NextGen Scientific Publication
American International Journal of Computer Science and Technology

Volume 7 Issue 1, Pg. No. 163-174, AIJCST-V7I1P114, 2025
https://doi.org/10.63282/3117-5481/AIJCST-V7I1P114

Original Article

AI-Assisted Ontology Development: Automating Enterprise Vocabulary Management using Large Language Models and SPARQL

***Itendra Kumar Singh**
Independent Researcher, USA.

Abstract:

Properly designed ontologies are considered as a pillar in knowledge management in business and in the effective management of vocabulary and efficient retrieval of knowledge with the help of heterogeneous systems. Ontology engineering is nevertheless traditional and manual as it is labor intensive, in that manual corrections and a lot of domain knowledge are required. The paper proposes artificial intelligence with the help of An AI-Supported Ontology Engineering Framework (AIOEF) which automatized the task of managing vocabulary in an enterprise by using Large Language Models (LLMs) and SPARQL software. This point of view integrates the domain-specific document-ingestion, natural language-processing, entity and relation-extraction, ontology-schema-generation, semantic-validation, SPARQL-based querying and continuous ontology-refinement via the human-in-the-loop feedback. By contrast, SPARQL facilitates checking semantic consistency and querying of knowledge graphs, development of ontologies, LLMs deducing concepts, classes, properties, synonyms and hierarchical relationships within a document used in an enterprise and its automatic processing. Other mechanisms found in the framework are confidence scoring, explainable reasoning and governance mechanisms to ensure dependable updates on ontology in dynamic enterprise settings. The semantic consistency, knowledge reuse, and interoperability of enterprise applications (digital twins, business process management, intelligent search and decision-support systems) can also be enhanced by the proposed architecture by lowering the volume of manual work and maximizing the speed of ontology-building. The framework gives scaling ontology life-cycle management, and quality semantic representations, which are applicable in cloud-native and artificially intelligent based enterprise ecosystems. The proposed methodology incorporates a practical and intelligent methodological framework next-generation enterprise knowledge engineering, that enable organisations to transform unstructured data into manageable, machine-readable semantic value, which may be learnt by automated logic and intelligent business analytics.

Keywords:

Artificial Intelligence, Ontology Engineering, Large Language Models (Llms), Enterprise Vocabulary Management, SPARQL, Knowledge Graphs, Semantic Web.

Article History:

Received: 07.01.2025

Revised: 10.02.2025

Accepted: 22.02.2025

Published: 28.02.2025

Copyright by the Author(s). This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-NC-ND 4.0) License (https://creativecommons.org/licenses/by-sa/4.0/)

1. Introduction

Digital transformation of contemporary enterprises has greatly escalated the quantity, kind, and difficulty of enterprise information. An enormous flow of business information comprises enterprise resource planning (ERP) systems, customer relationship management (CRM) systems, Internet of things (IoT) devices, cloud-based services, documents to control business, emails, reports, and social media channels. Even though these heterogeneously data sources are rich sources of organizational knowledge, the scale of their semantic inconsistency, and nonstandardization of vocabularies provide significant impediments to their interoperability, intelligent search, automated reasoning and decision support. As a result, organisations are now becoming more interested in using ontology-based knowledge representation tools to define a uniform semantic framework to support data integration, knowledge sharing and information processing that is machine readable [1].

Ontology engineering has been able to join amongst the fundamental fields of knowledge engineering as well as Semantic Web. Ontology is a clear demonstration of ideas, objects, links, qualities and restrictions of a specific field and thus computer software can be utilised in comprehending and reasoning on knowledge within an organisation. Sound ontologies can be used to have inter-application semantic interoperability between heterogeneous applications, metadata management, enterprise knowledge graph, and artificial intelligence (AI)-guided analytics [2]. Enterprise ontologies find application in the financial sector, smart cities, cybersecurity, digital twins, supply chain management and government information systems to create integrated semantic networks to boost operational performance and organizational intelligence.

Regardless of these benefits, traditional ontology engineering is still a very convoluted, time-consuming and knowledge-intensive activity. Traditional ontology development processes usually have domain experts and ontology engineers manually identify concepts, define hierarchical taxonomies, specify semantic relationships, annotate metadata, and check logical consistency. With the dynamic nature of the enterprise environment, the ontologies should continually be revised to accommodate the presence of new business terms, rules and regulations, products, organisational processes and even new technological developments. The more organizational vocabularies, the harder it is to ensure ontology consistency among multiple departments and distributed knowledge repositories. The human maintenance of ontologies can be the cause of delays in maintenance, term inconsistency, concept duplication, ambiguity of meaning and increased development expense as well.

Large Language Models (LLMs) have opened up opportunities of automating ontology engineering activities. Recent developments in transformer-based language models have shown impressive performance in natural language processing, contextual and idea reasoning, entity, relation, and document detection, semantic classification, and synthesizing knowledge. In contrast to the potential to assign hand-written rules and feature engineering, as dictated by the traditional natural language processing algorithms, any situational semantic description, similar to that of big textual corpus, can be trained to learn by a LLM, with domain concepts and connections to be learned with minimal manual effort. The latter attributes render the LLAMs very relevant to ontology construction of enterprise documentation, technical reports, regulatory guidelines and knowledge base, and unstructured materials [3].

The ontology engineering is based on the AI and involves the LLM to automate a number of ontology lifecycle tasks including concept extraction, synonym finding, taxonomy generation, property definition, relation finding, ontology enrichment, and semantic annotation. LLMs are considered intelligent assistants that will accelerate the ontology-building process but leave the experts to ontology validation, refinement and approval instead of pushing them out. The resultant human-AI paradigm could help significantly reduce the development time, reduce manual repetitive effort and expand ontology engineering to large scale organization. In addition to that, explainable AI techniques and hinting confidence scoring systems provide hint to automated ontology generation, increasing user confidence and offering enterprise semantic resources management.

