

Ontology-Based Prompt Engineering: A Narrative Review

Jin Ma¹, Ehab Seedahmed², Essa Alyounis³, Mohammad Nazir Ahmad⁴

Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Bangi, Malaysia^{1,4}

Health Information Management and Technology Department, King Faisal University, Al-Hofuf, Saudi Arabia^{2,3}

Abstract—Large Language Models (LLMs) have significantly advanced Prompt Engineering (PE); however, current prompting strategies still rely heavily on unstructured natural language and lack explicit semantic guidance. Ontology has recently emerged as a promising approach for enhancing PE by providing structured domain knowledge and reasoning capabilities, yet existing studies remain fragmented across application domains and lack a unified theoretical perspective. This narrative review examines six core studies published between 2024 and 2026, selected from literature covering 2020 to 2026, to investigate different ontology–LLM interaction patterns in prompting-related applications. The studies are compared according to their use of ontology as input knowledge, structural guidance, or an output target, as well as their prompting strategies and evaluation evidence. The review synthesizes four ontology knowledge capabilities—knowledge representation, concept hierarchy, semantic relationships, and logical reasoning—and introduces a study-by-capability evidence matrix to distinguish conceptual relevance from available empirical support. Five research gaps and corresponding future research directions are identified. The Prompt Canvas serves as a conceptual reference; the mapping proposed in this review between its elements and ontology capabilities remains unvalidated.

Keywords—Large language models; ontology; prompt engineering

I. INTRODUCTION

Recent advances in transformer-based architectures and large-scale pretraining have enabled LLMs to become one of the most influential technologies in artificial intelligence (AI), driving significant progress across a wide range of research and application domains. Trained on massive-scale corpora, LLMs exhibit remarkable emergent capabilities, including advanced pattern recognition, extensive world knowledge, and the ability to understand, generate, and reason with natural language. Recent models, such as GPT-4, can produce high-quality, contextually appropriate, and human-like text, representing a significant milestone in the development of AI. These capabilities have supported applications across a range of domains, including ontology-informed construction activity inference from images [1]. More broadly, LLMs are also being explored for language generation, information processing, and decision-support tasks. As their reasoning and language understanding continue to improve, LLMs are becoming increasingly versatile and widely adopted tools for both academic research and real-world problem solving [2].

As LLMs continue to demonstrate increasingly sophisticated reasoning and generation capabilities, effectively interacting

with these models has become an important research issue. PE plays an important role in effectively leveraging LLMs and generative AI across diverse applications, acting as a crucial bridge between human intent and machine output [3], [4]. Its importance stems from its ability to significantly enhance LLM performance and the quality of generated content, enabling models to produce precise, coherent, pertinent, and accurate responses for specific tasks [5]. Compared with parameter-intensive fine-tuning approaches, PE offers a more cost-effective, flexible way to adapt LLMs to downstream tasks. Furthermore, advanced prompting strategies, such as Chain-of-Thought (CoT) prompting, improve reasoning performance by decomposing complex problems into intermediate reasoning steps, thereby enhancing task accuracy [6].

Despite its remarkable success, PE still faces several fundamental challenges that limit its robustness, scalability, and generalizability across different tasks and application domains. Previous literature has identified several recurring challenges in PE, including data and resource constraints, model dependency, task complexity, user-related barriers, ethics and safety, and evaluation limitations [2], [4]–[10]. Collectively, these challenges indicate that current PE still faces difficulties in systematically designing, evaluating, and transferring prompts across tasks and application domains. Many existing approaches continue to rely on manually designed prompts and empirical trial-and-error strategies, which may result in inconsistent prompt quality and limited transferability. Such approaches often lack explicit semantic guidance and structured domain knowledge.

Ontologies provide explicit conceptual representations, structured semantic relationships, and domain knowledge that may be incorporated into LLM prompting and other ontology-related tasks. Depending on the application, an ontology may serve as an input knowledge source, a semantic constraint, or a target representation generated or processed by an LLM. Whether ontology integration improves task performance, reasoning consistency, or interpretability depends on the specific system design and available comparative evidence. In recent years, researchers have increasingly explored the integration of ontology with PE to enhance the performance, interpretability, and domain adaptability of LLMs. Despite the emerging body of research, existing studies of ontology–LLM interactions in prompting-related applications [1], [8], [11]–[14] remain fragmented in terms of research objectives, application domains, and methodological approaches. In addition, structured prompt engineering frameworks have emerged to organize prompt components and prompting techniques, but

their relationship with ontology-based semantic knowledge remains insufficiently explored. More importantly, the selected studies primarily address individual application scenarios and provide limited evidence explaining when and how different ontology roles contribute to prompting-related task outcomes. Furthermore, the selected studies do not provide a shared conceptual framework or coordinated research agenda for organizing these different ontology–LLM interaction patterns. To address these gaps, this study conducts a narrative literature review to synthesize and critically interpret existing ontology-guided PE studies while considering structured prompt engineering frameworks as a conceptual reference. Specifically, this review aims to identify the major research themes and application patterns, develop a capability-oriented analytical framework explaining the role of ontology in PE, and propose a future research agenda to support subsequent theoretical development and practical implementation.

