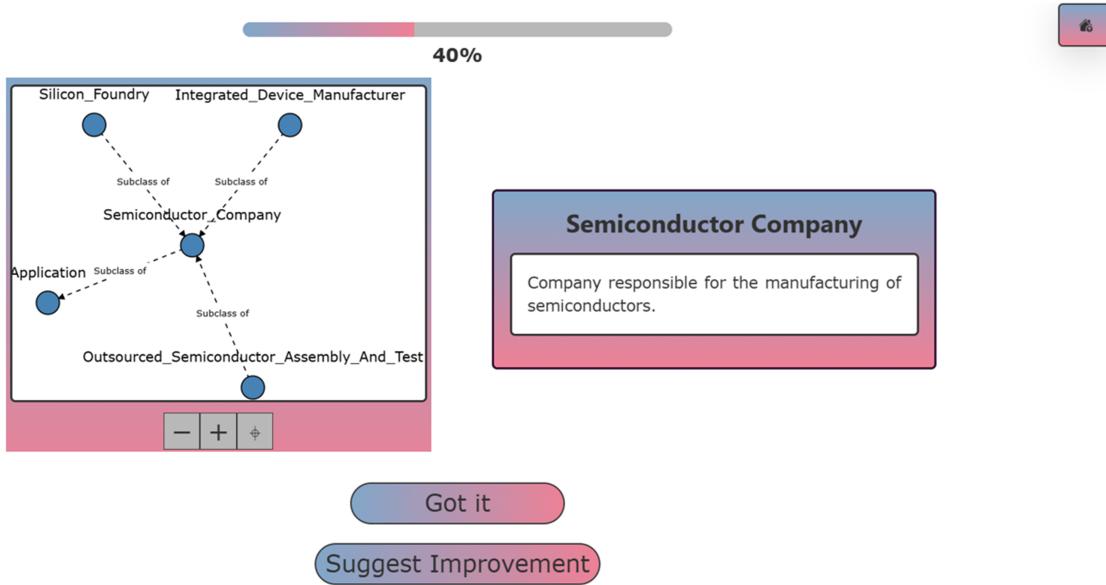


Figure 2: Visualization of the Semiconductor Company Class.

environment. Initial user feedback improved around 20% of the DR classes, with changes introduced during regular updates and primarily related to definition modifications and missing connections.

Moving forward, this tool has the potential to be extended with even more personalized and advanced features to make the learning experience richer and more tailored. For instance, users could create personal accounts where their progress is tracked over time, allowing them to return to previously explored classes, resume sessions, or track their overall mastery of the ontology.

4.2 AI-based System

The system aims to precisely interpret user queries and extract the requested information from the model. A key feature is its ability to provide clear definitions of ontology entities. For instance, when asked, *"What does Decision Making Unit mean?"* the system identified the entity and retrieved its definition from the DR ontology using the query:

```
MATCH (n {name: 'Decision Making Unit'}) RETURN n.definition
```

The result provided a clear, human-readable explanation of the entity, making the ontology accessible to non-experts. Furthermore, the system can uncover indirect relationships between entities. For instance, in response to the query: *"What is the relation between Workshop and Current WIP?"* the system explored the graph, identified multi-step paths connecting Workshop and Current WIP and generated the query:

```
MATCH p = shortestPath( (a:Workshop) -[*] - (b: 'Current WIP')) RETURN p
```

This allows users to visualize and understand complex relationships that are difficult to manually trace. Importantly, the methodology not only facilitates querying direct and hidden relationships but also enables users to retrieve definitions of entities, explore related classes, and answer general structural questions about the ontology, such as how many nodes the ontology contains, how many relationships a specific class has, or the total number of distinct entity types represented.

The proposed tools, OntoLearn and the AI-based querying framework, can integrate with semiconductor industry systems to enhance training, supply chain management, and digital twin applications. OntoLearn could support training modules by enabling employees to explore and refine the DR ontology, while the querying framework could assist in real-time supply chain monitoring and bottleneck identification.

Additionally, integration with digital twin systems could leverage ontology data to optimize production processes and synchronize virtual and physical systems, demonstrating scalability for industrial use.

4.2.1 Evaluation and Validation

The system's performance was tested on the following five tasks using natural language queries to validate its robustness and flexibility:

1. Retrieving definitions of specific entities.
2. Exploring direct and indirect relationships between supply chain components within the DR model.
3. Understanding the structure of the model by querying the total number of nodes and relationships in the ontology or per lobe.
4. Investigating all entities related to a specific class.
5. Providing summaries of Digital Reference classes belonging to specific topics of interest.