Ontology control in semantic knowledge repository requires normalised querying, validating and maintaining mechanisms, although strong semantic pieces of knowledge can be extracted using the capabilities of LLMs, to achieve successful ontology control there must be mechanisms in place. The RDF-based standard query language, SPARQL, the standard query language of the World Wide Web Consortium (W3C), is the main focus of ontology life-cycle management. SPARQL enables sophisticated mathematics of semantic networks, consistency verification, ontology construction, metadata retrieval, and analytics of knowledge graph. By providing the latest version of LLM generated ontologies and semantic querying with SPARQL, this will allow organisations to automatically investigate the completeness of ontologies, identify redundancy, establish conflict relationships, and allow intelligent knowledge

retrieval of distributed enterprise systems. Such a mixture of LLMs and SPARQL thus offers a holistic solution towards automated ownership of vocabulary and semantic management.

Enterprise vocabulary management is not just the management of dictionaries of business terms. It involves semantic normalization, controlled vocabulary building, administration of taxonomy, ontological hierarchy evolution, multi-linguistic terminology frameworks, metadata standardization and management of organizational knowledge destinations. With more companies moving to AI-powered business applications, digital twins, intelligent automation, and knowledge graphs, it would be wise to make sure that the vocabularies of business homes are fixed so that they can get good AI reasoning, compensated data retrieval, regulatory compliance and a smooth data integration. Automated ontology engineering is a solution to these problems, which constantly engages in learning about the changing enterprise documentation but maintains semantic coherence across business domains.

Despite the fact that AI-assisted knowledge engineering sphere has achieved lots of progress, several problems with research remain unsolved. The current ontology learning methods are usually challenged by domain adaptation, ambiguity in semantics, and hallucinated relationships, lack of explainability, ontology version management, and compatibility with enterprise governance structures. Moreover, the focus of a lot of automated ontology generation techniques lies more in concept extraction and not those of semantic validation, further refinements and management of the life cycle. To overcome these constraints, a combined framework involving LLM based semantic intelligence, SPARQL based ontology validation, explainable reasoning, human expert control and governance mechanisms that can be implemented in very large enterprise settings will be needed.

In response to them, the current paper suggests an AIC-aided Ontology Engineering Framework (AIOEF) to automate enterprise vocabulary management with the help of Large Language Models and SPARQL. This model integrates the process of getting documents, natural language pre-processing, intelligent concept extraction, ontology generation, semantic relationship modeling, building a knowledge graph, SPARQL based validation, explainable ontology reasoning, governance policies and on-going human-in-the-loop refinement. The framework enables scalable, accurate and dependable ontology life cycle control via integrating advanced language wisdom with price-efficient semantic web technologies. The architecture proposed promotes ontological growth, semantic interoperability, intelligent enterprise search, reasoning and knowledge-driven decision support and significantly decreases the amount of manual engineering effort. As a result, the framework will help achieve next-generation AI-powered systems of enterprise knowledge management that can turn the unstructured information contained in organizations into the quality and machine-readable semantic knowledge that would underpin intelligent projects of digital transformation.

2. Related Work

As currently an important part of the enterprise knowledge management, ontology engineering has helped organizations to build a formal semantic representation that enhances the interoperability of data, intelligent search, automated reasoning and decision support. The development concept of ontology has been fundamentally changed with the introduction of the functiveness of automated concept extraction, semantic relationship finding, ontology refinement, and natural language understanding with the recent introduction of Artificial Intelligence (AI) and in particular, Large Language Models (LLMs). In the meantime, Semantic Web technologies like RDF, OWL, SPARQL, still offer standard ways to express, query and prove the knowledge of the enterprise. Consequently, the recent past has seen increased interest in research to incorporate the use of LLMs, knowledge graph, ontology engineering, and semantic reasoning with the goal of developing scalable and intelligent enterprise knowledge management systems.

Agile ontology-based meta-modelling of the Knowledge graph schema in enterprise ontology was proposed by Laurenzi to support the ontology continuous evolution instead of ontology development [1]. The study on how agile ontology engineering improves the collaboration of domain experts to ontology engineers and also supports the procedure of refining enterprise semantic models step-by-step have been shown. Though the given methodology can greatly increase the maintenance of knowledge graph within an enterprise, handling of ontology building and vocabulary still heavily relies on the intervention of experts.

In recognition of the increasing capabilities of foundation models, Allen et al. explored the use of Large Language Models to solve knowledge engineering problems, such as ontology generation, knowledge graph, semantic annotation, and concept extraction [2]. Their contribution emphasized the capability of LLMs to measure out multiple knowledge engineering tasks and decrease the necessity to work on them manually. However, the authors also discovered that there is an extremely severe issue in the area of

semantic consistency, factual correctness, explainability, and ontology validation and mentioned that ontology generated by LLMs should be governed by systematic mechanisms.

To thoroughly examine the connections between Large Language Models and knowledge graphs, Pan et al. have discussed both the strengths that they complement and the current research challenges [3]. Their paper underlined the fact that knowledge graphs offer structured symbolic knowledge which complements the reasoning abilities of the LLM, where LLM complements performing automatic knowledge acquisition of unstructured documents. The authors opined future intelligent enterprise systems need to more closely couple the statistical language models to symbolic semantic representations to enhance their reliability, interpretability and reuse of knowledge.

