Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

ENHANCED ONTOLOGY EXTRACTION: INTEGRATING GPT AI WITH HUMAN KNOWLEDGE ON THE EXAMPLE OF EU STANDARDS RELATED TO SEMICONDUCTOR SUPPLY CHAINS

George Dimitrakopoulos[1], Hans Ehm[2], and Eleni Tsaousi[1]

[1]Dept. of Informatics and Telematics, Harokopio University, Athens, GREECE
[2]Dept. of Supply Chain Innovations, Infineon Technologies AG, Munich, GERMANY

ABSTRACT

This paper addresses challenges in creating ontologies for the semiconductor supply chain. Ontologies are crucial for seamless data exchange within the semantic web, enabling initiatives like GAIA-X and CatenA-X. Traditionally, ontology creation is complex. Here, we propose a novel AI-assisted method using large language models (LLMs) like ChatGPT 4 Turbo to support human experts. This collaboration aims to expedite ontology generation while maintaining quality. While initial tests show promise, refining the human-AI interface for clear content generation remains a focus. By improving this collaboration, we expect to create more accurate and complete ontologies, fostering efficient information sharing and strengthening the meaningfulness of standards within the semiconductor supply chain.

1 INTRODUCTION

As the semiconductor industry moves toward greater technological sophistication, the role of ontologies in the Semantic Web has become critical in managing the complex information that supports supply chain collaboration. The work of Yun et al. (2022) showed how building knowledge graphs from these ontologies can optimize searchability and retrieval of product details within supply chain. These ontological constructs, rich in classes, object properties, and data properties, are not merely theoretical entities but serve as crucial facilitators for enhancing semantic capabilities that align with the nuanced demands of semiconductor products. At the same time, Ehm et al. (2019) identified the adaptability of product structure representations as essential to meet the demands of a market characterized by rapid innovation and change. In line with Ramzy et al.(2022) and the GAIA-X and CatenA-X initiatives, which strive for a federated and secure data infrastructure in Europe, our study integrates these concepts into the semiconductor field. By aligning semantic models with a robust Disruption Management Process (DMP), as highlighted by recent global events, we can predict and mitigate the effects of supply chain disruptions more effectively. Such disruptions, which extend beyond individual entities to entire interconnected networks, highlight the need for a holistic and integrated approach to managing the performance of supply chains.

Based on these insights, our study presents an innovative approach that leverages the capabilities of Large Language Models (LLMs), particularly ChatGPT 4 Turbo, for ontology generation. LLMs, known for their excellent ability to digest and generate human-like text, offer a powerful means of processing and structuring the large-scale documents inherent in semiconductor manufacturing (Bombieri et al. 2024). By engaging these models in a specific context of detailed queries, we aim to extract accurate ontological information from extensive datasets (Li et al. 2024). Our initial efforts using ChatGPT 4 Turbo to produce GPT ontologies produced promising results, demonstrating a deeper understanding of the language that closely aligns with the evolving terminologies required by the industry.

Although we have taken a significant step forward, we have also faced some challenges. In building GPT ontologies that are both comprehensive and accurate, we encountered problems with the logical consistency and ambiguity of AI-generated content. Identifying and acknowledging these challenges is a critical step, as it points to the need to improve the interaction framework between human operators and ChatGPT 4 Turbo. Addressing these obstacles is necessary to ensure the accuracy of ontologies and their effective contribution to the Semantic Web.

This paper aims to describe our method, which demonstrates its effectiveness and potential to contribute significantly to future research in this area. In addition, we aim to address the issues that other researchers have faced, thereby improving the interaction framework between humans and ChatGPT 4 Turbo. In this way, we strive to enhance the credibility of ontologies and their impact on the Semantic Web, particularly in the context of Supply Chain Collaboration (SCC) and the semiconductor industry's growing reliance on Digital Reference (DR) content. Essentially, our research attempts to bridge the theoretical frameworks set by Yun et al.(2022) and Ehm et al.(2020) with practical innovations based on artificial intelligence, opening a new path for ontological development that could redefine the operational dynamics of semiconductor supply chains.

This research paper is organized into four primary sections, each aimed at systematically exploring and expanding upon the intersection of artificial intelligence and ontology generation within the semiconductor industry's supply chain dynamics. Section 1, the Introduction, lays the foundation of the study by highlighting the importance of ontologies in the Semantic Web and setting the context for the integration of ChatGPT 4 Turbo and human expertise. Section 2, Related Work, delves into the existing literature and studies to position our methodology within the broader academic dialogue and to acknowledge the progress and pitfalls encountered in the field. Section 3, Methodology, offers an in-depth look at the novel approach employed in this study, detailing the collaborative process and the steps undertaken to leverage both AI and human judgment in creating ontologies. Lastly, Sections 4 and 5, Results and Conclusions, present the outcomes of the applied methodology and discusses the broader implications, potential improvements, and future directions for the research. The structure supports a clear narrative flow, guiding readers from the conceptual framework through the practical application, leading to conclusions that underscore the paper's contributions to the field.