Testing and validation processes demonstrated that the system performs well across all given tasks, reliably delivering the expected results. Despite occasional limitations, the system consistently supported meaningful interactions with the ontology and provided accurate outputs aligned with user queries. To evaluate the reliability of the system, we tested three core components: the error handling logic, the Cypher query generation, and the interpretation of results. These steps are critical to ensuring that users receive meaningful responses even when input queries are ambiguous or partially incorrect.

First, the error handling mechanism validates each query by checking for syntax issues, verifying that entities exist in the ontology, and filtering out irrelevant or malformed identifiers. If a query fails or yields no results, fallback templates are triggered to maintain usability and provide informative feedback.

Second, the system dynamically generates Cypher queries to explore relationships within the ontology. An example is shown below, where the system identifies how the entities `Workshop` and `Current WIP` are connected:

```
MATCH (n1:Entity), (n2:Entity)
WHERE toLower(n1.uri) CONTAINS 'workshop'
    AND toLower(n2.uri) CONTAINS 'current wip'
MATCH path = shortestPath((n1)-[*..6]-(n2))
RETURN [rel IN relationships(path) | type(rel)],
       [n IN nodes(path) | n.uri]
```

Finally, the output is translated into a natural language explanation. In this case, the system identified an indirect link: a `Workshop` contains an `Equipment Group`, which references a `Current WIP` entity through primary and foreign keys. This allows users to understand complex connections without inspecting raw graph data.

Performance metrics such as query accuracy (alignment between generated queries and user intent) and response accuracy are widely recognized as essential for evaluating query systems(Potdar et al. 2025). Query accuracy measures how well queries match user requirements, while response accuracy, often related to user satisfaction, assessed through surveys or trials, provides insights into accessibility and usability.

We applied a similar evaluation approach to assess the system's performance across five tasks. Query accuracy was measured as the ratio of correctly generated queries to the total number of attempted queries. A query was considered "correct" if it was semantically and syntactically aligned with the ontology's structure and schema. Validation involved checking the generated query against the ontology file to ensure that the referenced entities, relationships, and attributes existed within the Digital Reference schema. For example, queries such as "*What is the definition of Planning Processes?*" and "*What does Decision Making Unit mean?*" were validated by confirming the correct identification of the target class and its definition. Response accuracy was evaluated using two methods depending on the query type. For deterministic queries, such as retrieving structural information, the system's outputs were directly compared to the ontology file

and deemed correct if they matched the expected results. For more complex queries requiring domain-specific interpretation, responses were assessed by domain experts. For example, thematic summaries of related classes were evaluated by experts for relevance to the provided topic. Both query and response accuracies were quantified using standardized metrics, with query accuracy and response accuracy expressed as percentages of correct outputs relative to total attempts, ensuring rigorous validation across diverse tasks and query complexities.

For **retrieving definitions**, a test set of 30 queries about classes' definitions across 14 lobes achieved 80% query accuracy and 100% response accuracy. Errors arose from ambiguous queries leading to misassigned classes, while accurate definitions were consistently retrieved once the correct class was identified. For **exploring relationships**, 10 one-hop and 10 multi-hop queries (e.g., *"What is the relation between Workshop and Current WIP?"*) achieved 70% query accuracy and 100% response accuracy. Lower query accuracy was due to challenges in interpreting complex queries. For understanding the model's structure, queries about nodes, lobes (e.g., *"How many lobes are in the DR?"*), and class counts achieved 100% accuracy due to the well-structured ontology.

For **investigating related classes**, 20 queries (e.g., *"Which entities are related to Supplier?"*) achieved 80% query accuracy and 100% response accuracy, with errors caused by misinterpreted queries or misassigned classes. Figure 3 illustrates the system's response to a query about the Equipment class, showing its relationships with other entities and demonstrating the system's ability to capture and visualize class connections. For **summarizing classes by topic**, five thematic areas (e.g., Semiconductor Demand Forecasting and Order Management) were tested. While relevant lists were generated, inconsistencies across iterations highlighted challenges in precision and completeness, requiring standardization.

Welcome To Ontology Explainability Tool

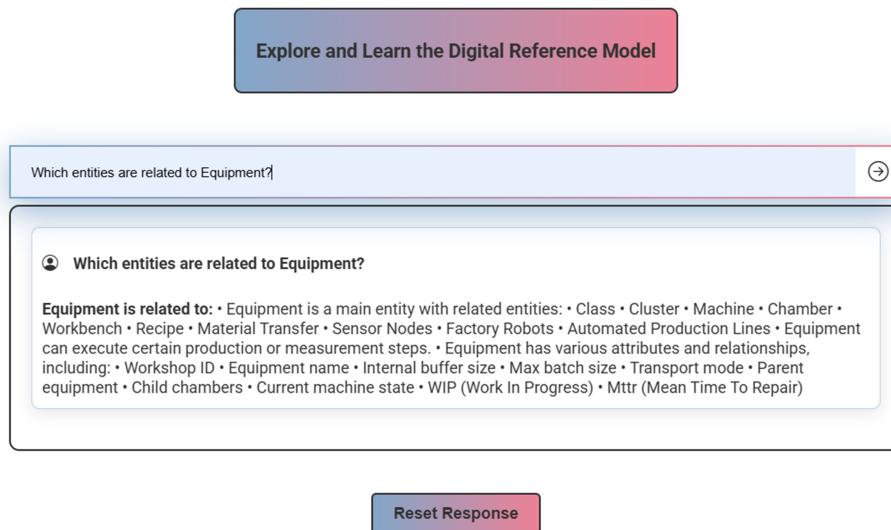


Figure 3: Entities related to the Equipment Class. Response of the system.