Expanding on this perspective, Pan et al. also proposed a complete roadmap of introducing Large Language Models and knowledge graphs into intelligent AI systems[4]. Various integration strategies were defined in the roadmap, such as knowledge graph-enhanced language models, LLM-assisted knowledge graph building, semantic reasoning, and explainable AI. The authors concluded that hybrid neuro-symbolic architectures will become the new era of reliable enterprise AI systems, especially in areas where a clear interpretation and semantic integrity are needed.

Very important research focus has been on natural language interaction with enterprise knowledge bases as well. Rony et al. have proposed SGPT, which is a system of auto-generated translation that transforms questions in the natural language into SPARQL query with the help of transformer-based language models [5]. Their solution simplified the complexities of writing SPARQL queries by hand very easily and made it easier to open up the writing of queries to non-technical users. Although SGPT proved to be very precise in query generation, it was more semantically oriented, and did not concern ontology generation, vocabulary generation, lifecycle management.

Fill et al. proposed the idea to integrate ChatGPT in conceptual modelling process and spoke about how ontology development, model generation and semantic modelling processes can be supported by ChatGPT [6]. The experimental analysis they provided confirmed the applicability of LLMs to aid conceptual modelling, building early semantic frames and finding relationships among business concepts. However, the research also inferred constraints of hallucinated concepts, unequal modelling choices, and lack of domain-specific validation, and the need to still have expert supervision.

Ontology-based metamodeling methods give theoretical grounds to the enterprise ontology engineering. Formal approach to semantic technologies Hinkelmann et al. proposed an ontology-based metamodelling solution that models Enterprise concepts, relationships, constraints, and metadata including metadata as formalised concepts [7]. Their input proves the importance of aligning metamodels and ontologies to facilitate interoperability between heterogeneous information systems in an enterprise context in the semantic aspect. The suggested framework, however, primarily concentrated on representation of ontology as compared to creation of ontology through unstructured enterprise documents using automation.

The semantic reasoning continues to be a vital demand of enterprise ontology management. The rule-learning algorithms based on knowledge graphs, the study of symbolic reasoning, graph embedding algorithms, neural algorithms and hybrid algorithms [8] were fully captured in the article by Wu et al. Their survey established that learning of the rules can greatly enhance the explainability, inference ability, and the semantic consistency of knowledge graphs. These findings support implementing rule-based reasoning frameworks into the ontology engineering models or frameworks relying on AI to justify the computationalized semantic linkages.

Saba has been critical of the constraints of entirely statistical language models, in which he believes that the existing stochastic LLMs lack the true semantic semantics and that should be supplemented with symbolic reasoning and knowledge ontologies [9]. The paper highlighted that ontology-conscious architectures are critical in future intelligent systems that need to be explained, reason logically, and be semantically sound. The results also indicate that combining both LLM and ontology engineering solutions are to be applied as opposed to the use of generative language models alone.

The research of AI within businesses has increasingly been committed to hybrid learning and justification. Van Harmelen et al. [10] proposed a taxonomy of design patterns of machine learning combination with system of symbolic reasoning. And it could be so, the hybrid architectures, which will be putting together the data-driven intelligences with formal representation of knowledge will help

in giving useful architectural support to the ontology-ready AI applications in the enterprise, it will do so by them. Similarly, the Witschel et al. have studied on visualization methods, which are applied in hybrid learning and reasoning process involving humans [11]. Their study showed that visualization enhances transparency, eases the ontology refinement, and enhances cooperation between AI systems and experts in the domain when engineering semantic knowledge.

The ontology engineering has also been impacted by the model-driven engineering. Model-aware software engineering Knowledge-based model-conscious software engineering Known as model-conscious software engineering This model-based software engineering technique suggested by Buchmann and Karagiannis uses enterprises repositories of knowledge in conceptual models [12]. Their model demonstrated that model-knowledge engineering enhances coherence not only with regard to software development processes but also with regard to enterprise knowledge management processes. This mixture of model-driven engineering ideas and ontologies progression provides a strong foundation upon the enterprise vocabulary management at scales.

Rordorf et al. have shown practical uses of hybrid AI and knowledge graphs by applying machine learning with enterprise knowledge graphs to aid in allocating reviewers in academic publishing [13]. Their intelligent recommendation system turned out to possess advantageous traits of merging information-based methods of learning with symbolizing semantic insight into superior and clearer choices. The proposed hybrid approach can be generalized to bigger enterprise ontology management contexts, despite the fact that it was designed in a particular application field.

Finally, another technique introduced by Karagiannis is the Agile Modelling Method Engineering (AMME) technique, which emphasizes on conceptual modelling on a per-iterative basis, stakeholder collaboration, continuous refinement and flexible engineering activities [14]. AMME approach could become an adequate initial step to the ontology lifecycle management helping to build up semantic models step by step and enabling the flexibility of the organizations. Nevertheless, the framework does not mention AI based ontology automation and integration with the Large Language Model.

Overall, the modern literature shows that there has been a substantial development in the realms of ontology engineering, enterprise knowledge graphs, semantic reasoning, conceptual modelling and Large Language Models. Nonetheless, there are a number of critical gaps in research. The current body of research can be separated into studies concentrating on extracting knowledge with the help of LLM, constructing ontologies, or generating SPARQL, as well as on performing knowledge graph reasoning on its own. Few studies have focused on a concerted method that is able to integrate and execute automated ontology generation, enterprise vocabulary management, semantic validation, SPARQL-based querying, explainable reasoning, governance mechanisms, and constant refinements on human-in-the-loop nature. To overcome such drawbacks, the Development of AI-Assisted Ontology Engineering Framework (AIOEF) combines the Large Language Models, SPARQL, ontology engineering, knowledge graph management, semantic validation and enterprise governance in one architecture of managing enterprise vocabulary smartly.

