Issues in Information Systems
Volume 25, Issue 4, pp. 299-310, 2024

DOI: https://doi.org/10.48009/4_iis_2024_124

Exploring large language models for ontology learning

Olga Perera, *Dakota State University, olga.perera@trojans.dsu.edu*
Jun Liu, *Dakota State University, jun.liu@dsu.edu*

Abstract

Ontology Learning aims to facilitate automatic or semi-automatic ontology development based on machine learning techniques in context of big data. Recent evolution of technology has introduced Generative Artificial Intelligence (AI) capable of creating new data, extracting insights from the existing data, and generating coherent texts from various inputs. This ability supports analysis of text data, providing insights and annotations that reduce human effort. This study explores the emerging field of Generative AI, specifically, Large Language Models for ontology learning. We conducted a survey of the current state of Generative AI research with focus on applicability and efficacy for ontology development tasks, and assessment of evaluation techniques. We discussed challenges related to explainability and interpretability of Generative AI and outlined directions for future research.

Keywords: large language models, LLM, Generative AI, ontology learning, deep learning

Introduction

Ontologies are defined as a set of primitives to model a domain of knowledge or discourse, which are presented by classes (or sets), attributes (or properties), and relationships (Alqahtani & Rilling, 2017). Ontology of a domain contains four key components: (1) concepts - abstract or concrete entities derived from specific instances or occurrences (2) attributes - characteristics of the concepts (3) taxonomy – structure that provides hierarchical relations between the concepts (4) non-taxonomic relations – specific non-hierarchical semantic relationships between the concepts (Punuru, 2007). Some ontologies are generic and cover many areas, while others are more specific to a domain. Traditional manual methods of ontology development are labor-intensive and time-consuming, often unable to keep pace with the rapidly evolving digital landscape (Yang et al., 2021).

Ontology Learning (OL) is a basic unit of ontology development that aims to facilitate automatic or semi-automatic ontology construction. OL utilizes Machine Learning (ML) techniques, generally in the context of "big data" (Mahmoud et al., 2018). The intersection of ML and OL has emerged as a pivotal area of research, allowing for more efficient, accurate automatic extraction of semantic structures from large sets of documents. The incorporation of Natural Language Processing (NLP) techniques, such as tokenization, part-of-speech tagging, and named entity recognition, serve as foundational tools in identifying and categorizing key terms and relationships in texts (Hari & Kumar, 2023; Yang et al., 2021).

NLP techniques facilitate the understanding of context and semantic meaning, thereby enriching ontologies with more accurate and contextually relevant information (Kersloot et al., 2020). The effectiveness of NLP in ontology learning is further underscored by its ability to discern and capture idiomatic expressions, technical jargon, and domain-specific terminologies. A handful of automatic ontology generation tasks use

supervised learning to extract concepts or keywords. For instance, Support Vector Machine (SVM) is considered as one of the most used and reliable algorithms that provides data analysis for classifications and regressions. Unsupervised learning, on the other hand, thrives in exploratory settings where no labeled data is available. It uncovers hidden patterns and relationships in data, making it ideal for initial ontology construction or discovering relationships between concepts. Semi-supervised learning, a blend of the two, utilizes a small amount of labeled data to guide the learning process on a larger unlabeled dataset, offering a balanced approach for ontology development where partial knowledge is available.

Deep Learning (DL), a subset of machine learning, attempts to build appropriate models by simulating the structure of the human brain (Ren et al., 2021). DL role in ontology development is primarily centered around the automation and refinement of ontological structures. DL approaches have gained significant popularity across various domains, pointing to notable advantages over conventional ontology engineering tools (He et al., 2023). Active Learning (AL) is known as query learning (Ren et al., 2021). In AL, a small set of labeled data is used in a model, then the model selects the unlabeled data and sends queries to label data actively. These new labeled data can be used to do another round of learning. However, Ren et al. (2021) argue that the classic AL algorithm is challenged by high-dimensional data. The combination of DL and AL, often referred to as deep active learning, has been recognized to achieve better results. This approach has been used for numerous ML tasks, including text classification.