2 RELATED WORK

In recent studies, the integration of ontologies and knowledge graphs (KGs) within the Semantic Web has garnered considerable attention, particularly for its potential to streamline complex supply chain activities in the semiconductor industry. This heightened interest has led researchers to harness the capabilities of Large Language Models (LLMs) for the development of ontologies, a process foundational to the construction and optimization of KGs (K. Kommineni et al. 2024). These ontologies act as structured repositories of knowledge, pivotal in formulating competency questions (CQs) that guide the systematic assembly of information, thereby enhancing the accessibility and searchability of intricate product details. Despite the initial promise shown by these efforts, the application of LLMs has surfaced challenges, such as the generation of irrelevant KG content, underscoring the nascent stage of this technology (He et al. 2023). Additional complexities arise from the multilingual output generated by LLMs, which, while demonstrating the models' versatility, also highlight the necessity for meticulous human oversight to ensure relevance and accuracy in the context of specific industrial applications (He et al. 2023). Innovations like OntoFact and ontology-based reinforcement learning (ORL) engines have been explored to mitigate the scattered distribution of facts within the knowledge processed by LLMs (Shang et al. 2024). These pioneering frameworks have marked significant advancements in ontology creation.

Nonetheless, they also reveal an ongoing need to fortify the dependability of LLMs in producing content that is not only accurate but contextually attuned to the subtleties of domain-specific knowledge (Shang et al. 2024). A consistent challenge observed across various studies is the LLMs' intermittent understanding of ontological concepts, coupled with inconsistencies in memorizing and applying these concepts within the appropriate context (Luo et al. 2024). This issue is fundamental to the ongoing debate over LLMs' effectiveness in complex domains, such as the semiconductor supply chain, where precision in information is paramount. Methodologies like ANGEL have attempted to capitalize on the hierarchical nature of ontologies by pairing them with the generative abilities of LLMs. Despite the potential of such integrations, difficulties in identifying synonyms accurately and providing suitable concept labels have persisted, indicating that even advanced AI models like LLMs require further tuning to address these problems (K. Kommineni et al. 2024; Neuhaus 2023). Furthermore, research focusing on GPT-3.5 and Flan-T5 within the field of ontologies has emphasized the crucial role of fine-tuning the interplay between human expertise and LLMs. Such fine-tuning is essential for effective ontology construction, underscoring the need for a collaborative approach where human knowledge and AI capabilities complement each other (Neuhaus 2023). Moreover, studies have shown that ChatGPT can facilitate the creation of KGs by adhering to detailed instructions. These findings suggest potential methods to improve the precision of AI-generated content and prevent unlikely or erroneous alignments (Trajanoska et al. 2023).

Yet, these approaches have not fully achieved the dynamic ontology creation needed to match the industry's rapidly evolving vocabularies or to enhance information sharing within supply chains effectively. To bridge these gaps, our study presents a new methodology that employs ChatGPT 4 Turbo to produce ontologies tailored to the semiconductor industry. By improving the interaction between human operators and ChatGPT 4 Turbo, we aim to boost the reliability of ontologies, thereby positively impacting supply chain collaboration. Our approach addresses the pressing challenges of ensuring accuracy and the dynamic generation of ontologies, with the goal of driving innovation and efficiency within supply chain operations. In conclusion, our method aims to overcome the prevailing limitations by cultivating a productive collaboration between human cognitive abilities and the computing power of ChatGPT 4 Turbo. In doing so, we aim to enhance the robustness and applicability of the Semantic Web, particularly in the era of Supply Chain Collaboration (SCC) and the semiconductor industry's growing reliance on Digital Reference (DR) content. This paper will clarify our approach and its implications for the future of ontology development, potentially redefining the operational dynamics of semiconductor supply chains.

3 METHODOLOGY

To improve supply chain collaboration and address existing challenges in ontology creation in the semiconductor industry, we describe a detailed methodology designed to effectively integrate ontologies into practical applications.

Our research approach is based on leveraging the collaboration between artificial intelligence, especially ChatGPT 4 Turbo, and human expertise. By leveraging the strengths of each, we aim to optimize the ontology creation process. AI excels at processing massive amounts of data and creating content at scale, while human input brings domain knowledge, understanding of context, and discerning judgment to ensure the accuracy and relevance of ontological structures.

This dynamic interaction enables the extraction of comprehensive and contextually accurate information from complex data sets, while ensuring that the ontologies produced are closely aligned with industry-specific terminologies and evolving requirements. Human understanding helps guide AI in creating content that is not only grammatically correct but also semantically meaningful and contextually relevant. Furthermore, this collaborative approach promotes the continuous improvement and adaptation of ontologies to evolving vocabularies and industry requirements. By facilitating effective communication and collaboration between AI and human operators, our methodology improves the ontology creation process, leading to more efficient and effective results.

Additionally, the methodology for creating the classes and subclasses can be replicated by different users when running the same document. The structured framework and detailed steps outlined in our methodology—including initial document evaluation, identifying frequent terms, text segmentation, iterative feedback, and standard integration—ensure that other users can reproduce similar ontological structures, maintaining consistency and replicability in the creation of classes and subclasses.