Overall, the system demonstrates strong performance in providing users with accurate information about the structural aspects of the ontology, as shown in Table 1. However, it requires further improvements to better interpret user queries. A detailed evaluation of its functionality across the specific tasks outlined

earlier highlights these strengths and areas for improvement. Future evaluations incorporating refined metrics and a larger test set will enable a more comprehensive assessment of the system’s capabilities.

Task	Query Accuracy (%)	Response Accuracy (%)	Validation Method	Key Observations
Retrieving definitions	80	100	Ontology-based	Issues with semantically similar classes
Exploring relationships	70	100	Ontology-based	Accurate for one-hop, challenges with multi-hop
Understanding model structure	100	100	Ontology-based	Strong performance due to well-structured ontology
Investigating related classes	80	100	Ontology-based	Misinterpretation of user queries caused errors
Summarizing classes for topics	N/A	N/A	Domain expert evaluation	Inconsistent lists across query iterations

Table 1: Summary of Query and Response Accuracy Metrics.

5 CONCLUSIONS

This paper presents a hybrid methodology combining classical and AI-based approaches to improve the explainability, usability, and accessibility of complex ontologies, as demonstrated through the Digital Reference model for semiconductor supply chains. Two tools were developed: OntoLearn, which provides interactive visualization for exploring and refining ontology classes, and an AI-based querying tool that allows natural language interaction without requiring expertise in query languages like Cypher.

OntoLearn supports user-friendly visualization and collaborative ontology refinement, while the AI-based tool incorporates query syntax validation, schema verification, and fallback templates to ensure semantic accuracy and robust responses. These mechanisms enhance accessibility for non-technical users and improve system reliability. Together, these tools enable tasks such as retrieving ontology definitions, exploring relationships, and querying ontology structures using clear, natural language.

While effective, the system faces limitations, including occasional paraphrasing inaccuracies, computational inefficiencies, and difficulties in interpreting user queries. Addressing these challenges will be key to optimizing real-time industrial deployment.

The methodology’s scalability makes it applicable beyond semiconductors to domains such as manufacturing and logistics. Further enhancements, including customizable OntoLearn sessions, Retrieval-Augmented Generation (RAG) for richer contextual responses, and expanded ontology visualizations, could improve usability and broaden its applicability in real-world systems.

REFERENCES

Amara, J., B. König-Ries, and S. Samuel. 2024. “Enhancing Explainability in Multimodal Large Language Models Using Ontological Context”. *arXiv preprint arXiv:2409.18753* <https://doi.org/10.48550/arXiv.2409.18753>.

Baldazzi, T., L. Bellomarini, S. Ceri, A. Colombo, A. Gentili, E. Sallinger *et al.* 2024. “Explaining Enterprise Knowledge Graphs with Large Language Models and Ontological Reasoning”. In *The Provenance of Elegance in Computation-Essays Dedicated to Val Tannen (2024)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Berners-Lee, T., D. Connolly, L. Kagal, Y. Scharf, and J. Hendler. 2008. “N3Logic: A Logical Framework for the World Wide Web”. *Theory and Practice of Logic Programming* 8(3):249–269.

Berners-Lee, T., J. Hendler, and O. Lassila. 2001. “A New Form of Web Content That Is Meaningful to Computers Will Unleash a Revolution of New Possibilities”. *Scientific American* 284(5):34–43.

D’Abramo, J., A. Zugarini, and P. Torroni. 2025, May. “Investigating Large Language Models for Text-to-SPARQL Generation”. In *Proceedings of the 4th International Workshop on Knowledge-Augmented Methods for Natural Language Processing*, edited by W. Shi, W. Yu, A. Asai, M. Jiang, G. Durrett, H. Hajishirzi, and L. Zettlemoyer, 66–80. Albuquerque, New Mexico, USA: Association for Computational Linguistics <https://doi.org/10.18653/v1/2025.knowledgenlp-1.5>.

de Sousa Ribeiro, M., and J. Leite. 2021. “Aligning Artificial Neural Networks and Ontologies Towards Explainable AI”. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Volume 35, 4932–4940.