II. BACKGROUND

A. What is Prompt Engineering?

PE refers to the systematic process of designing, structuring, and optimizing prompts that guide LLMs or other generative AI systems toward producing accurate, reliable, and task-oriented outputs. Unlike fine-tuning, which modifies model parameters, instruction-based PE guides model behavior through natural language instructions without updating those parameters. It involves iterative prompt construction, testing, evaluation, and refinement to improve response quality, reasoning ability, and task alignment [4], [7], [10]. PE has emerged as a lightweight yet powerful approach for adapting foundation models to downstream tasks without expensive retraining or fine-tuning [4].

B. Structured Prompt Engineering Frameworks

As prompt engineering has evolved, researchers have increasingly attempted to organize fragmented prompting techniques into structured frameworks. One representative approach is the Prompt Canvas proposed by Hewing and Leinhos, which consolidates key prompt components and prompting techniques into a unified framework for designing effective prompts. The Prompt Canvas organizes prompt construction around elements such as persona/role, audience, task and intent, context, step-by-step instructions, references, output, tonality, tooling, and recommended prompting techniques [15].

The Prompt Canvas provides a practical and structured representation of prompt construction and helps bridge the gap between fragmented prompt engineering knowledge and practical application. However, its components primarily describe what information and instructions should be incorporated into a prompt, rather than providing a formal semantic representation of the domain knowledge underlying those components. Moreover, the framework itself acknowledges the need for further quantitative and qualitative validation and emphasizes that the canvas should evolve as prompt engineering develops.

From this perspective, structured prompt engineering frameworks provide an important foundation for organizing prompt construction, but they do not fully address how domain

knowledge can be represented, related, and reasoned over within prompts. This limitation motivates the investigation of ontology as a potential semantic knowledge layer for enhancing structured prompt engineering.

To further examine this potential, the major elements of the Prompt Canvas are conceptually mapped to the ontology knowledge capabilities identified in this review. The mapping is intended as an analytical interpretation rather than a modification proposed by the original Prompt Canvas framework. It illustrates how structured ontological knowledge may enhance selected prompt components by providing explicit concepts, semantic relationships, and reasoning support. Table I summarizes these potential ontology enhancements.

TABLE I. CONCEPTUAL MAPPING BETWEEN PROMPT CANVAS ELEMENTS AND ONTOLOGY KNOWLEDGE CAPABILITIES

Prompt Canvas Element	Relevant Ontology Capability	Potential Ontology Enhancement
Persona / Role	Concept Hierarchy	Represent roles and user types at different abstraction levels
Audience	Concept Hierarchy	Organize user groups and audience characteristics hierarchically
Task / Intent	Knowledge Representation	Represent tasks, goals, and intents as structured concepts
Context	Semantic Relationships	Explicitly model contextual dependencies among concepts
References	Knowledge Representation	Organize and retrieve structured domain knowledge
Step-by-Step Instructions	Logical Reasoning	May specify ontology-based reasoning constraints when appropriate inference mechanisms are available
Output	Knowledge Representation	Represent output requirements and semantic constraints
Recommended Techniques	Semantic Relationships	Associate prompting techniques with tasks, contexts, and objectives

As shown in Table I, this review proposes that an ontology could complement the Prompt Canvas by providing structured semantic knowledge for selected prompt components. Concept hierarchies may organize roles and audiences, while knowledge representation may structure tasks, intents, references, and output requirements. Semantic relationships may make contextual dependencies explicit, and formal reasoning may support reasoning constraints when suitable inference mechanisms are available. This mapping is an analytical proposition developed in this review, and its effectiveness remains to be evaluated.

C. What Is Ontology?

Ontology has become one of the most widely adopted approaches for representing domain knowledge in AI, the Semantic Web, and knowledge engineering. With the rapid adoption of LLMs, ontology has increasingly attracted attention as a structured knowledge source capable of supporting PE. Unlike unstructured textual knowledge resources, formal ontologies explicitly represent domain concepts, relationships,

and semantic constraints in a machine-interpretable form. Such structures can support knowledge organization, semantic interoperability, and, when appropriate formal reasoning mechanisms are available, logical inference. In the reviewed literature, ontology-related structures serve different purposes, including guiding summarization [11], supporting ontology population [8], [12], and representing extracted domain knowledge [14].