Performance of ML models is typically measured using precision, recall, and F1 score, which collectively evaluate model's ability to correctly identify relevant ontology components while minimizing the inclusion of irrelevant or incorrect content. Precision represents exactness and Recall represents completeness of all concepts and properties in the learned ontologies. F-value is the weighted average of the precision and recall (Mahmoud et al., 2018). Recent developments in this field have seen the integration of detailed evaluation metrics and techniques, such as incorporating domain-specific validation and user feedback, to further enhance the quality and applicability of the generated ontologies (Poveda-Villalón et al., 2022; Rane et al., 2023).

The debate on the efficacy of different ML approaches in ontology learning is ongoing, with researchers exploring the optimal balance between automation and human involvement. While supervised learning provides a high degree of accuracy in well-understood domains, its dependence on labeled data is a limitation in areas where such data is scarce or expensive to obtain. Unsupervised and semi-supervised methods, while more flexible, often require additional post-processing to refine and validate the generated ontologies, highlighting the need for continued human oversight. The ontology development process is long and difficult, as it usually requires domain experts to be involved in the definition of the knowledge to be modeled (Spoladore & Pessot, 202). By combining computational efficiency with human insights, we ensure the developed ontologies are not only technically sound but also practically meaningful.

The emerging field of Generative AI offers promising avenues to address these challenges. However, their practical application in ontology learning is fraught with complexities. One major issue is the accurate capture and interpretation of industry language and specialized terminologies in documents. The complex nature of such texts, often laden with ambiguity, poses significant hurdles in ontology development (Hari & Kumar, 2023). In addition, potential misinterpretation and incorrect inferences, amplification of biases, and challenges related to semantic inconsistencies and ambiguities, present certain risks when employing these technologies (Božić, 2023). LLMs have shown great promise in aiding us in various informational tasks, however, their usage must be done with responsibility and sufficient validation (Shah et al., 2023). Determining how these technologies can effectively extract relevant information and represent underlying relationships and hierarchies in domain-specific ontologies remains a pivotal research question.

Research Objectives

This study aims to explore the emerging field of Generative AI, specifically, to review LLM-based approaches for ontology learning from text-based data. The first objective is to evaluate the applicability of Generative AI and LLMs in ontology learning. This involves an in-depth assessment and review of LLM approaches in prior research. The second objective focuses on assessing the efficacy of LLMs in ontology learning. The goal is to determine how LLMs can be effectively used to generate ontologies in automatic and semi-automatic ways. We aim to assess how Generative AI capabilities can be tailored to meet the domain specific needs. Evaluating LLM-generated ontologies is another critical objective. This involves reviewing these mechanisms to ensure quality, accuracy, and reliability of ontologies generated through Generative AI-driven methods, thereby mitigating the risks of biases and errors.

Finally, the research aims to contribute to the body of knowledge by reviewing methodologies and empirical evidence on the use of Generative AI and LLM in ontology learning, which is not only addresses the existing gaps in ontology learning research but also paves the way for future research and development in the state-of-the-art AI field. To fulfill these objectives, we formulated following research questions:

RQ1. *What LLM approaches for Ontology Learning have been proposed in prior research?*

RQ1-1. *Can LLMs be adapted for Ontology Learning (applicability)?*

RQ1-2. *How were these LLM-based approaches evaluated?*

RQ1-3. *What is the efficacy of the proposed approaches?*

Methodology

To address research objectives, we conducted a literature search in arXiv digital library (arxiv.org), which is an open-access archive for scholarly articles in multiple disciplines, including physics, mathematics, computer science, statistics, etc. Due to the novelty of the research topic, we expanded our literature search with the following search phrases: “ontology learning LLM”, “ontology generation LLM”, and “ontology engineering LLM”.

The PRISMA approach for systematic literature review (Page et al., 2020) was employed to the search and selection process (presented in Figure 1). The PRISMA methodology consists of the checklist for reporting of systematic reviews focused on evaluating interventions. However, many items are applicable to systematic reviews with other objectives and applicable to quantitative and qualitative studies.

The selection process was based on multiple exclusion criteria as follows: (1) research that does not focus on textual data (2) research that does not use LLMs-based OL. The initial search has resulted in 69 articles, combined, for all search phrases. During the selection process, we excluded 24 duplicate records. Title and abstract screening comprised of close examination of each article’s title and abstract, which resulted in exclusion of irrelevant studies according to the exclusion criteria. Next, we performed full-text screening, which resulted in the final selection of 3 relevant articles.

