

Iterative AI-Assisted Ontology Engineering: Lessons Learned from Developing the CASTON ontology for Metal Casting, Materials and Manufacturing

He Tan^{1,*}, Ziyu Li², Xuan Liu³ and Anders E.W. Jarfors²

¹Department of Computing, School of Engineering, Jönköping University, Sweden

²Department of Materials and Manufacturing, School of Engineering, Jönköping University, Sweden

³Department of Mechanical Engineering, National University of Singapore, Singapore

Abstract

Recent advances in foundation models are transforming ontology engineering by enabling AI to support different ontology engineering tasks. However, practical guidance on how to integrate foundation models, evolving ontologies, and domain experts into effective ontology engineering workflows remains limited. This paper presents the lessons learned from the development of CASTON, a core ontology for metal casting, materials, and manufacturing. We describe an iterative AI-assisted ontology engineering workflow that combines expert-guided ontology scoping and corpus construction, LLM-assisted knowledge extraction, expert-driven ontology construction and refinement, and ontology-guided knowledge augmentation for fine-tuning domain-specific SLMs. From this work, we identify four practical lessons for iterative AI-assisted ontology engineering. We hope these lessons provide practical guidance for future AI-assisted ontology engineering in materials and manufacturing, and other knowledge-intensive domains.

Keywords

Ontology Engineering, AI-Assisted Ontology Engineering, Human-AI Collaboration, Materials and Manufacturing Ontology

1. Introduction

Ontology provides the semantic foundations for knowledge representation, knowledge sharing, interoperability, and intelligent applications across a wide range of scientific and engineering domains, including the field of metal casting, materials and manufacturing. However, ontology engineering remains a knowledge-intensive process that requires substantial effort in knowledge acquisition, conceptual modelling, ontology implementation, and expert validation [1]. Recent advances in foundation models have significantly influenced ontology engineering and created new opportunities for supporting the ontology development process.

Existing studies have demonstrated the use of foundation models for different ontology engineering tasks, including concept and relation extraction, ontology generation from competency questions, ontology population, and ontology refinement [2, 3, 4]. However, practical experience and design guidance on how foundation models should be integrated into a complete ontology engineering workflow remains limited. In particular, comparatively less attention has been paid to how domain experts, foundation models, and evolving ontologies can complement one another throughout successive stages of ontology engineering, from initial knowledge acquisition to continuous ontology development.

Our recent work on ontology engineering for metal casting, materials, and manufacturing provides a concrete case for exploring this perspective [5, 6, 7]. We first developed CASTON¹, a validated core ontology for the domain, through an AI-assisted workflow that combined expert-guided ontology scoping and corpus construction, Large Language Model (LLM)-assisted knowledge extraction, and expert-driven ontology construction and refinement. Rather than treating the ontology solely as the

SeMatS 2026: The 3rd International Workshop on Semantic Materials Science co-located with the 25th International Semantic Web Conference (ISWC 2026), October 25th, 2026, Bari, Italy

*Corresponding author.



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹<https://inmost-research.org/ontologies/caston/>

outcome of ontology engineering, we subsequently reused the validated core ontology to support ontology-guided knowledge augmentation and the fine-tuning of domain-specific Small Language Models (SLMs) for subsequent ontology development. As new domain knowledge was identified from newly collected literature and refined by domain experts, it was continuously incorporated into the ontology. This iterative process establishes a positive feedback loop in which foundation models support continuous knowledge acquisition, and the evolving ontology improves subsequent AI-assisted ontology engineering.

Reflecting on this work, we argue that the workflow did not consist of isolated applications of LLMs or SLMs to individual ontology engineering tasks. Instead, it evolved into an iterative human–AI collaboration in which domain experts, foundation models, and an evolving ontology each played distinct but complementary roles. The validated core ontology served as the starting point for an evolving ontology that gradually became an active component of the engineering process, supporting subsequent AI-assisted knowledge acquisition and continuously expanding through expert construction refinement and enrichment.