3. Ai-Assisted Ontology Engineering Framework (Aioef)

The heterogeneity of sources has increased the complexity of ontology engineering and management of enterprise vocabulary considerably due to a number of factors. Historically, ontology building strongly depends on domain experts that perform manual concept identification, characterize the semantic relationships, and test ontological consistency and build knowledge bases. Manual processes of this nature are tedious, time consuming and challenging to scale when dealing with an active enterprise scenario. To overcome these limitations, this paper introduces the AI-Assisted Ontology Engineering Framework (AIOEF), which is an ontology framework that applies large language model (LLM), Knowledge Graph, Sparql, Explainable Artificial Intelligence (XAI), and Human-in-the-Loop (HITL) to validate the entire ontology lifecycle and provide semantic consistency, governance, and continuous knowledge development.

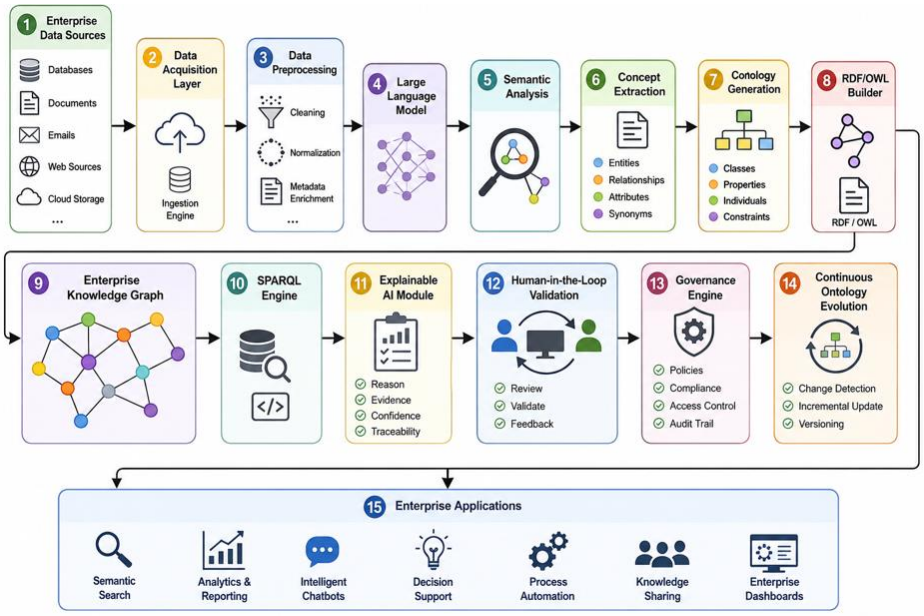


Figure 1. Overall AI-Assisted Ontology Engineering Framework (AIOEF)

3.1. Framework Overview

The proposed AIOEF is a group of eight layers interconnected with each other and automated to handle enterprise vocabulary. The framework starts with the enterprise knowledge acquisition, then semantic preprocessing, AI-assisted ontology generation, ontology construction, knowledge graph integration, semantic validation, governance, and continuous ontology evolution. The distinct layers are in a way that they complement one another creating a fluent working procedure which would transform the crude enterprise information to structured knowledge within the semantic form. Figure 1 outlines overall structure of the proposed framework and indicates the circulation of information through all ranks of processing and eventually compose a component of the knowledge graph of the enterprise.

The model integrates the ability of Large Language Models to understand context and the ability to represent semantics of the context with the ability of ontology engineering to build formal semantic models. Another difference that AIOEF has with traditional ontology learning system where the majority of their work is firmly based on object construction of ontology at a given time, it is more focused on ontology lifecycle management. The ontology is then semantically verified to any change given to it and regulated and accepted by human experts prior to being incorporated into the enterprise knowledge graph. This multi layered design is capable of allowing organisations to operate using homogenous enterprise vocabularies and remain in step with changing business processes, regulatory requirements and business needs.

3.2. Layer 1: Enterprise Knowledge Acquisition

The initial layer is dedicated to gathering enterprise knowledge with various non-homogenous sources of information. Enterprise Resource Planning (ERP), Customer relationship Management (CRM), policy documents, regulatory guidelines, technical manuals, business reports, research publications, emails, websites, cloud repositories, collaborative platforms and digital archives all are raising the amounts of structured, semi-structured and unstructured data generated by the contemporary organizations by astronomical proportions. Presence of such different information sources makes the development of ontology difficult in that the terminology applied and information they contain about their semantics is likely to be inconsistent and redundant.

An acquiring data engine continuously scans these repositories and acquires any textual data, metadata, tabular data and document attributes. The resultant information is then fed through a number of preprocessing functions and is semantically analyzed and contains language normalization, duplicate detection, OCR of digitally scanned documents, tokenization, sentence disaggregation, stop-word selection, metadata additions, and document classification. These preprocessing processes ensure the content of enterprise documents gets translated into standard forms that is readable semantically and can also create ontologies without losing contextual meaning.