From the perspective of PE, the value of ontology lies not merely in knowledge representation, but in its ability to provide structured semantic support throughout the prompting process. Specifically, ontology can be understood through four complementary knowledge capabilities that are particularly relevant to PE: knowledge representation, concept hierarchy, semantic relationships, and logical reasoning.

Knowledge representation forms the foundation of ontology. Ontologies provide a formal and machine-readable representation of domain knowledge by defining concepts, properties, relationships, and constraints. This structured representation enables knowledge to be consistently organized, shared, and reused across applications, thereby offering an explicit semantic foundation for downstream AI tasks [12], [14].

Concept hierarchy enables ontologies to organize domain knowledge at different levels of abstraction through superclass–subclass relationships. Such structures can provide different levels of conceptual specificity for ontology-related prompting and knowledge processing. However, their contribution depends on how hierarchical information is incorporated into the task. In the reviewed studies, hierarchical structures are used or examined in different contexts, including ontology design patterns for A-box extraction [8] and hierarchical correspondence identification in ontology alignment [13]. These applications do not establish a general performance benefit of hierarchy-based prompting.

Semantic relationships enable ontologies to represent explicit connections among concepts and entities, including attributes, associations, and dependencies. These relationships can provide structured contextual information for ontology-related prompting and knowledge-processing tasks. For example, semantic relationships are incorporated into ontology-informed construction activity inference [1] and domain-specific news summarization [11]. However, whether they independently reduce ambiguity or improve task performance cannot be established without comparisons that isolate their contribution from other system components.

Logical reasoning is a potential capability of formal ontologies when appropriate inference rules and reasoning mechanisms are available. Such mechanisms can derive implicit knowledge from explicitly represented concepts, relationships, and constraints. However, the use of structured ontology knowledge in LLM prompts should not be equated with the execution of formal ontology reasoning. In the reviewed studies, direct empirical evidence of formal ontology-based inference remains limited, and its contribution to prompting performance requires further investigation.

Collectively, these four knowledge capabilities distinguish ontologies from conventional knowledge resources by enabling

explicit semantic organization, contextual understanding, and logical reasoning. Rather than merely storing information, ontologies provide a structured knowledge layer capable of supporting prompt construction, semantic interpretation, and reasoning-intensive tasks. These characteristics suggest that ontology has the potential to function as an effective semantic foundation for ontology-guided PE.

D. Why Does Prompt Engineering Need Ontology?

Previous literature identified six major challenge dimensions in PE, including data and resource constraints, model dependency, task complexity, user-related barriers, ethical concerns, and evaluation limitations. Collectively, these challenges highlight that current PE is not solely constrained by prompt formulation itself, but also by the availability, organization, interpretation, and utilization of domain knowledge.

Although structured prompt frameworks such as the Prompt Canvas provide a systematic way to organize prompt components and prompting techniques, they primarily operate at the level of prompt formulation rather than formalizing the domain knowledge underlying these components. As illustrated in Table I, ontology can potentially complement such frameworks by providing structured concepts, semantic relationships, and reasoning support.

Structured knowledge grounding has also been explored in LLM-based applications without the explicit use of formal ontologies. For example, the MIND-SAFE framework [9] integrates therapeutic knowledge and retrieval-augmented generation (RAG) to support mental health chatbot interactions. However, it does not employ an explicit ontology with formally defined concepts, relationships, and logical constraints. This distinction highlights the difference between grounding prompts in domain-specific knowledge and grounding them in formal ontological structures.

From a knowledge perspective, these challenges indicate a broader need for structured semantic knowledge to support prompt construction, contextual understanding, and reasoning. Ontology addresses this need by explicitly formalizing concepts, relationships, and semantic constraints into a machine-interpretable representation. Such semantic organization provides a consistent and reusable knowledge foundation for ontology-guided PE.

Ontology provides explicit semantic representations through knowledge representation, concept hierarchy, semantic relationships, and logical reasoning. These complementary knowledge capabilities suggest that ontology may serve as a structured knowledge layer capable of supporting PE by improving semantic consistency, contextual understanding, and reasoning support. As illustrated in Fig. 1, the conceptual progression from current PE practices to ontology-guided PE is driven by the increasing need for structured semantic knowledge. The following parts therefore examine how existing studies have leveraged these ontology capabilities to support PE across different application domains.