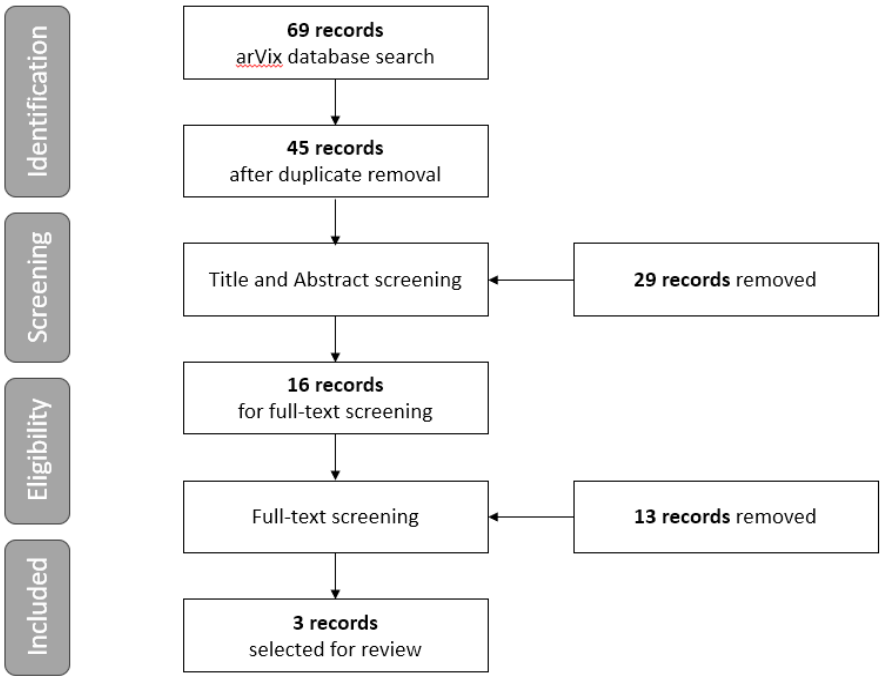

Figure 1. Selection Procedure

Results

Generative AI and Large Language Models

Generative AI represents a groundbreaking paradigm in computational science, generative in nature, functioning with a purpose of producing artifacts such as different types of text, images, audio, animations, or videos. The primary attribute distinguishing generative AI from classical AI/ML models lies in its capacity for creativity, enabling it to produce new, contextually relevant information based on learned patterns in data. At the core of Generative AI technologies are Foundation Models (FM). A foundation model is any model trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks (Bommasani et al., 2022), for example, encoder-decoder models, generative adversarial networks, and transformer models that are probabilistic in nature and capable of producing multiple, distinct outputs for a user's input. (Weisz et al., 2023).

Foundation models consist of two categories: Encoder-only or Encoder-Decoder (BERT style) and Decoder-only (GPT style) (Yang et al., 2023). Bommasani et al. (2022) stated that FMs are enabled by transfer learning approach and scale. Recent developments in generative AI have shown more sophisticated and context-aware models. Large Language Models are a subset of FMs applied for natural language processing capable of understanding and generating human language. (Chang et al., 2023). Prior research has covered multiple areas of Generative AI technology. A handful of researchers focused on investigation of different LLMs, for instance, GPT model was trained with 175 billion parameters, while other researchers explored into prompt generations and fine-tuning. Summary of LLMs adopted from Yang et al. (2023) is presented in Table 1.

Table 1. Summary of LLMs (Yang et al., 2023)

LLM type	Characteristics	Available LLMs
Encoder-Decoder or Encoder-only (BERT)	*Training*: Masked Language Models *Model type*: Discriminative *Pretrain task*: Predict masked words	ELMo, BERT, RoBERTa, DistilBERT, BioBERT, XLM, Xlnet, ALBERT, ELECTRA, T5, GLM, XLM-E, ST-MoE, AlexaTM
Decoder-only (GPT)	*Training*: Autoregressive Language Models *Model type*: Generative *Pretrain task*: Predict next word	GPT-3, OPT. PaLM, BLOOM, MT-NLG, GLaM,Gopher, chinchilla, LaMDA, GPT-J, LLaMA, GPT-4, BloombergGPT

LLMs can perform a wide range of tasks, including extracting structured knowledge from text, inferring patterns from simple word connections, encoding language semantics, and generating new text, which is crucial for OL. Giglou et al. (2023) stated that LLMs are trained on extensive and diverse text similar to domain-specific knowledge bases. Current LLMs are typically produced by Transformers with decoder-only models (e.g., GPT, Palm, Llama) becoming ever more dominant. A typical LLM has (hundreds of) billions of parameters and is able to infer and reproduce information from self-supervised training (Bakker et al., 2023).