This paper presents this iterative AI-assisted ontology engineering workflow and identifies four practical lessons for future ontology engineering. These lessons concern (i) expert-guided ontology scoping and corpus construction, (ii) collaborative expert refinement and ontology enrichment in complex engineering domains, (iii) the evolving role of ontologies in supporting subsequent ontology development, and (iv) the complementary roles of LLMs and SLMs within iterative ontology engineering. We believe these lessons provide practical guidance for developing AI-assisted ontology engineering workflows in materials and manufacturing, and other knowledge-intensive domains.

2. Iterative AI-Assisted Ontology Engineering

Figure 1 summarizes the iterative ontology engineering workflow developed during the construction of CASTON. The first phase establishes a validated core ontology through expert-guided knowledge scoping, corpus construction, AI-assisted knowledge extraction, and expert-driven ontology engineering. The second phase reuses the ontology as structured semantic knowledge to support ontology-guided knowledge augmentation, domain adaptation of SLMs, and continuous ontology development. Together, these two phases establish an iterative human–AI ontology engineering workflow in which the ontology continuously evolves while simultaneously supporting subsequent knowledge acquisition.

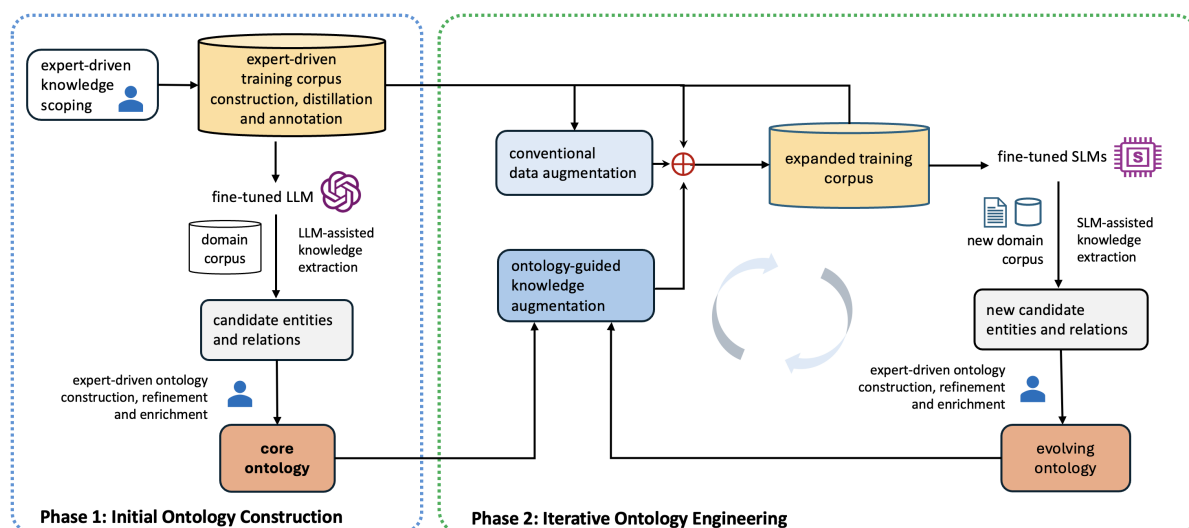


Figure 1: Overview of the Iterative AI-Assisted Ontology Engineering Workflow

Initial Core Ontology Construction

The objective of the first phase was to construct a validated core ontology for the domain of metal casting, materials, and manufacturing [7]. It focused on establishing a stable semantic foundation that captures the fundamental concepts and relationships of the domain. This core ontology subsequently serves as the starting point for continuous ontology development.

The workflow begins with expert-driven knowledge scoping. Domain experts first identified the fundamental knowledge categories that define the conceptual scope of the ontology. Based on these categories, they formulated a small set of targeted domain questions that guided corpus construction. Instead of relying on unrestricted document collection, these questions were used to identify and distill high-quality domain corpus from scientific publications and technical documents for training models and ontology development. This expert-guided corpus construction process ensured that the collected knowledge remained closely aligned with the intended semantic scope of the ontology.