Below, we delve into the specifics of our framework and the steps we followed to implement this collaborative process. We explain the importance of a structured framework for successful ontology creation in the semiconductor industry, highlighting how each phase leverages AI's advantages while incorporating human insights. Figure 1 presents a visual representation of our framework, illustrating the sequence of steps and their interdependencies, showing how AI and human collaboration produce robust ontologies.

Moving from general strategies to practical actions, we detail the specific steps of our methodology, as illustrated in Figure 1. The use of 300 pages for text segmentation was due to AI input size limitations. To ensure comprehensive analysis, texts exceeding 300 pages were segmented into smaller parts, facilitating iterative integration and maintaining coherence. Licensing considerations also influenced this choice, ensuring compliance and optimizing ChatGPT 4 Turbo's processing capability.

Each phase is described in detail, depicting a process that combines the analytical capability of ChatGPT 4 Turbo with the differentiated expertise of human professionals. This dual approach enables us to calibrate the ontology development process in detail - starting from an initial document evaluation to the complex mapping of relationships within the final knowledge graph.

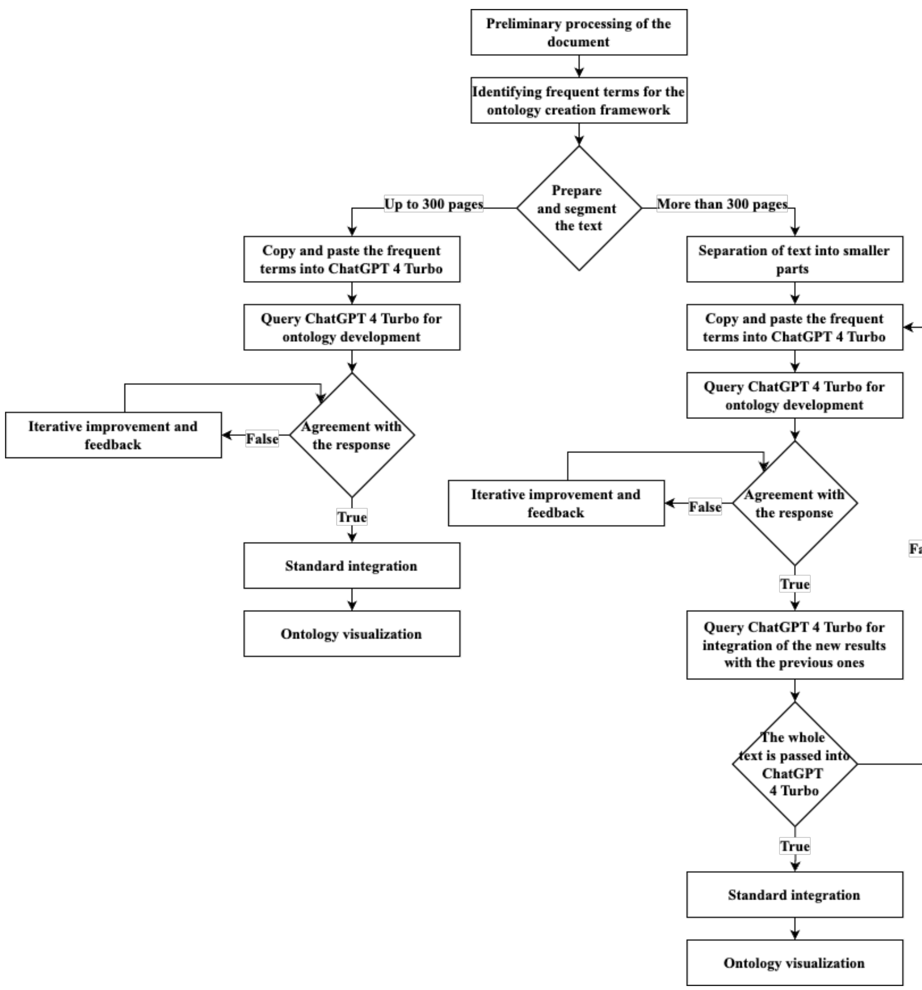

Figure 1: KG creation framework using ChatGPT 4 Turbo.

Step 1: Preliminary processing of the document:

Before dealing with the AI, it was important to familiarize ourselves with the content and context of the text. We skimmed the document to understand the general themes and key terminologies. This initial step helped us evaluate the ontology that ChatGPT 4 Turbo would create, ensuring that we could effectively judge its accuracy and completeness.

Step 2: Identifying frequent terms for the ontology creation framework:
Having established a general understanding, we proceeded to identify frequently occurring terms in the text. These terms, which likely represent core concepts, gave us a glimpse into possible classes and subclasses within the ontology. This proactive identification ensured that important terms were not neglected by the AI during the ontology creation process.

Step 3: Prepare and segment the text:
Document text prepared for import into ChatGPT 4 Turbo. Given the limitations of AI on input size, we segmented texts exceeding 300 pages into smaller, digestible parts. This ensured that every part of the document was analyzed in detail without being limited by the AI's processing capability.