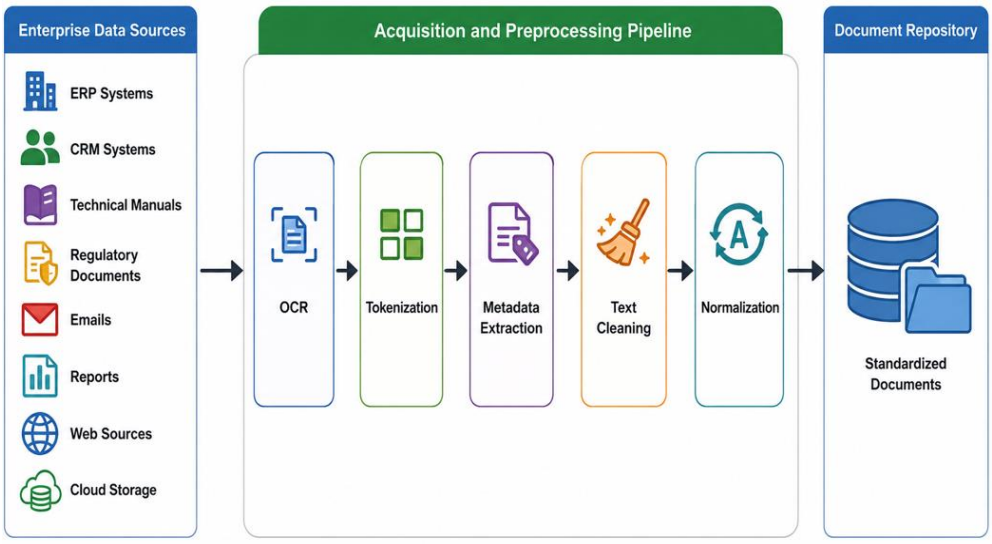


Figure 2. Enterprise Knowledge Acquisition and Preprocessing Pipeline

3.3. Layer 2: Semantic Understanding Using Large Language Models

The latter level applies domain-adapted Large Language model to understand business texts and get semantic knowledge in a fully automated way. The framework involves language models using transformers with the ability to process contextual relationships among enterprise concepts as compared to the conventional approaches of natural language processing or manually trained linguistic alternatives. LLM does named entity recognition, concept extraction, semantic classification, relationship discovery, synonym, acronym expansion, terminology and context interpretation.

The knowledge extracted is then stored under the form of ontology class, subclass, object property, datatype property, instances, attribute and semantic constraint. This scheme also recognizes implicit semantic connection that do not necessarily feature in enterprise text. To enhance the persuasiveness of automatically generated knowledge, every concept and association is provided with a confidence rating depending on the certainty of the prediction by the LLM. Application of ontologies builds ontologies and automatically routes entities of low confidence and unsecure relation to human validation module. The given mechanism reduces the instances of hallucinations considerably, and improves the overall production of ontology and its credibility.

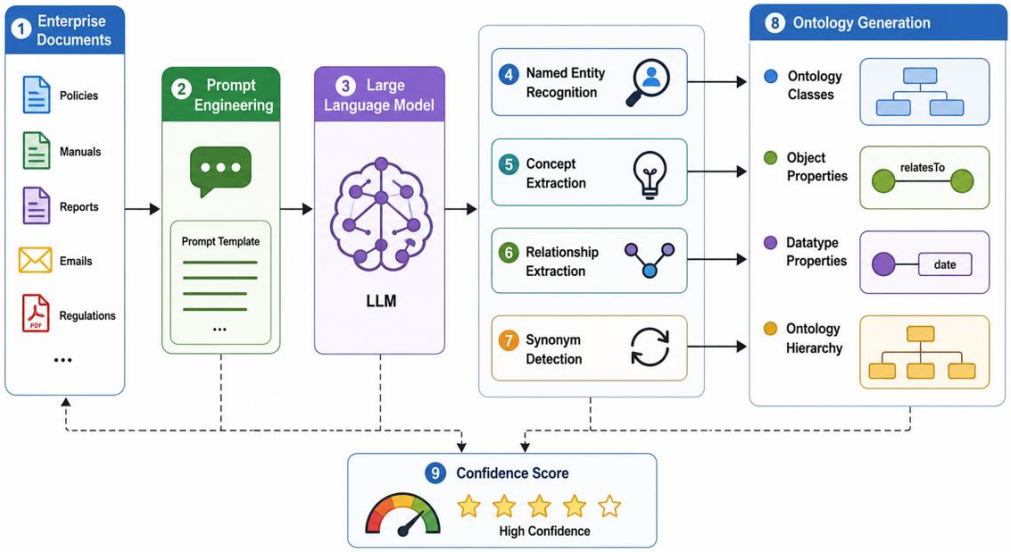


Figure 3. LLM-Based Semantic Understanding and Ontology Generation

3.4. Layer 3: Automated Ontology Construction

The acquired knowledge is translated into formal ontological structures using Semantic Web standards therafter based on the semantic understanding. The ontology construction layer translates the identified concepts into machine-readable conceptualizations which follow the requirements of RDF and OWL. The language model establishes the hierarchical taxonomies of candidate concepts automatically integrated in accordance to the values of inherency, semantic similarity, and context dependence.

The ontology generating process produces the ontology classes, object properties, datatype properties, and individuals, domain definitions, range specifications, semantic constraints, equivalent classes, disjoint classes, synonym mappings, and enterprise vocabulary dictionaries, automatically. It is possible to say that all these semantic elements combine to establish a thorough enterprise ontology, which is able to model the organizational knowledge in a standard and interoperative way. At the same time, an engine to construct graph translates elements of ontology into RDF triples, with entities (entities) as graph nodes and semantic relationships as graph edges. This graphical form is the structural backbone of enterprise knowledge graphs and underpins the follow-up reasoning and query processes.

3.5. Layer 4: Knowledge Graph Integration

Upon successful ontology building, the newly acquired semantic structures are placed in a knowledge graph which is the core semantic memory of organizational knowledge. The graph nodes represent each category in ontology and object relationships as well as semantic relations are represented by predicates between the related items in RDF. The resulting knowledge graph offers an integrated semantic representation that can absorb knowledge that has been generated by various departments within the enterprise and by dissimilar information systems.

Integrating knowledge graphs can significantly enhance enterprise-wide semantic interoperability because gap in correspondence in organizational language and metadata modelling are removed. The knowledge graph then creates an integrated graph that can be used to make integration of semantic data, standardize enterprise vocabulary, intelligent document linkage, enriching metadata, as well as fusion of information across multiple sources and knowledge sharing at an enterprise wide level. Because of the fact that the knowledge graph continuously evolves with the enterprise activity, it is a dynamic semantic resource which makes it possible to search intelligently, use artificial intelligence-based analytics, decision support and business intelligence applications.