Thus, LLMs can better comprehend the meaning of unstructured text and produce human-like response across various tasks. LLMs obtain exceptional performance on a variety of tasks along with other pre-trained language models that demonstrated compelling performance on generating responses by leveraging the pre-train / fine-tune paradigms (Varshney et al., 2023). Table 2 presents an overview of the current LLMs adopted from Carugati (2023). The core difference between traditional and generative AI is in knowledge extraction mechanism. Traditional models rely on testing specifically trained tasks while LLMs rely on testing their ability to generate responses.

Table 2. LLMs overview (Carugati, 2023)

Model	Release date	Developer	Type	Permitted use
Bloom	2022	Big Science	Open source	Commercial and non-commercial with restrictions
GPT-4	2023	Open AI	Closed source	Non-commercial and commercial
PaLM	2022	Google	Closed source	Non-commercial and commercial
LLaMA	2023	Meta	Closed source	Non-commercial
ERNIE 3.0 Titan	2021	Baidu	Open source	Non-commercial and commercial
Wu Dao 2.0	2021	BAAI	Open source	Non-commercial and commercial
YaLM	2022	Yandex	Open source	Non-commercial and commercial
Claude	2023	Anthropic	Closed source	Non-commercial and commercial
Amazon Titan FMs	2023	Amazon	Closed source	Non-commercial and commercial
Jurassic-2	2023	AI21 labs	Closed source	Non-commercial and commercial

LLM approaches in Ontology Learning

Currently, there is limited research that explicitly addresses application of LLMs in ontology learning. During literature search and selection, we discovered 3 research papers each addressing different aspects of ontology learning and offering unique perspective on OL using LLMs. Table 3 presents a summary of these approaches.

Table 3. Summary of LLM-based approaches in OL

Article	RQ1: OL approach	RQ1-1: Applicability in OL	RQ1-2: Evaluation	RQ1-3: Efficacy
LLMs4OL: Large Language Models for Ontology Learning (Giglou et al.,2023)	LLMs4OL ontology learning framework	Researchers report promising results, suggest task-specific finetuning	Use of multiple domains: lexicosemantic, geographical, biomedicine, web content representation	Finetuning outperforms models with more parameters
Domain Knowledge Distillation from Large Language Models: An Empirical Study in the Autonomous Driving Domain (Tang et al., 2023)	Automated and semi-automated framework for domain knowledge distillation	Fully automated ontology distillation process is possible but may lead to irrelevant ontology	Evaluated via multiple observations	Effective with proper design of prompts and manual supervision
Dynamic Retrieval Augmented Generation of Ontologies using Artificial Intelligence (DRAGON-AI) (Toro et al., 2023)	Ontology generation method employing LLMs and Retrieval Augmented Generation (RAG)	Feasibility of incorporating AI into ontology development workflows, but with caution	Evaluated by comparing multiple ontologies using model accuracy metrics	Best performing model-GPT4. The results are generally correct, but may be incomplete

LLMs4OL

Giglou et al. (2023) introduced LLMs4OL paradigm to test Large Language Models for ontology learning for the first time. They presented a conceptual framework for ontology learning in various knowledge domains, using LLMs to obtain high-quality results, and utilizing manual ontology validation by domain experts. LLMs4OL framework consists of the following OL tasks: (1) corpus selection and preparation (2) terminology extraction - identifying and extracting relevant terms (3) term typing – grouping similar terms into concepts (4) taxonomy construction – establishing "is-a" hierarchy (5) relationship extraction – identifying semantic relationships (6) axiom discovery – finding constraints and rules for the ontology.
LLMs4OL framework was extensively tested on 11 LLMs across following OL tasks: (1) Task A – term typing (2) Task B – taxonomic hierarchical relation prompts (3) Task C – non-taxonomic hierarchical relation prompts. Comparison of different LLM models on OL tasks was performed to support LLMs4OL evaluation by assessing model performance, identifying areas for model improvements, and finding new areas of research. The evaluation of the results indicates applicability of LLMs in OL, but there is a necessity for additional task-specific finetuning. The authors proposed to adopt FLAN collection (Flan-T5 LLM) as a method for instruction, which resulted in output improvements across all tasks.