Starting with Step 4, we engage in what we call " Human-in-the-Loop Ontology Extraction for EU Semiconductor Standards with GPT-AI". This pivotal phase of our methodology highlights the interactive cycle where human experts refine and validate the information processed by ChatGPT 4 Turbo. AI's ability to manipulate data is aligned with human oversight, ensuring that ontologies reflect the latest EU standards and industry practices. This combination of human acumen and advanced AI analytics underpins our approach to tackling the complexities of semiconductor standardization in the EU.

Step 4: Query ChatGPT 4 Turbo for ontology development:
With the text segmented, we used ChatGPT 4 Turbo through specific and structured queries. For example, we asked the AI to "identify all classes, subclasses, and discover their relationships, indicating the direction and nature of each relationship using a variety of descriptive verbs, not just simple verbs like include." Such a detailed prompt led the AI to build a richly detailed ontology, capturing a wide range of relationships within the text.

For documents exceeding the 300-page limit, we adopted an iterative approach. Each time ChatGPT 4 Turbo analyzed a section and identified classes and subclasses, we requested the integration of these new findings with the ontology developed from the previous sections. This iterative integration was instrumental in maintaining the continuity and coherence of the ontology throughout the body of the document.

Step 5: Iterative improvement and feedback:
When AI results did not meet standards of clarity and relevance, we provided detailed feedback. For example, we would highlight unclear points or possible inaccuracies, such as "This connection between concepts is not clear or may not be directly related. Review and improve." This feedback was very important in guiding the AI, allowing it to recalibrate its responses in subsequent iterations. Moreover, we tried to supply ChatGPT with more examples that model the type of logical relationships we expected to see. Another method was to simplify our prompts, breaking down more complex requests into smaller, more targeted queries that the AI could process more effectively.

Step 6: Standard integration:
After evaluating the initial AI-generated ontology for robustness and alignment with the document content, we instructed ChatGPT 4 Turbo to incorporate specific industry standards into the structure. ChatGPT 4 Turbo's ability to understand and implement multifaceted standards allowed us to refine the ontology to be industry-compliant and semantically robust.

Step 7: Ontology Visualization:
Using visualization tools like draw.io, we translated the complex relationships and structures identified by AI into a comprehensive knowledge graph. This visualization acted as a tangible representation of the ontology, facilitating an intuitive understanding and further analysis.

This methodological approach was created to take advantage of AI in large-scale data processing, while leveraging human expertise to guide and improve AI results. In this way, we were able to create an ontology that resonates with the specific terminologies and requirements of the semiconductor industry and is also adaptable and up to date with current industry standards.

Of course, we addressed some potential intellectual property (IP) issues related to the use of ChatGPT within semiconductor manufacturing companies. Since ChatGPT 4 Turbo was utilized internally at Infineon, it was reassuring to know that our data remained secure within our own infrastructure. This internal deployment meant that proprietary semiconductor designs and processes were handled in a controlled environment, reducing the risk of sensitive information being inadvertently exposed. By addressing these potential IP issues, we ensured that the integration of ChatGPT into semiconductor manufacturing workflows adhered to legal and ethical standards, safeguarding both the company's proprietary information and the integrity of the AI-generated content.

4 RESULTS

In our study, we used the methodology described to create knowledge graphs for the Union's Annual Work Programme for European standardization 2024, with documents of considerable size and complexity. Given the sheer scale of these documents and the intricate details they contain, manual creation of ontologies would be time-consuming and prone to human error and oversight. The methodology we described earlier was tested with these extensive documents. This approach was chosen to bypass the tedious task of manually parsing the content and leverage the speed and pattern recognition capabilities of ChatGPT 4 Turbo.

Our first step involved a review of the entire document. This was not just a cursory glance, but a strategic overview to familiarize ourselves with the key themes and terminology of the text. This fundamental understanding was crucial to critically evaluate the accuracy of the AI-generated output in the next steps.

After highlighting frequently occurring terms that could indicate critical concepts within the text, we parsed the documents into manageable parts for ChatGPT 4 Turbo and submitted specific, structured messages to the AI. The prompts were created to get the AI to reveal not only the classes and subclasses, but also the complex web of their relationships, ensuring a comprehensive understanding of the document structure.

When the AI delivered its initial set of results, our domain expertise allowed us to critically evaluate the result. Where clarity was lacking or connections seemed weak, we provided specific feedback in ChatGPT 4 Turbo. We flagged ambiguous sections or potential inaccuracies, asking the AI to review and improve its answers. This feedback loop is crucial as it serves as a fine-tuning mechanism, allowing us to iteratively improve AI performance. It is through this iterative process that we tease out a refined, accurate ontology from the raw AI output.

For example, during the creation of the knowledge graph for the green electricity system, part of the Union's Annual Work Programme, we applied our methodology to ensure that the graph accurately reflects the complex network described in the text — of interconnected networks that ensure supply of high-quality electricity to integrate renewable energy sources. The resulting knowledge graph, which is depicted below in Figure 2, is a visual synthesis of the interdependencies and nuances of infrastructure critical to Europe's electricity networks.