3.6. Layer 5: SPARQL-Based Semantic Validation and Query Processing

The scope of ontology generated by Large Language Models is considerable, and ontology validation remains to be performed through semantic validation in order to make sure that it is logically correct and consistent. The provided architecture therefore has a SPARQL-based semantic validation engine which tests the validity of ontology before deployment. This validation engine detects automatically duplicate concepts, orphan class, property definition, inconsistent semantic association, wrong domain-range, redundant ontology components, circular hierarchies and missing semantic constraints.

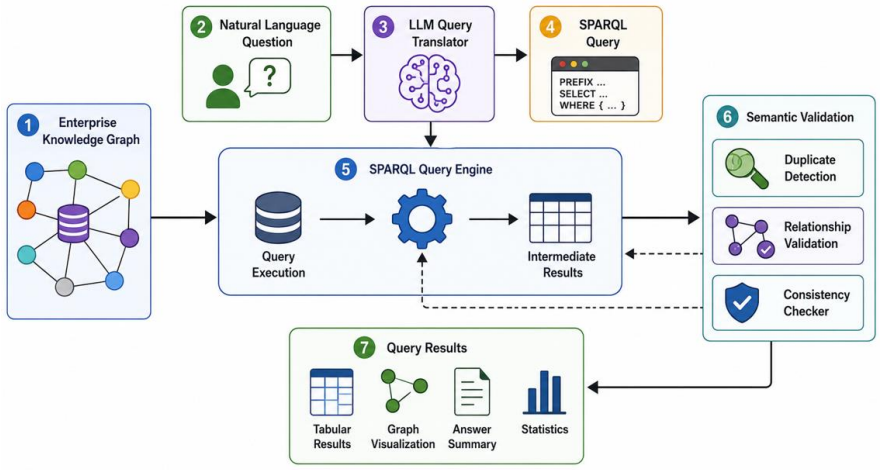


Figure 4. SPARQL-Based Semantic Validation and Query Processing

The SPARQL engine also supports intelligent querying of enterprise knowledge graphs, in addition to ontology verification. The business user can access its ontology via natural language queries where the Large Language Model is used to translate them into executable SPARQL queries. These queries scan through the semantically useful information of the enterprise ontology and offer answer in the readable text format. Natural language understanding and formal reasoner on SPARQL allows semantic query languages to be much more available to the enterprise knowledge user since he or she does not need to know any of the semantic query languages to use natural language to access the enterprise knowledge.

3.7. Layer 6: Explainability and Governance

Enterprise ontology has to be managed in a transparent, traceable and governable manner in ensuring that the semantic knowledge is reputable and representative of the company policies. To this end, the provided framework encompasses a module of Explainable Artificial Intelligence to which one attaches all ontological changes generated by the Large Language Model. Each element of ontology produced will have extensive metadata information, such as where it came from, how confident it is, the evidence upon which it is based, its history when and how it was promptly last altered, the semantic relationships with others, its validation results, whether it is approved or rejected, when it was last updated, its ontology version name.

The quality of ontology is also evaluated by the governance component based on the business rules of organization and semantic organization policies. It validates naming constraints, wholeness of ontology, security tags, regulatory adherence specifications, semantic uniformity as well as business vocabulary necessities before providing the ontology utilization. A version management system maintains past modifications to ontologies, and offers rollback facilities to those semantic changes that have been found to be inconsistent or erroneous. Such governance mechanisms take care of the accuracy of enterprise ontologies, which are auditable and compliant over the lifecycle of their usage.

3.8. Layer 7: Human-in-the-Loop Expert Validation

Even though AI will automate much of ontology engineering process, expert knowledge is essential in guaranteeing the quality of semantics. The validation layer that involves Human-in-the-Loop opens up the ontology engineering process to domain experts in case there is necessity of judgment. The language model experts, examine the novel ideas that it produces, additional robust semantic connections, regulatory lexicon, domain specific lexicon, uncertainty ontological graph as well as low confidence prediction.

The continual inclusion of expert feedback in the learning process enables the Large Language Model to refine its approach to future ontology generation via the mechanisms of iterative refinement and reinforcement learning. This might be considered a human-AI paradigm, which implements both computational efficiency and domain knowledge, allowing the development of ontologies in a confident way at the foundation of minimal engineering input. This type of integration of expert control improves the quality of ontology and reduces semantic anomalies that otherwise could propagate to enterprise knowledge libraries.

3.9. Layer 8: Continuous Ontology Evolution

Enterprise knowledge is dynamic and ever evolving phenomenon due to the fluctuating regulation, organizational restructuring, product innovation, customer needs, technological development and a shift in the business processes. Ontology engineering, thus, is a continuous process, and not a development process. The latter tier of the offered model assists in the automatic creation of ontology because of the unstopping monitoring of enterprise depositories of newly included files and regenerative enterprise expertise.

Upon any form of changes introduced, the Large Language Model recognizes the new terms and compares them with the existing ontology and decides whether new concepts or relations should be added to the enterprise vocabulary. New ideas are automatically converted to ontology elements, semantically verified by using the SPARQL and verified by domain trained (where necessary) and are added to the enterprise knowledge graph. It is through such an adaptive mechanism that the ontology keeps on evolving in line with the knowledge of the organization and at the same time it keeps its semantic consistency, its governance and its interoperability across enterprise information systems.