Domain Knowledge Distillation

Tang et al. (2023) presented an automation and semi-automation framework for domain knowledge distillation using prompt engineering and ChatGPT LLM to construct a driving scenario domain ontology. This framework aims to "distil" the domain knowledge and to encapsulate all relevant domain entities, their relationships, and associated events and activities, to generate a scenario. The authors referred to it as a scenario ontology. The "Domain Knowledge Distillation" framework proposed in this research consists of three components: (1) Task Workflow (2) Prompt Engineering (3) Execution Loop. Experiments with ChatGPT were performed to identify a range of concepts, hierarchies, relationships, and concept properties, which formed the basis for domain ontology. Prompt engineering consisted of three parts: (1) domain context (why) (2) task instruction (what) (3) response format (how). Each execution loop involved prompt generation, response processing, and ontology updating, which continued until the "stop" criteria is met. This loop mechanism is intended to improve distillation quality and lessen the butterfly effect due to manual supervision and early optimization.

Domain knowledge distillation approach was applied to the road traffic domain. The researchers utilized OpenXOntology as a seed ontology. Multiple iterations (execution loops) were performed and ChatGPT responses were observed and evaluated by this research team. Based on the results of this study, automated ontology distillation is possible but may lead to unpredictable and irrelevant results in the ontology generation process. Manual supervision and early adjustments might be required to produce accurate domain ontology.

DRAGON-AI

Toro et al. (2023) presented Dynamic Retrieval Augmented Generation of Ontologies using AI (DRAGON-AI) method for ontology development using LLMs and Retrieval Augmented Generation (RAG). This method can generate textual and logical ontology components from existing ontological knowledge and unstructured textual sources. The main goal is to produce various requisite parts of ontology such as textual description and relationships based on specific concepts (label, name, definition). In other words, DRAGON-AI assists in AI-based auto-completion of ontology objects. Initially, all ontology terms and additional context are indexed into JSON format with predefined schema. Then LLM is used to generate different ontological components. During this stage, RAG is used for prompt generation to provide additional context, for example, existing relevant terms, documentation written for ontology developers, or related articles. Next, prompt is passed to LLM, which produces a response that is parsed into JSON parser, then after additional processing, the results are merged with the input object to form the predicted object.
DRAGON-AI evaluation was performed against 10 different ontologies in a broad range of domains. The core ontology and testing set of 50 terms were used to evaluate ontology generation tasks such as relationships, definitions, and logical definitions. The following LLMs are used for testing purposes: GPT-4, GPT-3.5 turbo, nous-hermes-13b-ggml. The evaluation of DRAGON-AI method offers valuable insights and feasibility of using Generative AI for ontology development. For instance, relationship generation task demonstrated high precision, but moderate recall/F1 scores, which indicates that results might be incomplete. For general term definitions, LLM generated definitions rank close but lower than human authored ones. In conclusion, the researchers stated that AI generated content should be used with caution.

Discussion

Key findings from this literature review underline the importance of research in Generative AI. Recent developments have shown a growing trend towards more sophisticated and context-aware models. LLMs can comprehend the meaning of unstructured text and produce human-like responses across various tasks,

offering remarkable potential in ontology learning field. Introduction of LLMs continues to prove to be more efficient and scalable compared to traditional ML and manual methods, which, although accurate, are time-consuming and labor-intensive. AI-based methods excel in handling large, diverse datasets and exhibit adaptability to evolving domain knowledge, a critical requirement in many rapidly advancing industries.

The results of this study indicate feasibility of LLM-based approaches. Debates and discussions within the academic community highlight the transformative impact of LLMs in many areas of research. While there is a consensus on its potential, there are also concerns regarding the quality and reliability of AI-generated content. Some critics argue that LLMs rely on statistical patterns rather than true understanding and reasoning, and may generate plausible but incorrect responses, such as hallucinations Pan et al. (2023). Another debate progresses around biases. Yang et al. (2023) argued that LLMs are susceptible to majority label, positional, and common token biases.

However, recent studies show, for instance, that positional bias can be mitigated by selecting proper prompts. On the other hand, proponents argue that bias is not inherent to LLMs but embedded in the data. In regards of OL research, Toro et al. (2023) indicated that novice ontology editors might be "tricked" by AI if they had lower confidence in the domain, accepting AI generated responses on face value. Giglou et al. (2023) stated that integrating human-in-the-loop approaches with expert involvement will enhance ontology relevance and accuracy.