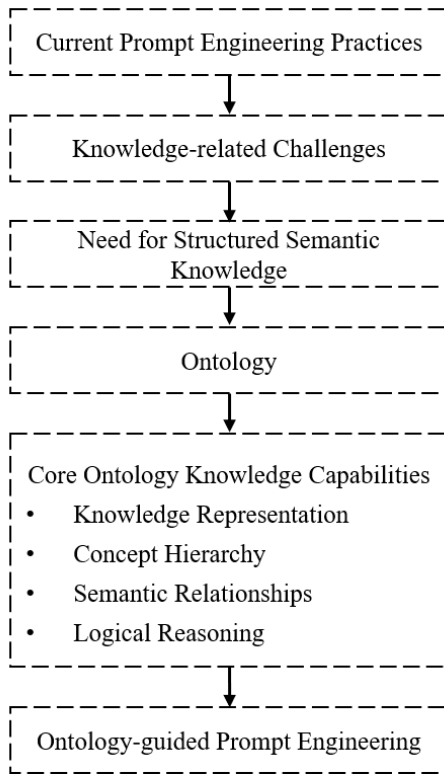


Fig. 1. Conceptual progression from conventional prompt engineering to ontology-guided prompt engineering.

III. METHODOLOGY

A. Research Design

This study adopts a narrative literature review (NLR) to synthesize and critically interpret the emerging research on ontology-guided PE. The NLR was selected for its effectiveness in synthesizing diverse sources of literature and providing a holistic understanding of an emerging research topic [16]–[18]. Unlike systematic review approaches, which typically focus on well-defined research questions and structured evidence synthesis [19], [20], NLR provides greater flexibility in exploring diverse research domains and synthesizing interdisciplinary insights [21]. This flexibility is particularly appropriate for ontology-guided PE, where existing studies remain limited and fragmented across different application domains. This approach has also been adopted in recent research to synthesize interdisciplinary literature and develop conceptual insights, supporting its suitability for emerging research areas [22].

B. Literature Search Strategy

This review searched the literature published between 2020 and 2026 using four major academic databases, including Web of Science (WoS), Scopus, IEEE Xplore, and ScienceDirect. Search strings were constructed by combining keywords related to ontology, PE, and large language models using Boolean operators (AND/OR), following the literature search strategy recommended for review studies [17]. The detailed search strategy, including the databases, search fields, keywords, publication years, and document types, is summarized in Table II.

TABLE II. REPORTED LITERATURE SEARCH SCOPE AND ELIGIBILITY PARAMETERS

Item	Description
Databases	WoS, Scopus, IEEE Xplore, ScienceDirect
Years	2020-2026
Language	English
Search Fields	Title, Abstract, Keywords
Keywords	Ontology; Prompt Engineering; Large Language Models; Ontology-guided Prompt; Ontology-enhanced Prompt; Semantic Prompting; Knowledge-guided Prompt.
Document Type	Journal Articles, Conference Papers

1) *Inclusion criteria:* The studies included in this review met the following criteria:

- Published in English between 2020 and 2026;
- Peer-reviewed Journal Articles or Conference Papers;
- Focused on Ontology-guided PE or Ontology-supported Prompting;
- Involved LLM-based Applications; and
- Explicitly Discussed the Role of Ontology in Prompt Construction, Semantic Understanding, or Reasoning.

The screening process involved title and abstract screening followed by full-text assessment against the inclusion criteria. Studies addressing prompt engineering without an ontology component, or ontology-related applications without a direct connection to LLM-based prompting, were excluded. For the analytical synthesis presented in this review, ontology-supported prompting encompasses the use of ontology knowledge in prompts, prompting for ontology population, and prompting for ontology construction or alignment. Six studies addressing these prompting-related ontology–LLM interactions were selected for synthesis. Their inclusion indicates relevance to the review’s analytical scope, rather than evidence that ontology independently improves prompting performance. Additional literature on structured prompt engineering, including the Prompt Canvas [15] and MIND-SAFE [9], was used only as supporting conceptual background and was not included in the six-study synthesis.

As this study adopts a narrative literature review approach, the six core studies were purposively selected based on their relevance to ontology–LLM interactions and their contributions to the research objectives. The selection was intended to provide a focused analytical synthesis rather than an exhaustive account of all eligible publications. The original retrieval and screening counts were not retained; therefore, a complete record-flow reconstruction and quantitative reporting of exclusion reasons were not possible. This limitation is acknowledged when interpreting the scope and generalizability of the findings.