Specifically, Figure 2 is the knowledge graph of Green Electricity System created in draw.io. This visual mapping is particularly useful for confirming the consistency and accuracy of the relationships described by ChatGPT 4 Turbo. It provides a clear, graphical representation of the ontology, which is particularly beneficial when dealing with large-scale and complex systems such as the green electricity grid.

In this knowledge graph, we used a color-coded system to improve clarity and distinguish between various elements within the ontology:

Main classes are highlighted in bright light blue color. ChatGPT 4 Turbo identified these main classes by recognizing recurrent, high-level concepts critical to the green electricity system. These include overarching categories like Regulatory Frameworks, Technological Innovations, Electricity Grids, Interconnectivity, Investment and Financing, Financial Instruments, Supply Chain Development, Stakeholder Engagement, Cybersecurity, and Regulatory Authorities. The main classes are foundational elements that define the structure and operation of the green electricity system.

Subclasses and properties are multicolored to categorize them into subclasses of the classes, making the hierarchy and classification immediately visible. ChatGPT 4 Turbo employed algorithms such as clustering, dependency parsing, and co-occurrence analysis to discover these subclasses and properties. By analyzing the context in which terms appear together, the AI was able to discern finer-grained categories within the main classes.

- Subclasses/properties of Regulatory Frameworks are highlighted in light purple. These include the rules and guidelines that govern the operation and development of the green electricity system, ensuring compliance with legal and safety standards.
- Subclasses/properties of Technological Innovations and Standards are highlighted in yellow. These focus on the latest technological advancements and the establishment of standards for manufacturing and operational processes in the green electricity sector.
- Subclasses/properties of Electricity Grids are highlighted in light green. These relate to the infrastructure and management of electricity distribution and transmission networks, essential for the integration of green energy sources.
- Subclass/property of Interconnectivity and Energy Transition is highlighted in grey. This subclass deals with the modernization and enhancement of grid connectivity to support the transition to greener energy solutions.
- Subclass/property of Investment and Financing Needs is highlighted in light pink. This includes the financial requirements and funding sources necessary for developing and sustaining the green electricity system.
- Subclasses/properties of Financial Instruments are highlighted in beige. These outline various financial mechanisms and instruments that provide funding for green electricity projects and infrastructure.
- Subclass/property of Supply Chain and Workforce Development is highlighted in purple. This subclass focuses on developing the supply chain and workforce necessary to support the green electricity system.
- Subclass/property of Stakeholder Engagement and Public Permits is highlighted in pink. This involves engaging stakeholders and managing the permitting processes required for the implementation of green electricity projects.
- Subclass/property of Cybersecurity and Resilience is highlighted in light blue. This ensures the security and resilience of the green electricity infrastructure against cyber threats and risks.
- Subclasses/properties of Regulatory Authorities and Entities are highlighted in green. These include the organizations and bodies responsible for regulating and overseeing the green electricity system.
- Subclass/property of Network Development and Planning is highlighted in salmon color. This covers the strategic planning and development of electricity networks to meet future energy demands.
- Subclasses/properties of Infrastructure Projects are highlighted in lime light. These are specific projects aimed at building and upgrading the infrastructure necessary for the green electricity system.

Subclasses related to digital reporting are embedded in red to highlight their relevance and specificity in the chart. This integration demonstrates ChatGPT 4 Turbo's ability to understand and implement hierarchical relationships within a given domain. Digital reporting includes tools and systems for data collection, analysis, and dissemination, which are crucial for monitoring and managing the green electricity system.

The "standards" class is depicted in green. Standards are essential for ensuring consistency, safety, and interoperability within the green electricity system. The effectiveness of ChatGPT 4 Turbo in detecting and incorporating patterns suggests that it leverages its extensive training, which likely includes exposure to a variety of pattern-related documents and data. ChatGPT 4 Turbo uses machine learning models trained on large corpora of text to understand linguistic patterns and relationships. In this way, it can identify patterns and their relative context in the field literature, allowing it to accurately place those patterns in the knowledge graph relative to other classes and subclasses.

The "Infineon Technologies (IFX) products" class is color-coded in blue. IFX products are integrated into the graph because of their critical role in providing semiconductor technologies essential for various applications within the green electricity system. AI's understanding of the concept and application of IFX products is likely attributed to its underlying algorithms, which process massive data sets during training, allowing it to recognize and make connections based on product features, usage, and industry context.