A small training corpus was also constructed and manually annotated for named entity recognition (NER) and relation extraction (RE) by labeling relevant domain-specific entities and the semantic relationships between them. The annotated training corpus was then used to support AI-assisted knowledge acquisition. After fine-tuned on the training corpus, the LLM was used to perform NER and RE on the domain corpus [5]. Rather than generating the ontology directly, the fine-tuned LLM supported the identification of candidate concepts and semantic relationships, which were subsequently used as input for ontology engineering.

The candidate concepts and relations were then transformed into a validated core ontology through expert-driven ontology construction, refinement, and enrichment. Domain experts reviewed the extracted knowledge, resolved conceptual ambiguities, organised the ontology hierarchy, introduced missing concepts and relations, and refined semantic definitions through multiple rounds. The resulting core ontology provides a validated semantic foundation for subsequent ontology engineering.

From Core Ontology to Iterative Ontology Engineering

Once the validated core ontology had been established, its role changed. Instead of serving solely as the outcome of the first development phase, it became structured symbolic domain knowledge that supports subsequent ontology engineering.

The validated ontology was first used to guide ontology-guided knowledge augmentation. Unlike conventional data augmentation, which expands a training dataset through transformations or variations of existing samples, this approach incorporates ontology-based semantic knowledge to construct an enriched training corpus. The expanded corpus was subsequently used to fine-tune domain-specific SLMs, continuously improving their performance in acquiring knowledge from newly collected domain literature [6]. The fine-tuned SLMs extracted new candidate entities and semantic relations, which must again be reviewed, refined, and enriched by domain experts before being incorporated into the ontology.

This process establishes an iterative ontology engineering workflow. As new knowledge is incorporated, the ontology evolves beyond the original core ontology into a progressively richer domain ontology. The evolving ontology simultaneously improves ontology-guided knowledge augmentation and supports subsequent AI-assisted knowledge acquisition, creating a continuous human–AI collaboration in which domain experts, foundation models, and the ontology each play complementary roles throughout successive ontology engineering cycles.

3. Lessons Learned

The development of CASTON and the associated AI-assisted ontology engineering workflow provided an opportunity to reflect on how foundation models, ontologies, and domain experts can be effectively integrated into ontology engineering. Our experience suggests that successful ontology engineering depends on a carefully designed workflow in which human expertise and AI complement each other

throughout the development lifecycle. The following lessons summarize the key observations from our work.

Lesson 1: Domain Experts Should Guide Ontology Scoping and Corpus Construction

Our first lesson is that ontology engineering should begin with expert-guided knowledge scoping rather than unguided information extraction. In our work, domain experts first identified a small set of fundamental knowledge categories that defined the conceptual scope of the ontology. These categories provided the initial semantic framework for ontology development and guided subsequent corpus construction and information extraction. Rather than collecting documents through general keyword searches, the experts formulated a small set of targeted domain questions aligned with these knowledge categories. The questions were subsequently expanded using an LLM. The expanded question set was evaluated by a domain expert and then used within a retrieval-augmented workflow to identify and distill high-quality corpus from scientific publications and industrial technical reports for training foundation models and ontology development.

This question-guided corpus construction process is an important step in developing a high-quality core ontology. Instead of allowing the LLM to explore the domain without constraints, expert knowledge was used to determine what knowledge should be collected before ontology construction began. The resulting corpus provided broad domain coverage while remaining closely aligned with the intended conceptual scope of the ontology.

Lesson 2: Complex Engineering Domains Require Expert Refinement and Enrichment

Although foundation models substantially reduced the effort required for concept and relation extraction, constructing the core ontology remained an expert-driven activity. Metal casting, materials, and manufacturing integrate knowledge from multiple scientific and engineering disciplines, including materials science, metallurgy, manufacturing processes, equipment, defect mechanisms, and process-structure-property relationships. Such knowledge is highly specialised and often distributed across different communities of practice.