C. Narrative Analysis

The selected studies were analyzed using a narrative conceptual synthesis approach. Rather than comparing quantitative outcomes, the analysis aimed to identify recurring conceptual patterns regarding the role of ontology in PE.

To enable a consistent cross-study comparison, each reviewed study was analyzed using a predefined set of analytical criteria, including application domain, research objective, ontology type and design, ontology role in prompt engineering, prompting strategy, research methodology, evaluation evidence, key finding, and limitation. These criteria were selected to examine not only what problems existing studies address, but also how ontology is designed and utilized to support prompt engineering across different application domains.

To ensure analytical consistency, the reviewed studies were interpreted through the four ontology knowledge capabilities introduced in Section II: knowledge representation, concept hierarchy, semantic relationships, and logical reasoning. These four capabilities served as the theoretical lens for synthesizing and interpreting the reviewed studies. Although the literature is presented on a paper-by-paper basis in Related Work, the capability-oriented analytical framework was applied in Section V to synthesize the findings across studies from a conceptual perspective rather than merely describing individual applications. This analytical perspective provides the theoretical basis for identifying research gaps and proposing a future research agenda for ontology-guided PE. The overall methodological process adopted in this review is illustrated in Fig. 2.

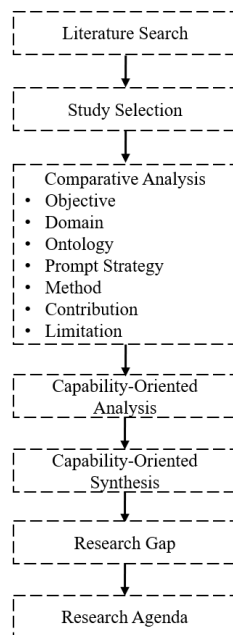


Fig. 2. Methodological framework for comparative and capability-oriented narrative analysis.

As shown in Fig. 2, the methodological process consisted of seven sequential stages: literature search, study selection, comparative analysis, capability-oriented analysis, capability-oriented synthesis, research gap identification, and research agenda development. Following this process, the reviewed studies were first systematically compared using the predefined analytical attributes and were subsequently synthesized through the capability-oriented analytical framework. This two-stage analytical process enabled both cross-study comparison and theory-driven interpretation of ontology-guided PE.

IV. RELATED WORK

A. From Prompt to Graph: Comparing LLM-Based Information Extraction Strategies in Domain-Specific Ontology Development [14]

1) *Objective*: The study [14] aims to investigate and compare three LLM-based strategies—pre-trained prompting, in-context learning (ICL), and fine-tuning—for semi-automating domain-specific ontology construction. In this study, ontology construction is the primary objective, rather than the use of an existing ontology to guide LLM prompting. Using the casting manufacturing sector as a representative case, the authors seek to transform fragmented industrial documents into structured ontological knowledge, thereby improving knowledge management, semantic interoperability, and intelligent decision-making in manufacturing.

2) *Methodology*: The study adopts a multi-stage framework consisting of data preparation, information extraction, and ontology construction. Domain knowledge is first summarized from technical literature using RAG. Three LLM-based strategies—optimized prompting with CoT reasoning, ICL, and fine-tuning—are then comparatively evaluated for extracting ontology concepts and semantic relationships. The extracted knowledge is subsequently represented as relation triples and incorporated into a casting ontology within a graph database, enabling structured knowledge organization for downstream manufacturing applications.

3) *Ontology-guided prompting strategy*: In this study, the ontology primarily serves as the target knowledge representation for LLM-assisted information extraction. Domain-specific concepts and relationships are extracted using prompting, ICL, and fine-tuning, and the resulting knowledge is organized into ontology triples within a graph database. Although predefined domain categories inform the extraction process, the study does not primarily investigate the use of an existing ontology as a knowledge source for prompt construction. Therefore, its findings provide evidence for LLM-assisted ontology construction rather than direct evidence that ontology-guided prompts improve LLM reasoning performance.

4) *Main contribution*: The principal contribution of this study is its comparative evaluation of three LLM-based information extraction approaches for ontology construction. Within the evaluated casting-domain setting, fine-tuning achieved the best overall performance in extracting domain-specific concepts and semantic relationships. However, these findings concern information extraction and ontology development, rather than the independent effect of ontology-based prompt guidance. The proposed framework successfully supports the semi-automatic development of a casting ontology by converting unstructured industrial documents into structured semantic knowledge. The study further demonstrates the feasibility of integrating PE with ontology development to enhance knowledge acquisition and intelligent manufacturing systems.