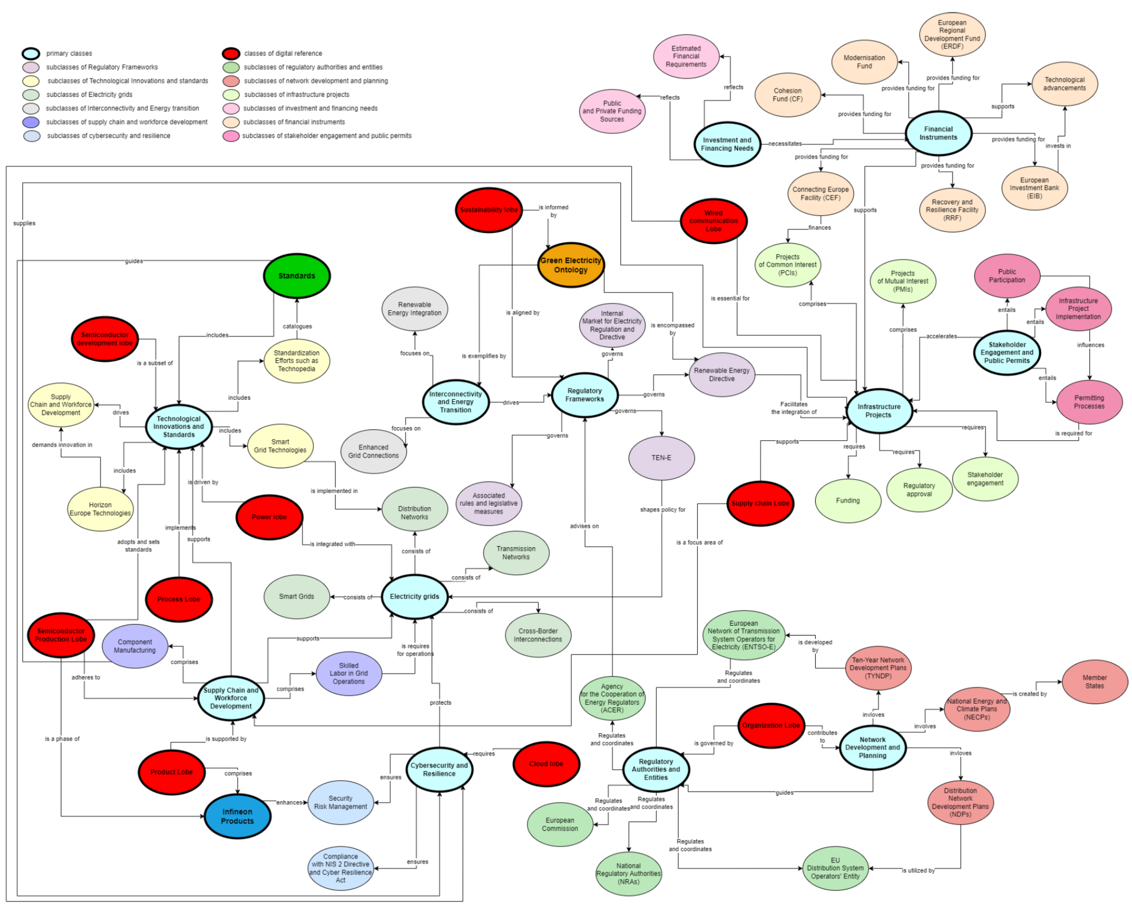

Figure 2: KG of green electricity system.

By using this approach, we created several KGs for the Union's Annual Work Programme for European standardization 2024, such as the knowledge graph of Artificial Intelligence and the knowledge graph of Critical Row Materials which are represented in Figures 3 and 4 , respectively.

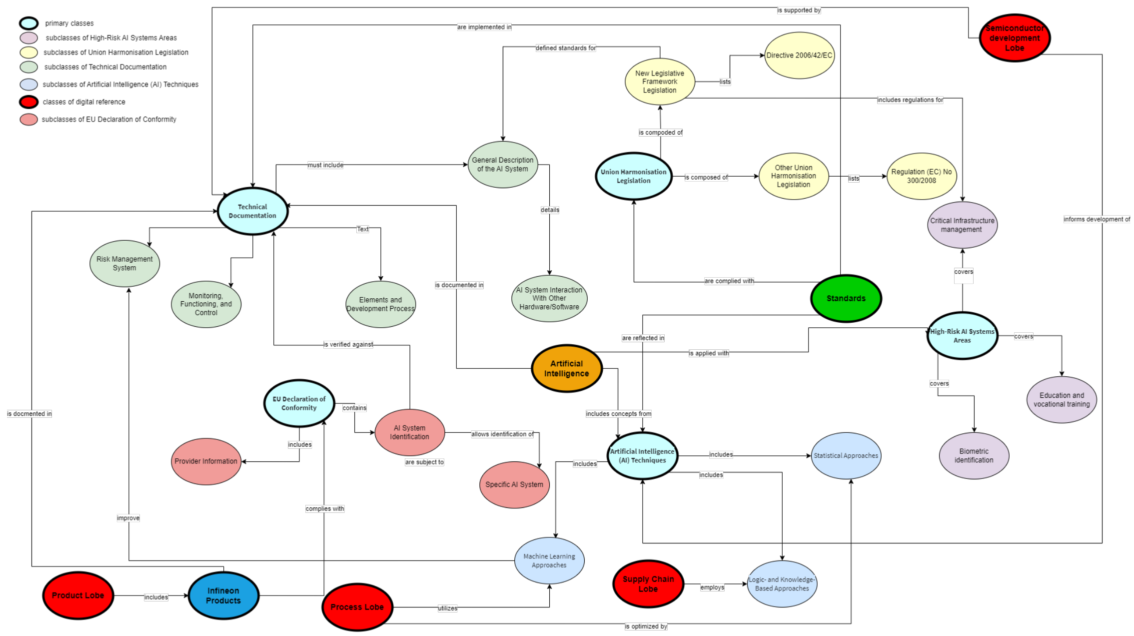

Figure 3: KG of artificial intelligence.