Dimitrakopoulos, G., H. Ehm, and E. Tsaousi. 2024. “Enhanced Ontology Extraction: Integrating GPT AI with Human Knowledge on the Example of EU Standards Related to Semiconductor Supply Chains”. In *2024 Winter Simulation Conference (WSC)*, 1955–1965. IEEE <https://doi.org/10.1109/WSC63780.2024.10838760>.

DR. “Digital Reference Website”. Accessed 4th April 2025. Available at: <https://ifx-dr.github.io/DigitalReference/>.

Guizzardi, G., and N. Guarino. 2024. “Explanation, Semantics, and Ontology”. *Data & Knowledge Engineering* 153:102325 <https://doi.org/10.1016/j.datak.2024.102325>.

Herding, R., L. Mönch, and H. Ehm. 2021. “Design and Application of an Ontology for Demand Fulfillment in Semiconductor Supply Chains”. In *2021 Winter Simulation Conference (WSC)*, 1–12. IEEE <https://doi.org/10.1109/WSC52266.2021.9715435>.

Naqvi, M. R., L. Elmhadhbi, A. Sarkar, B. Archimede, and M. H. Karray. 2024. “Survey on Ontology-Based Explainable AI in Manufacturing”. *Journal of Intelligent Manufacturing*:1–23 <https://doi.org/10.1007/s10845-023-02304-z>.

OntoLearn 2025. “OntoLearn Tool”. Accessed 4th April 2025. Available at: <https://learningtool.opaix.com/>.

Potdar, S., D. Lee, O. Attia, V. Embar, D. Meng, R. Balaji, *et al.* 2025. “Comprehensive Evaluation for a Large Scale Knowledge Graph Question Answering Service”. *arXiv preprint* <https://doi.org/10.48550/arXiv.2501.17270>.

Rahayu, N. W., R. Ferdiana, and S. S. Kusumawardani. 2022. “A Systematic Review of Ontology Use in E-Learning Recommender System”. *Computers and Education: Artificial Intelligence* 3:100047 <https://doi.org/10.1016/j.caai.2022.100047>.

Ramzy, N. 2022. *Semantic Data Integration for Supply Chain Management: With a Specific Focus on Applications in the Semiconductor Industry*. Ph. D. thesis, Leibniz Universität Hannover <https://doi.org/10.15488/13083>.

Ramzy, N., H. Ehm, V. Wiens, and L. Kohnen. 2021. “The Digital Reference: Semantically Connecting Semiconductor Supply Chains to Customers—The Open Online Sales and Marketing Vision”. In *Proceedings of the CASE Conference, Lyon, France*, 23–27.

Sadruddin, S., J. D’Souza, E. Poupaki, A. Watkins, H. Babaei Giglou, A. Rula, *et al.* 2025. “LLMs4SchemaDiscovery: A Human-in-the-Loop Workflow for Scientific Schema Mining with Large Language Models”. In *2025 Winter Simulation Conference (WSC)*, edited by E. Curry, M. Acosta, M. Poveda-Villalón, M. van Erp, A. Ojo, K. Hose, C. Shimizu, and P. Lisena, 244–261. Cham: Springer Nature Switzerland https://doi.org/10.1007/978-3-031-94578-6_14.

AUTHOR BIOGRAPHIES

MARTA BONIK is a Ph.D. candidate at Infineon Technologies AG, researching language technologies and the Semantic Web for supply chain knowledge management. She holds degrees in Philology and Computational Linguistics and focuses on large language models, generative AI, and industrial knowledge graphs to enhance semiconductor supply chains. Her email address is marta.bonik@infineon.com.

ELENI TSAOUSI is a research associate at Harokopio University of Athens and a Master’s student in Data Science at the University of Zurich. She interned at the Athena Research Center and Infineon Technologies AG, working on NLP, chatbots, and ontology development for supply chains. Her research interests include ontologies for industrial applications and AI-driven semantic data integration. Her email address is etsaousi@hua.gr.

HANS EHM is a Senior Principal Engineer Supply Chain at Infineon Technologies AG with nearly 40 years of experience in the semiconductor industry. He has held leadership roles across manufacturing, assembly, and global supply chains and has led multiple international projects, including SC³ (Semantically Connected Semiconductor Supply Chains). His research focuses on semantic web, AI, and quantum computing in supply chain systems. His email address is hans.ehm@infineon.com.

GEORGE DIMITRAKOPoulos is an associate professor at Harokopio University of Athens and a research coordinator at Infineon Technologies AG. He holds a Ph.D. in Electrical and Computer Engineering and is the author of three books and over 210 publications. His research includes cognitive algorithms for vehicular communications and wireless network optimization. He is ranked among the top 2% of scientists worldwide (Stanford, 2021–2023). His email address is gdimitra@hua.gr.