5) *Limitation*: The study is constrained by the relatively small size of the annotated dataset, which may limit the scalability and generalizability of the proposed framework across broader industrial domains. Although fine-tuning achieved the highest extraction performance, it remained susceptible to relation-label substitution errors when encountering previously unseen concepts. Furthermore, the ICL approach was highly dependent on the quality and diversity of retrieved examples, resulting in inconsistent synonym recognition and relation extraction. These limitations indicate a need for more robust and scalable LLM-assisted information extraction and ontology construction methods for industrial knowledge engineering.

B. *Ontology-Based Prompt Tuning for News Article Summarization [11]*

1) *Objective*: The study [11] investigates an ontology-based prompt tuning framework for improving domain-specific news summarization. Using sports news as the application domain, the framework incorporates ontology-derived concepts and semantic relationships into prompt generation to support contextually relevant summarization. Unlike ontology construction studies, this work uses an existing domain ontology as a knowledge source for guiding the summarization process.

2) *Methodology*: The proposed framework combines ontology-based prompt tuning with an abstractive text summarization model based on the T5 transformer. News articles are first preprocessed through text cleaning, tokenization, and lemmatization. An ontology integration module then extracts domain entities and semantic relationships, which are mapped to predefined ontology classes and properties. Based on the extracted semantic knowledge, domain-specific prompt templates are dynamically generated to guide the summarization process. The proposed framework is evaluated against several summarization models using automatic evaluation metrics. However, comparisons between different summarization architectures do not necessarily isolate the contribution of ontology-based prompt tuning. A controlled comparison using the same underlying model, with and without ontology-derived prompt information, would be needed to establish its independent effect.

3) *Ontology-guided prompting strategy*: In this framework, ontology serves as an input knowledge source for prompt construction. Domain entities and semantic relationships are identified through ontology mapping and incorporated into domain-specific prompt templates. These templates are subsequently used to guide the T5-based summarization process. This approach illustrates how ontology-derived knowledge can be integrated into prompt generation to provide domain-specific semantic context. However, the extent to which ontology integration independently improves semantic consistency or factual accuracy depends on the experimental comparisons and evaluation metrics reported in the study.

4) *Main contribution*: The principal contribution of this study is the development and evaluation of an ontology-based prompt tuning framework for domain-specific news summarization. The authors report that the proposed framework outperforms several baseline summarization models, including GPT-3.5 and BART, on the evaluated automatic metrics. These findings provide empirical evidence of the performance of the complete ontology-integrated summarization framework within the sports news domain. However, they should not be interpreted as conclusive evidence that ontology integration alone improves semantic understanding or factual consistency. The study also demonstrates a practical mechanism for incorporating ontology-derived semantic knowledge into structured prompt templates.

5) *Limitation*: The framework is primarily evaluated within the sports news domain, limiting conclusions about its generalizability to other domains and languages. Its effectiveness may also depend on the completeness, accuracy, and maintenance of the underlying ontology. Automatic summarization metrics may not fully capture factual consistency, semantic correctness, or human readability. Furthermore, comparisons involving different summarization architectures do not necessarily isolate the independent contribution of ontology-derived prompt information. Future research should therefore consider controlled ablation experiments, broader cross-domain evaluation, and complementary human assessments.

C. *Ontology-Based Prompting with Large Language Models for Inferring Construction Activities from Construction Images [1]*

1) *Objective*: The study [1] investigates an ontology-based prompting framework for inferring construction activities from site images. It integrates a domain-specific construction ontology (ConSE) with LLM prompting to transform extracted visual information into structured semantic representations. Unlike studies that primarily use LLMs to construct ontologies, this work uses an existing ontology to guide activity inference. The framework aims to improve the interpretation of construction scenes while reducing dependence on extensive annotated datasets.

2) *Methodology*: The proposed framework combines ontology-guided prompting with visual feature extraction and CoT reasoning. Construction entities, attributes, and spatial relationships are extracted from images and mapped to the ConSE ontology to produce structured semantic representations. These representations are incorporated into prompt templates alongside task instructions and in-context examples. CoT prompting is used to organize the activity inference process into sequential reasoning steps. The framework is evaluated using GPT-based models through comparative and ablation experiments that investigate the contributions of different framework components.

3) *Ontology-guided prompting strategy*: In this framework, ontology serves as a structured semantic knowledge source that

connects visual information with LLM-based activity inference. Extracted construction entities and spatial relationships are mapped to the ConSE ontology, and the resulting semantic representations are incorporated into prompt templates. This enables the model to use ontology-derived contextual information alongside task instructions and CoT prompting. Unlike ontology construction approaches, the ontology functions primarily as an input for activity inference. The reported framework is designed to reduce ambiguity and support interpretable reasoning, although the contribution of each component should be interpreted according to the reported ablation results.