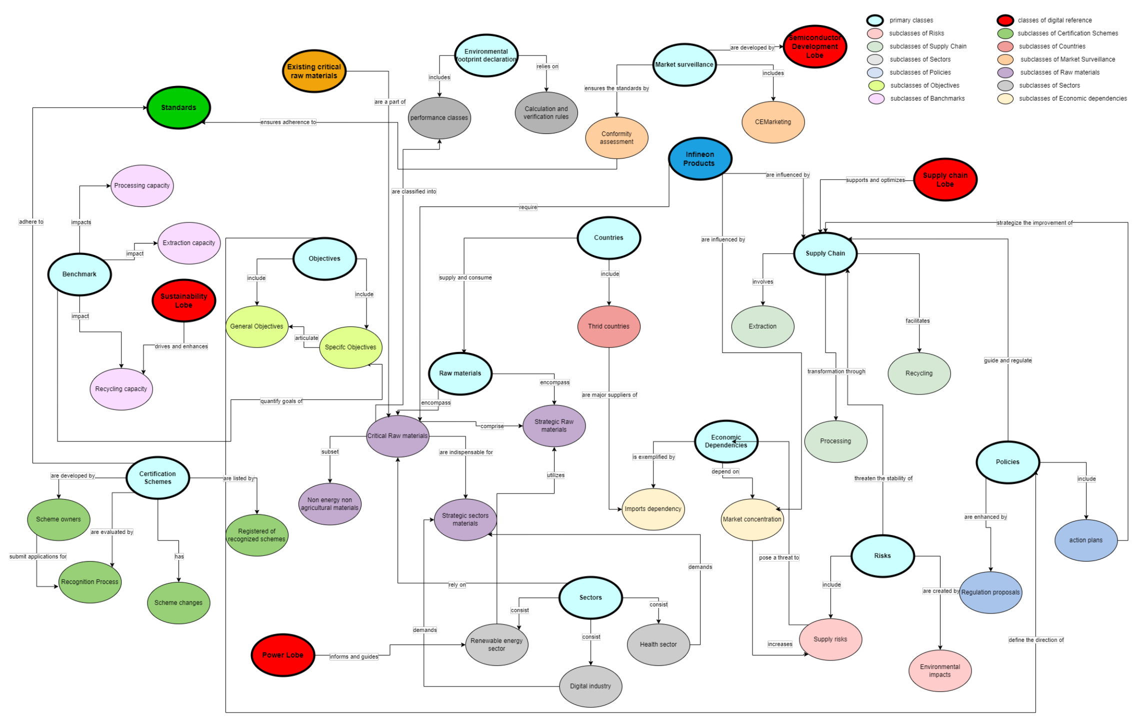

Figure 4: KG of critical raw materials.

However, sometimes the result of ChatGPT was ambiguous. For example, sometimes ChatGPT's responses were a bit confusing. When this happened, we applied some of the strategies mentioned earlier and this usually helped to give us a better answer. In this way, we achieved a higher degree of accuracy in AI-generated content.

It is made clear that the research for full logical consistency remains ongoing. The challenges we faced highlight the necessity of human involvement in reviewing and correcting AI-generated data, reflecting the current state of AI technology where human-AI collaboration is essential to achieve the best results.

5 CONCLUSIONS

This paper introduced a joint effort between the algorithms of ChatGPT 4 Turbo and humans to enhance the creation of ontologies within the semiconductor industry's supply chain. The collaborative methodology showed significant improvement in the speed and reliability of ontology construction and ensured that the generated knowledge graphs remained aligned with industry standards, thus facilitating more efficient information sharing mechanisms.

Our research confirmed that human expertise is important in guiding artificial intelligence and building complex knowledge graphs and ontologies, ensuring that technological developments remain realistic and workable. Humans offer benefits that AI cannot—such as our ability to understand complex contexts and make judgments based on experience. These uniquely human characteristics ensure that ontologies are not only comprehensive but also meaningful and practically applicable in the real world.

Finally, several avenues of research emerge from our work. Future studies may delve into the development of more sophisticated interaction models between human experts and artificial intelligence, which could lead to even more accurate and usable ontological structures. Exploring the application of our methodology in different complex domains may reveal broader implications and benefits, contributing to a generalizable framework for AI-assisted ontology creation. Still, the fine-tuning of LLMs for specific domain ontologies can be further explored to improve the accuracy and relevance of AI-generated content. In addition, the integration of emerging technologies such as blockchain could be explored to ascertain how they could improve data integrity and ontology accuracy in complex supply chain environments. Further research could address ethical issues, ensuring the responsible use of AI in knowledge management and supporting the transparency of AI-generated ontological content.

REFERENCES

Bombieri, M., P. Fiorini, S. P. Ponzetto, and M. Rospocher. 2024. "Do LLMs Dream of Ontologies?". *arXiv preprint arXiv:2401.14931*.