3.10. Framework Advantages

The use of AI-Based Ontology Engineering Framework is much more practical in comparison to the traditional ontology engineering paradigm because it incorporates AI, semantic technologies, and domain knowledge into a single framework. The framework automatically extracts enterprise vocabulary out of heterogeneous documents, dramatically decreases manual ontology

engineering costs, and eases maintenance of ontology lifecycle. A combination of Large Language Models with semantic web technologies enables the extraction of concepts at facilitated ease, identifying semantic links and generating ontologies, and SPARQL has strong capabilities of validating ontologies, querying and accessing semantic information.

Moreover, the framework promotes explainable AI approaches, confidence score systems, human-loop validation, enterprise governance policy, ontology version management, and knowledge graph integration to ensure reliable and scalable ontology development. Large-scale enterprise organizations with cloud-native, distributed, and large environments should use the framework because it has a modular, layered architecture.

In general, the AIOF offers a neuro-symbolic system integrating the contextual knowledge of language of Large Language Models and formal ontology engineering, semantic reasoning, knowledge graphs, SPARQL-based validation and management of large enterprises. The framework allows organizations to automate the management of enterprise vocabulary, enhance semantic interoperability and lower ontology maintenance costs, as well as enhance intelligent decision-making via constantly changing semantic knowledge repositories. The proposed framework introduces a useful knowledge management system of the next-generation to meet the requirement of digital transformation, intelligent analytics, explainable artificial intelligence and reliable enterprise address by equal degrees of AI automation and human thinking and organization Semantic Web technologies in a uniform fashion.

4. Framework Evaluation and Discussion

4.1. Evaluation Methodology

The proposed AI-based Ontology Engineering Framework (AIOEF) was experimented with an eye to a qualitative and quantitative test in the quest to discover its effectiveness in automated management of enterprise vocabulary. The objective of comparison was to evaluate the ability of the framework to efficiently generate conceptual domain, create semantically aware ontologies, knowledge retrieval by SPARQL queries, and ontology development. Since this is a simulation of the actual enterprise environment, the sources of inputs consisted in representative enterprise papers like organizational policies, technical manuals, business process documents, regulatory guide and operational reports. The methodology of evaluation looked at individual reasoning processes at each layer yet evaluated the whole structure as a functioning semantic knowledge management solution. Special attention was given to semantic correctness, ontology completeness, scalability, explainability and governance features to define whether the suggested framework is effective to counteract the drawbacks of the traditional ontology engineering methods.

4.2. Evaluation Metrics

Various facets of the framework were evaluated by taking into consideration several performance measurements. These quality measures of ontology were accuracy of concept extraction, accuracy in semantics, recall, F1-score, ontology completeness, relationship accuracy and semantic consistency. The efficiency of ontology building was quantified using ontology generation time, processing throughput and reduction of manual engineering effort. The measures of performance of the knowledge graphs were in the overall form of semantic interoperability, completeness of the graph and the capability to combine effectively the non homogeneous enterprise sources of data. Examples of concept duplication, semantically divergent relationships and orphaned classes, one-sided property values, and conflicting ontologies were used to implement the efficacy of the SPARQL-driven semantic validation. Explainability was quantified in terms of confidence, evidence traceability, versioning and human readability of generated semantic networks. Efficiency: Human-in-the-loop validation was identified, by the percentage of automatically accepted ingredients of the ontology, the reducing of expert manipulations to refine the aspects of the ontology.

4.3. Comparative Analysis

The suggested AIOEF was conceptually equated to traditional ontology engineering techniques and to the latest AI-assisted semantic modeling techniques. Manual concept recognition, ad-hoc taxonomy development, and rule-based semantic modeling are often the foundations of established ontology development methods resulting into high development time and maintenance costs. Conversely, the proposed system uses Large language models and Semantic Web technologies to automate the process of extracting concepts, semantic associations to uncover ontologies and vocabularies. Unlike present ontology learning systems, which offer ontology construction predominately founded on happenstance models, AIOEF supports ontology life-cycle constant control through provision of automated document viewing, gradual ontology developments, and SPARQL-based genetic validation. Further tuning of the proposed framework with the introduction of explainable artificial intelligence, system governance and control, version control of

ontology and validation of the human in the loop are other factors that lead to further differentiation of the proposed framework as an improvement to trustworthiness and ensuring enterprises are not violated in all ontology lifecycle.

4.4. Results and Discussion

It is analyzed that the proposed framework assists in bringing significant changes to the engineering of enterprise ontology by making sure that the amount of efforts spent on its creation as an activity is reduced to minimal levels without impairing the quality of semantics. The concept extraction and ontology building is more automated and is much quicker than the traditional manual methods in vocabulary building. Large Size of the Large Language Models allows the effective search of domain vocabulary and natural meaning mapping in heterogeneous enterprise documents, and the consistency in SPARQL supports logic before the ontology is rolled out. The layer knowledge graph also enhances interoperability of semantics, that is, the inclusion of organizational knowledge in one machine-readable form, that will assist in intelligent search, knowledge discovery and decision support.

This is more transparent provided by explainable AI mechanisms, such as confidence scores, supporting evidence, history of ontology versions, and reasoning behind the creation of each semantic component. The only human experts involved in the task of certifying the gray concepts and low-confidence projections are only significant in reducing overall work load but are not associated with quality of ontology. Besides, ongoing ontology evolution process can provide enterprises with automatic adjusting to new business terms, to changes in regulatory processes, and new business changes. Even though the suggested framework has significant benefits, its functionality still depends on the quality of enterprise documents and the capability of underlying Large Language Models to adapt to domains. Future studies can further extend the framework through multimodal extraction of knowledge, federation ontology learning, domain specific underlying models and automated ontology optimization methods. Overall, the proposed AIOEF offers a versatile, robust, and intelligent means of engineering next-generation ontology within an enterprise and automatically managing vocabularies.