4) *Main contribution*: The principal contribution of this study is the development and evaluation of an ontology-based prompting framework for construction activity inference from images. The framework integrates ConSE-based semantic representations, spatial information, and CoT prompting to support LLM-based activity inference. Through comparative and ablation experiments, the study examines the performance of the overall framework and the contributions of its components. These experiments provide relevant empirical evidence for ontology-guided prompting in a construction-specific multimodal task, although the strength of the evidence for individual components depends on the experimental controls and reported results.

5) *Limitation*: The framework depends on accurate extraction of visual entities and spatial relationships, and errors at this stage may propagate through ontology mapping to activity inference. Its evaluation on a relatively small dataset limits the generalizability of the reported findings across construction scenarios. Furthermore, the development and maintenance of the ConSE ontology introduce additional knowledge-engineering requirements. The reported variation across LLMs also suggests model-dependent performance. Although the study includes ablation experiments, the extent to which ontology representation, spatial information, and CoT contribute independently should be assessed using the specific experimental configurations and reported results.

D. *Ontology-Driven Generation of Parameters for Health Technology Assessment Models: a Prompt Engineering Study [12]*

1) *Objective*: The study [12] investigates the feasibility of using LLMs to support semi-automated ontology population for Health Technology Assessment (HTA). It integrates prompt engineering with the Ontology for Simulation Modeling (OSDi) to generate ontology individuals representing HTA model parameters. In this approach, the existing ontology provides a predefined semantic structure, while the generated ontology instances constitute the target output. The study aims to reduce manual ontology population effort while maintaining the relevance and consistency of generated information.

2) *Methodology*: The framework integrates prompt engineering with the OSDi ontology to support semi-automated ontology population. ChatGPT-4.1 and Perplexity are

evaluated under different prompting strategies, serialization formats, and information sources to generate ontology individuals for selected HTA use cases. Domain experts assess the generated instances in terms of consistency, relevance, completeness, and adequacy. Statistical analyses are used to examine associations between the evaluated prompting configurations and output quality, including the occurrence of hallucinations. Because multiple design factors are investigated, the effects of prompt design, ontology representation, and information sources should be interpreted according to the specific experimental comparisons.

3) *Ontology-guided prompting strategy*: In this framework, the OSDi ontology serves as a predefined semantic structure for guiding LLM-based ontology population. Its concepts, relationships, and modeling requirements inform prompt construction and provide constraints for generating ontology individuals. The ontology therefore plays two distinct roles: an input structure that guides the generation process and a target schema into which the generated instances are incorporated. This approach illustrates ontology-constrained instance generation rather than ontology construction from scratch. However, the extent to which ontology-based constraints improve output consistency or reduce hallucinations must be interpreted in relation to the reported experimental comparisons and expert assessments.

4) *Main contribution*: The principal contribution of this study is the development and evaluation of a semi-automated ontology population framework that combines prompt engineering with the OSDi ontology for HTA applications. The authors examine how prompting strategies, ontology representations, and serialization formats relate to the quality of generated ontology individuals, including their completeness and the occurrence of hallucinations. The study provides empirical evidence concerning the feasibility and quality of LLM-assisted ontology population under different prompting configurations. However, the reported findings should not be interpreted as demonstrating that ontology integration alone improves efficiency or reliability unless supported by controlled comparisons.

5) *Limitation*: The framework still requires expert validation to assess the correctness and suitability of generated ontology individuals, limiting the extent of automation. Difficulties in handling structurally complex ontology components and representations expressed using the Web Ontology Language (OWL) or Resource Description Framework (RDF) further constrain its applicability. The evaluation covers a limited number of HTA use cases and expert assessments, restricting generalizability to broader healthcare settings. In addition, the study does not establish the full operational costs of ontology preparation, prompt configuration, expert review, and subsequent maintenance. Future research should investigate automated validation, broader use-case coverage, and the trade-off between generation quality and human oversight.

E. Pattern-Based Prompting for Accurate Extraction of Ontology Assertional (A-box) Axioms Using LLMs for Ontology Population [8]

1) *Objective:* The study [8] investigates a pattern-based prompting strategy for extracting assertional (A-box) axioms from unstructured text to support ontology population. It uses Ontology Design Patterns (ODPs) to provide structured guidance for LLM-based extraction of ontology assertions. In this approach, ontology modeling patterns inform prompt construction, while the extracted A-box axioms constitute the target output. The study aims to improve extraction accuracy and reduce hallucinations in ontology population tasks.