Ehm, H., N. Ramzy, P. Moder, C. Summerer, S. Fetz and C. Neau. 2019. "Digital Reference–a Semantic Web for Semiconductor manufacturing and Supply Chains Containing Semiconductors". In *2019 Winter Simulation Conference (WSC)*, 2409–2418. IEEE https://doi.org/10.1109/WSC40007.2019.9004831.

Ehm, H., E. Schoitsch, J. W. van der Weit, N. Ramzy, L. Luo and D. L. Gruetzner. 2020. "Digital Reference: a Quasi-Standard for Digitalization in the Domain of Semiconductor Supply Chains". In *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, Volume 1, 563–570. IEEE.

He, Y., J. Chen, H. Dong, and I. Horrocks. 2023. "Exploring Large Language Models for Ontology Alignment". *arXiv preprint arXiv:2309.07172*.

K. Kommineni, V., B. König-Ries, and S. Samuel. 2024. "From Human Experts to Machines: An LLM Supported Approach to Ontology and Knowledge Graph Construction". *arXiv e-prints*:arXiv–2403.

Li, N., T. Bailleux, Z. Bouraoui, and S. Schockaert. 2024. "Ontology Completion with Natural Language Inference and Concept Embeddings: An Analysis". *arXiv preprint arXiv:2403.17216*.

Luo, Z., Y. Wang, W. Ke, R. Qi, Y. Guo and P. Wang. 2024. "Boosting LLMS with Ontology-Aware Prompt for Ner Data Augmentation". In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 12361–12365. IEEE.

Neuhaus, F. 2023. "Ontologies in the Era of Large Language Models–a Perspective". *Applied Ontology* 18(4):399–407.

Ramzy, N., S. Auer, H. Ehm, and J. Chamanara. 2022. "Mare: Semantic Supply Chain Disruption Management and Resilience Evaluation Framework". *arXiv preprint arXiv:2205.06499*.

Shang, Z., W. Ke, N. Xiu, P. Wang, J. Liu, Y. Li *et al*. 2024. "Ontofact: Unveiling Fantastic Fact-Skeleton of LLMs via Ontology-Driven Reinforcement Learning". In *Proceedings of the AAAI Conference on Artificial Intelligence*, Volume 38, 18934–18943.

Trajanoska, M., R. Stojanov, and D. Trajanov. 2023. "Enhancing Knowledge Graph Construction Using Large Language Models". *arXiv preprint arXiv:2305.04676*.

Yun, T., M. Patrick, R. Nour, and E. Hans. 2022. "Investigating Semantic Web as Enabler for Semiconductor Supply Chain Collaboration".

AUTHOR BIOGRAPHIES

GEORGE DIMITRAKOPOULOS is an associate professor at Harokopio University of Athens (HUA) - Greece, department of Informatics and Telematics (2010-). He is also a coordinator (ext.) at the department of research and development funding at Infineon Technologies A.G. (IFAG) in Germany (2018-), where he manages large scale projects. He has been acting also as an RD consultant for several large enterprises, as well as startups in the area of cognitive communication systems. He received his bachelor degree in Electrical and Computer Engineering in 2002 and his Ph.D. in 2007, from the National Technical University of Athens, and the University of Piraeus, respectively. He is the author of 3 books on vehicular communications and autonomous driving, as well as more than 210 publications in international journals and conferences. In the last 20 years he has been involved in numerous internationally funded RD projects. He is listed among the top 2% of scientists worldwide according to Stanford University rankings (2021, 2022, 2023). His research interests include the design and development of cognitive (knowledge-based) management algorithms for vehicular communications, as well as the design and development of strategies for the optimization of wireless networks based on cognitive networking principles. His email address is gdimitra@hua.gr.

HANS EHM is Senior Principal Engineer Supply Chain of Infineon Technologies AG. He holds degrees in Physics from Germany and a M.S./OSU. In almost 40 years in the Semiconductor industry, he was granted managing and consulting positions at Wafer Fabrication, at Assembly Test and for the global Supply Chains - on production site, on business unit and on corporate level. He was a Board member of camLine Holding AG, an IT company for supply and quality chains. He led and leads Governmental supported projects on national and international level in the context of IT, Semiconductor Manufacturing and Supply Chains - within such projects he initiated the EU/US MIMAC co-operation on capacity modeling in the late 90s or SC^3 (Semantically Connected SemiConductor Supply Chains). He is a member of INFORMS. His research interest is in semiconductor supply chains and supply chains containing semiconductors. He is using methods like discrete event-, agent based- system dynamics simulation, complexity management, semantic web and ontologies, Machine Learning, AI and Human AI interaction as well as Quantum Annealing and Universal Quantum Computing. His email address is hans.ehm@infineon.com.

ELENI TSAOUSI has received her Bachelor's degree in Informatics and Telematics at Harokopio University of Athens. She is doing her internship in the Semantic Web team at Infineon Technologies AG, in Munich, dealing with ontologies, graphs, and data, while also working as a research associate at Harokopio University on projects related to autonomous driving and data spaces in manufacturing. She has already done another internship at Athena Research Center, creating an AI Chatbot by using Pandas, NumPy, and NLP techniques. Her email address is it219104@hua.gr.