Tang et al. (2023) argued that due to randomness in the response and butterfly effect, the quality of generated content is not guaranteed. This indicates that fully automated LLM-based approaches in ontology learning are feasible but challenging, necessitating human oversight at the current state of Generative AI. As AI technologies continue to advance, so will the strategies for overcoming these challenges.

Explainability and interpretability further challenge AI-based techniques. Skeptics of LLMs argue that these models lack transparency and interpretability, making it difficult to understand how they arrive at their answers or recommendations due to their "black box" nature. Some researchers argue that Chain-of-Thoughts (CoT) may improve the explainability of LLMs, although question decomposition and precisely answering sub-questions with LLMs are still far from being solved. Proponents of LLMs acknowledge the challenge of explainability but argue that recent research efforts are improving LLMs through techniques like attention mechanisms and model introspection (Pan et al., 2023).

Proponents emphasize the efficiency gains and the ability of AI to stay abreast of the latest developments, which is critical in rapidly evolving fields (Yang et al., 2023). Since it is not possible to have full access to the model, there is a risk that the model can be changed. Another challenge is in the availability of intermediate scores that cannot be retrieved (Hertling & Paulheim, 2023).

These debates reflect the ongoing efforts to strike a balance between automation and accuracy, underscoring the need for continuous refinement of AI algorithms and models. Recent research and developments reflect an evolving effort to address these challenges, highlighting the dynamic and progressive nature of Generative AI technology.

Figure 2 outlines key areas for future research.

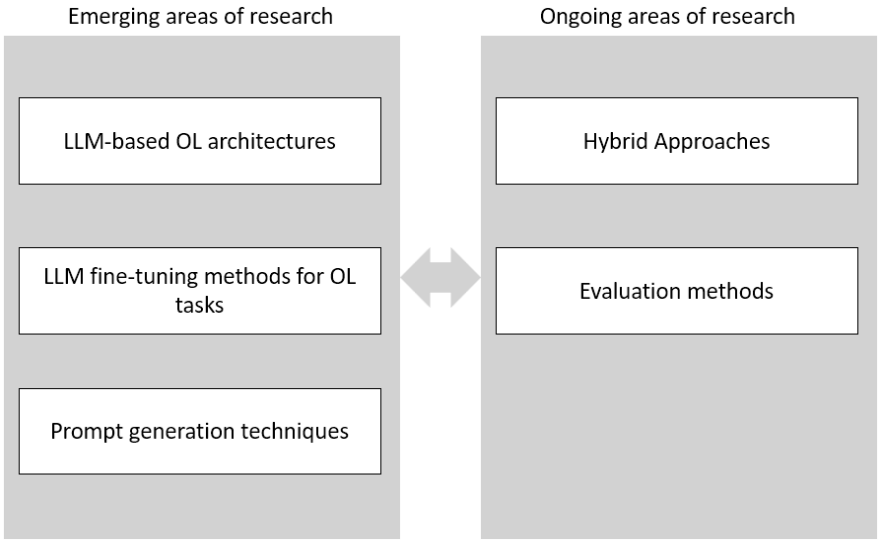

Figure 2. Key research areas

There are two primary areas for future research, the emerging field of research based on novel aspects of generative AI technology, and ongoing research based on enhancing existing ontology learning research. First, future research can focus on the development of novel LLM-based OL conceptual frameworks, architectures, and methods to enhance resulting ontological structures and improve specific areas of ontology learning process. Second, prior academic literature outlines the emergence of prompt engineering aspect. Prompt engineering focuses on designing optimal instructions for LLMs, which in some cases, include processing step-by-step thought process to arrive at a relevant answer. Prompt engineering can aid in the development of techniques that result in more accurate subsequent ontology created by Generative AI model. Third, future research can investigate LLM fine-tuning techniques to assist in ontology learning using large language models trained on domain specific data.

Development of robust evaluation techniques is another important area of research as it provides researchers and industry leaders with methods to assess their ontology-based systems in a specific domain setting. Standardizing these evaluation metrics as well as developing hybrid approaches are important research directions. Expanding these areas will assist in advancing existing research in OL leading to advancements across multiple domains and organizational enterprises. Continued exploration in the Generative AI area is vital in the field of ontology learning, ultimately contributing to the broader landscape of artificial intelligence and its application in various domains.