An important observation from our work is that expert involvement extended far beyond validation and correction. During the multi-round ontology review, domain experts not only refined the ontology hierarchy and resolved conceptual ambiguities, but also introduced entirely new branches of domain knowledge that were missing from the ontology. In several cases, professor-level experts identified important concepts and conceptual structures that were underrepresented or absent in the ontology because they were insufficiently covered by the domain corpus. These experts extended the ontology by introducing new concepts, semantic relations, and hierarchical structures based on their accumulated scientific and engineering knowledge.

Furthermore, integrating this knowledge required more than transferring domain terminology into the ontology. It involved translating domain-specific understanding, professional conventions, and industrial practices into formal structures suitable for computational representation. Such organising principles are not always explicitly documented in textbooks and research papers, which often describe individual phenomena without fully capturing the classifications, standards, and relationships used by practitioners. Cross-disciplinary collaboration between domain experts and ontology engineers was essential for representing this practical and standards-based knowledge within ontology.

This observation highlights an important limitation of corpus-driven ontology construction. While foundation models can efficiently extract knowledge contained in the corpus, they cannot recover important domain knowledge that is missing, implicit, or insufficiently represented in the available documents. For highly specialised engineering domains, expert-driven ontology construction therefore remains an indispensable component of ontology engineering.

Lesson 3: An Evolving Ontology Becomes an Active Component of Ontology Engineering

Another important lesson concerns the changing role of the ontology itself. In many ontology engineering projects, the ontology is regarded primarily as the outcome of the engineering process. Our experience suggests a different perspective. Once the validated core ontology had been established, it became structured symbolic domain knowledge that actively supported subsequent ontology engineering.

In our workflow, the ontology was reused to guide ontology-guided knowledge augmentation for constructing expanded training corpora used to fine-tune domain-specific SLMs. As newly acquired knowledge was reviewed and incorporated into the ontology, the ontology itself evolved into a progressively richer representation of the application domain. The improved ontology subsequently provided better semantic guidance for later development cycles.

Rather than acting as a static repository of knowledge, the ontology became an active participant in the ontology engineering workflow. Each development cycle improved both the ontology and the AI models that supported subsequent knowledge acquisition, creating a positive feedback loop between symbolic knowledge and AI-assisted learning.

Lesson 4: LLMs and SLMs Play Complementary Roles in the Ontology Engineering Workflow

The fourth important lesson from our work is that LLMs and SLMs should not be viewed as competing technologies, but as complementary components of an iterative ontology engineering workflow. Their strengths differ considerably, making them suitable for different stages of ontology engineering.

In the initial phase of ontology development, LLMs provide broad domain knowledge and strong language understanding and knowledge generalisation capabilities that are particularly effective for expert-guided corpus construction, knowledge extraction, and ontology-guided knowledge augmentation. These capabilities make LLMs well suited for bootstrapping ontology development, where broad coverage and generalization are more important than computational efficiency.

Once a validated core ontology has been established, the ontology itself can be reused to guide knowledge augmentation and generate high-quality training data for domain adaptation. Fine-tuning an SLM using this ontology-guided training corpus enables more efficient knowledge acquisition from newly collected domain literature. Compared with repeatedly relying on LLMs, domain-adapted SLMs provide lower computational cost, greater deployment flexibility, and more consistent extraction behaviour within the target domain [8, 9].

Rather than replacing LLMs, SLMs complement them. LLMs remain valuable throughout the workflow for tasks requiring broad reasoning capabilities and knowledge augmentation, while SLMs provide an efficient solution for repeated domain-specific knowledge acquisition. Together, they support an iterative ontology engineering process in which each model contributes according to its strengths.