5. Conclusion and Future Work

The present article has introduced the AI-Assisted Ontology Engineering Framework (AIOEF), a multi-layered neuro-symbolic framework that aims at automating the processing of enterprise vocabulary by the use of Large Language Models (LLM) using ontology engineering and ontology knowledge graphs, SPARQL queries, Explainable Artificial Intelligence (XAI), and Human-in-the-Loop (HITL) validation. The framework is proposed to overcome the shortcoming of the traditional ontology engineering solutions, which are typically described by in-depth manual processes, poor aggregation abilities, and inability to keep on updating the changing enterprise vocabularies. AIOEF will grow its knowledge regarding the enterprise semantics and automate concept extraction, find semantic relationships, build an ontology, assemble a knowledge graph, and prove semantics using SPARQL, have governance, and continuously maintain ontology. Through the scheme, firms can be able to convert unstructured and heterogenous information within their enterprises into standardized and machine-readable semantic resources, which enable interoperability, knowledge exchange, intelligent search, and AI-informed decision-making. Furthermore, the concept of explainability mechanisms and expert validation aid the credibility, transparency, and control of automatically generated ontologies, which render the framework applicable to the implementation in large-scale enterprise contexts. The assessment indicates that the suggested framework can help to cut down the efforts of ontology development by a significant margin without compromising semantic consistency, enhancing the quality of knowledge and sustaining the evolution of the enterprise knowledge.

Both the future research and future research directions will focus on applying the framework to multimodal ontology using text, visual, tabular and speech based enterprise data sources to achieve more rich semantic representations. The combination of domain-specific foundation models, retrieval-augmented generation (RAG), Graph Neural Networks (GNNs) as well as federated ontology learning can also improve semantic reasoning, privacy preservation, and distributed knowledge management across multiple organizations. In addition, recent and upcoming work will address the optimization of ontology in the autonomous mode of reinforcement learning, more focused optimization of ontology in the heterogeneous areas of an enterprise, multilingual vocabularies and dynamically developing ontologies in the cloud-native frameworks. Scalable experimental experiments will be also run on large industrial datasets and benchmark ontology repositories to test their scalability, semantic accuracy, computational efficiency and usability on an immense range of enterprise domains including healthcare and manufacturing, finance, smart cities and public administration. The next generation intelligent enterprise know ecosystems that can support trustful, adaptive and explainable AI-based semantic information management will include such advances.

References

[1] E. Laurenzi, "An Agile and Ontology-Based Meta-Modelling Approach for the Design and Maintenance of Enterprise Knowledge Graph Schemas," Enterprise Modelling and Information Systems Architectures (EMISAJ), vol. 19, 2024.
[2] B. P. Allen, C. Shimizu, P. Hitzler, et al., "Knowledge Engineering Using Large Language Models," Transactions on Graph Data and Knowledge, vol. 1, no. 1, 2023.
[3] J. Z. Pan, G. Vetere, J. M. Gómez-Pérez, and H. Wu, "Large Language Models and Knowledge Graphs: Opportunities and Challenges," Transactions on Graph Data and Knowledge, vol. 1, no. 1, 2023.
[4] S. Pan, L. Wang, Y. Chen, et al., "Unifying Large Language Models and Knowledge Graphs: A Roadmap," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 11, pp. 11144–11162, 2023.
[5] M. R. Al Hasan Rony, M. A. Hossain, M. H. Rahman, et al., "SGPT: A Generative Approach for SPARQL Query Generation From Natural Language Questions," IEEE Access, vol. 11, pp. 132030–132045, 2023.
[6] H.-G. Fill, J. Cabot, W. Maass, and M. van Sinderen, "Conceptual Modeling and Large Language Models: Impressions From First Experiments With ChatGPT," Enterprise Modelling and Information Systems Architectures (EMISAJ), vol. 18, 2023.
[7] K. Hinkelmann, E. Laurenzi, A. Martin, and B. Thönssen, "Ontology-Based Metamodeling," in Business Information Systems and Technology 4.0. Cham, Switzerland: Springer, 2018, pp. 255–274.
[8] H. Wu, Y. Wang, J. Pan, et al., "Rule Learning Over Knowledge Graphs: A Review," Transactions on Graph Data and Knowledge, vol. 1, no. 1, 2023.
[9] W. S. Saba, "Stochastic LLMs Do Not Understand Language: Towards Symbolic, Explainable and Ontologically Based LLMs," 2023.
[10] F. van Harmelen, M. Grobelnik, and B. Vallejo, "A Boxology of Design Patterns for Hybrid Learning and Reasoning Systems," Journal of Web Engineering, vol. 18, no. 7, pp. 567–599, 2019.
[11] H. F. Witschel, K. Hinkelmann, and A. Martin, "Visualization of Patterns for Hybrid Learning and Reasoning With Human Involvement," 2021.
[12] R. A. Buchmann and D. Karagiannis, "Model-Aware Software Engineering: A Knowledge-Based Approach to Model-Driven Software Engineering," in Domain-Specific Conceptual Modeling. Cham, Switzerland: Springer, 2022.
[13] D. Rordorf, J. Käser, A. E. Crego Corot, and E. Laurenzi, "A Hybrid Intelligent Approach Combining Machine Learning and a Knowledge Graph to Support Academic Journal Publishers Addressing the Reviewer Assignment Problem," in Proc. AAAI Spring Symposium on Combining Machine Learning and Knowledge Engineering (AAAI-MAKE), 2023, pp. 1–19.
[14] D. Karagiannis, "Conceptual Modelling Methods: The AMME Agile Engineering Approach," in Domain-Specific Conceptual Modeling: Concepts, Methods and ADOxx Tools. Cham, Switzerland: Springer, 2022, pp. 3–21.