Conclusion

Key findings from this literature review underline the importance of research in the Generative AI-driven era. The introduction of LLMs continues to prove to be more efficient, scalable and excel in handling large, diverse datasets. However, these approaches are not without challenges with implications that are profound and multifaceted. Future research is poised to delve into making Generative AI-driven methods more transparent, thereby enhancing their reliability and acceptability in critical applications.

In closing, this study reveals both the strides made and the challenges that persist in the field of ontology learning. The transition from traditional to AI-driven methods marks a significant evolution, opening new possibilities for efficient and adaptive knowledge representation. However, the journey is far from complete.

References

Abasi, A. K., Khader, A. T., Al-Betar, M. A., Naim, S., Makhadmeh, S. N., & Alyasseri, Z. A. (2020). A novel ensemble statistical topic extraction method for scientific publications based on optimization clustering. *Multimedia Tools and Applications*, *80*(1), 37–82. https://doi.org/10.1007/s11042-020-09504-2

Achachlouei, M. A., Patil, O., Joshi, T., & Nair, V. N. (2023, August 18). *Document automation* architectures*: Updated survey in light of large language models*. arXiv.org. https://arxiv.org/abs/2308.09341

Allen, B. P., Stork, L., & Groth, P. (2023, October 1). *Knowledge engineering using large language models*. arXiv.org. https://arxiv.org/abs/2310.00637

Alqahtani, S. S., & Rilling, J. (2017). An ontology-based approach to automate tagging of software artifacts. *2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*. https://doi.org/10.1109/esem.2017.25

Azad, B., Azad, R., Eskandari, S., Bozorgpour, A., Kazerouni, A., Rekik, I., & Merhof, D. (2023, October 28). *Foundational models in Medical Imaging: A comprehensive survey and future vision*. arXiv.org. https://*arxiv*.org/abs/2310.18689

Bakker, R. M., Kalkman, G. J., Tolios, I., Blok, D., Veldhuis, G. A., Raaijmakers, S., & de Boer, M. H. T. (2023). Exploring knowledge extraction techniques for system dynamics modelling ... https://bnaic2023.tudelft.nl/static/media/BNAICBENELEARN_2023_paper_68.28ca9f2727dc77fc1bca.*pdf*

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., … Liang, P. (2022, July 12). *On the opportunities and risks of* Foundation *models*. arXiv.org. https://arxiv.org/abs/2108.07258

Bontcheva, K., & Sabou, M. (2006, January). Learning Ontologies from Software Artifacts: Exploring and Combining Multiple Sources. http://km.aifb.kit.edu/ws/swese2006/final/bontcheva_full.pdf

Božić, V. (2023, May). Semantic web and generative pre-trained transformer (GPT). https://www.researchgate.net/publication/370818668_Semantic_Web_and_Generative_Pre-trained_Transformer_GPT

Carugati, C. (2023). Competition in Generative AI Foundation models. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4553787

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., Wang, Y., Ye, W., Zhang, Y., Chang, Y., Yu, P. S., Yang, Q., & Xie, X. (2023, December 29). *A survey on evaluation of large language models*. arXiv.org. https://arxiv.org/abs/2307.03109

Giglou, H. B., D'Souza, J., & Auer, S. (2023, August 2). *LLMs4OL: Large language models for ontology learning*. arXiv.org. https://arxiv.org/abs/2307.16648

Hari, A., & Kumar, P. (2023). WSD based ontology learning from unstructured text using Transformer. *Procedia Computer Science*, *218*, 367–374. https://doi.org/10.1016/j.procs.2023.01.019

He, Y., Chen, J., Dong, H., Horrocks, I., Allocca, C., Kim, T., & Sapkota, B. (2023, July 6). *DeepOnto: A python package for Ontology Engineering with deep learning*. arXiv.org. https://arxiv.org/abs/2307.03067

Hertling, S., & Paulheim, H. (2023). Olala: Ontology matching with large language models. *Proceedings of the* 12th *Knowledge Capture Conference 2023*. https://doi.org/10.1145/3587259.3627571

Kersloot, M. G., van Putten, F. J., Abu-Hanna, A., Cornet, R., & Arts, D. L. (2020). Natural language processing algorithms for mapping clinical text fragments onto ontology concepts: a systematic review and recommendations for future studies. *Journal of biomedical semantics*, *11*, 1-21.