4. Conclusion

This paper presented the lessons learned from developing CASTON and the associated iterative AI-assisted ontology engineering workflow for metal casting, materials, and manufacturing. Rather than focusing on individual AI techniques, our experience demonstrates that successful ontology engineering depends on the effective integration of domain experts, foundation models, and evolving ontologies throughout the engineering process.

Our experience suggests four practical principles for future AI-assisted ontology engineering: expert-guided ontology scoping provides a strong foundation for ontology development; constructing high-quality ontologies in specialised engineering domains requires expert refinement and knowledge extension beyond corpus-derived knowledge; validated core ontologies can evolve into valuable semantic

knowledge resources that continuously support subsequent ontology development; and LLMs and SLMs play complementary roles within iterative ontology engineering workflows.

Although these lessons were derived from the development of CASTON, we believe they are applicable to ontology engineering in many other knowledge-intensive domains. Future work will investigate how this workflow can be applied to larger ontology ecosystems and interoperable ontology networks, as well as how evolving ontologies can more effectively support continuous AI-assisted knowledge engineering.

Acknowledgments

This work has been conducted in the research project "Industrial Large Knowledge Model for Smart Manufacturing (INMOST)" (Grant No. 20250058), which is funded by KK-stiftelsen (the Swedish Knowledge Foundation).

Declaration on Generative AI

During the preparation of this work, the authors used GPT-5.5 for refinement and linguistic polishing of human-authored content, including grammar correction, spelling correction, paraphrasing, and rewording. After using the tool, the authors reviewed, edited, and verified the content as needed and take full responsibility for the publication's content.

References

- [1] Y. Sure, S. Staab, R. Studer, Ontology engineering methodology, in: Handbook on ontologies, Springer, 2009, pp. 135–152.
- [2] H. Babaei Giglou, J. D'Souza, S. Auer, LLMs4OL: Large language models for ontology learning, in: International semantic web conference, Springer, 2023, pp. 408–427.
- [3] M. J. Saeedizade, E. Blomqvist, Navigating ontology development with large language models, in: European semantic web conference, Springer, 2024, pp. 143–161.
- [4] A. S. Lippolis, M. J. Saeedizade, R. Keskisärkkä, S. Zuppiroli, M. Ceriani, A. Gangemi, E. Blomqvist, A. G. Nuzzolese, Ontology generation using large language models, in: European Semantic Web Conference, Springer, 2025, pp. 321–341.
- [5] X. Liu, Z. Li, M. He, Z. Ma, X. Wu, G. Yilmaz, Y. Xia, B. Li, H. Tan, J. Y. Hsi Fuh, W. F. Lu, A. E. Jarfors, P. Jansson, From Prompt to Graph: Comparing LLM-Based Information Extraction Strategies in Domain-Specific Ontology Development, *Procedia Computer Science* 277 (2026) 50–60. Part of special issue on 7th International Conference on Industry of the Future and Smart Manufacturing (former International Conference on Industry 4.0 and Smart Manufacturing).
- [6] H. Tan, Y. Yang, T. Yuan, Z. Li, Bootstrapping Knowledge Acquisition from Domain Text with a Seed Ontology and Small Language Models, in: Proceedings of the 25th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2026), 2026. Accepted for publication.
- [7] Z. Li, X. Liu, H. Tan, X. Wang, L. Peng, B. Li, J. Y. H. Fuh, W. F. Lu, F. Bonollo, M. Tiryakioglu, M. Easton, A. Luo, S. Pogatscher, A. E. W. Jarfors, CASTON: An Ontology for Enabling AI in Metal Casting, Materials, and Manufacturing, 2026. Under review at *Scientific Data*.
- [8] T. Schick, H. Schütze, It's not just size that matters: Small language models are also few-shot learners, in: Proceedings of the 2021 conference of the North American chapter of the association for computational linguistics: Human language technologies, 2021, pp. 2339–2352.
- [9] L. C. Magister, J. Mallinson, J. Adamek, E. Malmi, A. Severyn, Teaching small language models to reason, in: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2023, pp. 1773–1781.