Mahmoud, N., Elbeh, H., & Abdlkader, H. M. (2018). Ontology learning based on word embeddings for text big data extraction. *2018 14th International Computer Engineering Conference (ICENCO)*. https://doi.org/10.1109/icenco.2018.8636154

Page, M. J., McKenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., mulrow, cindy, Shamseer, L., Tetzlaff, J., Akl, E., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., … Moher, D. (2020). The Prisma 2020 statement: An updated guideline for reporting systematic reviews. https://doi.org/10.31222/osf.io/v7gm2

Pan, J. Z., Razniewski, S., Kalo, J.-C., Singhania, S., Chen, J., Dietze, S., Jabeen, H., Omeliyanenko, J., Zhang, W., Lissandrini, M., Biswas, R., de Melo, G., Bonifati, A., Vakaj, E., Dragoni, M., & Graux, D. (2023, August 11). *Large language models and knowledge graphs: Opportunities and challenges*. arXiv.org. https://arxiv.org/abs/2308.06374

Poveda-Villalón, M., Fernández-Izquierdo, A., Fernández-López, M., & García-Castro, R. (2022). LOT: An industrial oriented ontology engineering framework. *Engineering Applications of Artificial Intelligence*, *111*, 104755.

Punuru, J. (n.d.). Knowledge-Based Methods for Automatic Extraction of Domain-Specific Ontologies. https://doi.org/10.31390/gradschool_dissertations.708

Rane, N., Choudhary, S., & Rane, J. (2023). Integrating ChatGPT, Bard, and leading-edge generative artificial intelligence in architectural design and engineering: applications, framework, and challenges.

Ren, X., Li, X., Ren, K., Song, J., Xu, Z., Deng, K., & Wang, X. (2021). Deep learning-based Weather prediction: A survey. *Big Data Research*, *23*, 100178. https://doi.org/10.1016/j.bdr.2020.100178

Shah, C., White, R. W., Andersen, R., Buscher, G., Counts, S., Das, S. S. S., Montazer, A., Manivannan, S., Neville, J., Ni, X., Rangan, N., Safavi, T., Suri, S., Wan, M., Wang, L., & Yang, L. (2023, September 14). Using *large language models to generate, validate, and apply user intent taxonomies*. arXiv.org. https://arxiv.org/abs/2309.13063

Spoladore, D., & Pessot, E. (2021). Collaborative ontology engineering methodologies for the development of decision support systems: Case studies in the healthcare domain. *Electronics*, *10*(9), 1060.

Tang, Y., da Costa, A. A. B., Zhang, J., Patrick, I., Khastgir, S., & Jennings, P. (2023, July 17). Domain knowledge distillation from large language model: An empirical study in the autonomous driving domain. arXiv.org. https://arxiv.org/abs/2307.11769

Toro, S., Anagnostopoulos, A. V., Bello, S., Blumberg, K., Cameron, R., Carmody, L., Diehl, A. D., Dooley, D., Duncan, W., Fey, P., Gaudet, P., Harris, N. L., Joachimiak, M., Kiani, L., Lubiana, T., Munoz-Torres, M. C., O'Neil, S., Osumi-Sutherland, D., Puig, A., … Mungall, C. J. (2023, December 18). *Dynamic retrieval augmented generation of ontologies using Artificial Intelligence (Dragon-AI)*. arXiv.org. https://arxiv.org/abs/2312.10904

Varshney, D., Zafar, A., Behera, N. K., & Ekbal, A. (2023). Knowledge grounded medical dialogue generation using augmented graphs. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-29213-8

Weisz, J. D., Muller, M., He, J., & Houde, S. (2023, January 13). *Toward general design principles for* Generative *AI applications*. arXiv.org. https://arxiv.org/abs/2301.05578

Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., Yin, B., & Hu, X. (2023, April 27). *Harnessing the power of LLMS in practice: A survey on CHATGPT and beyond*. arXiv.org. https://arxiv.org/abs/2304.13712

Yang, L., Cormican, K., & Yu, M. (2021). Ontology learning for systems engineering body of knowledge. *IEEE Transactions on Industrial Informatics*, *17*(2), 1039–1047. https://doi.org/10.1109/tii.2020.2990953