2) *Methodology:* The proposed framework adopts a pattern-based prompting approach for ontology population. Following preliminary comparisons with conventional few-shot prompting, the authors construct prompt templates based on ODPs, including hierarchical, binary-relation, and attribute-value patterns. These patterns are translated into structured instructions for extracting A-box axioms from unstructured text. The approach is evaluated in an agricultural case study by comparing generated assertions with expert-annotated reference data. This evaluation provides evidence concerning extraction performance in the tested setting, while the contribution of individual patterns depends on the reported experimental comparisons.

3) *Ontology-guided prompting strategy:* In this study, ontology engineering knowledge is introduced into prompt construction through ODPs. These patterns encode reusable modeling structures and are translated into task-specific instructions for extracting A-box axioms. The ODPs therefore serve as input guidance for prompting, whereas the extracted assertions constitute the intended output for ontology population. This interaction differs from approaches that use an existing domain ontology primarily to support downstream reasoning or summarization.

4) *Main contribution:* The principal contribution of this study is the development and evaluation of a pattern-based prompting approach that incorporates ODPs into LLM-based A-box extraction. The authors report improved extraction performance and reduced hallucinations when using pattern-based prompting in the evaluated agricultural setting. These findings provide empirical evidence for the potential value of ontology modeling patterns as structured prompt guidance. However, the magnitude and generalizability of these improvements depend on the reported evaluation metrics, baseline configurations, and experimental conditions. The study also illustrates how ontology engineering knowledge can be translated into reusable prompting instructions for ontology population.

5) *Limitation:* The framework relies on the manual selection of appropriate ODPs, requiring ontology engineering expertise and limiting the degree of automation. Its pattern-by-pattern processing approach may also introduce scalability challenges when handling complex ontologies. Furthermore, the evaluation is primarily conducted in the agricultural domain,

leaving cross-domain generalizability insufficiently established. The approach also assumes that suitable modeling patterns are available for the target ontology. Incorrect pattern selection or incomplete pattern coverage could introduce extraction errors or omit relevant assertions, although these potential failure modes require further empirical investigation. Future research should examine automated pattern selection, multi-pattern processing, and evaluation across diverse domains.

F. Prompting is Not All You Need: Evaluating GPT-4 Performance on a Real-World Ontology Alignment Use Case [13]

1) *Objective:* The study [13] evaluates the effectiveness of GPT-4-based prompting for ontology alignment in a real-world setting. Using the EuroVoc and TESEO ontologies, it examines whether zero-shot and few-shot prompting can identify semantic correspondences between concepts, including equivalence and hierarchical relationships. The study focuses on the limitations of prompt-based semantic matching when applied to ontologies with substantial size and semantic complexity.

2) *Methodology:* The study evaluates GPT-4 for ontology alignment using different zero-shot and few-shot prompt formulations. Conceptual information from the EuroVoc and TESEO ontologies is provided as task input, and the model is asked to identify semantic correspondences, including equivalence, broader, and narrower relationships. The predicted alignments are compared with a manually constructed reference alignment using precision, recall, and F1-score. This evaluation assesses the performance of prompt-based ontology matching rather than the independent effect of ontology integration on unrelated LLM tasks.

3) *Ontology-guided prompting strategy:* In this study, the two existing ontologies serve as task inputs for prompt-based semantic matching. Their concept labels, definitions, and hierarchical relationships are incorporated into GPT-4 prompts to support the identification of cross-ontology correspondences. The intended outputs are alignment relationships rather than newly constructed ontologies or ontology instances. This represents an ontology alignment interaction in which prompting is used to process existing semantic structures. The reported limitations of GPT-4 in identifying complex hierarchical correspondences also indicate that providing ontology information within prompts does not guarantee accurate semantic reasoning.

4) *Main contribution:* The principal contribution of this study is its empirical evaluation of GPT-4-based prompting for real-world ontology alignment. The reported results indicate that GPT-4 is more effective at identifying concept equivalence than complex hierarchical correspondences, such as broader and narrower relationships, within the evaluated setting. These findings highlight limitations of zero-shot and few-shot prompting for semantically complex ontology alignment tasks. They also motivate further investigation of hybrid approaches that combine LLM-based semantic matching with dedicated

ontology alignment methods, although the superiority of such hybrid approaches is not established by this evaluation alone.

5) *Limitation*: The study's findings are limited to the evaluated GPT-4 configurations and ontology alignment setting. Difficulties in identifying complex hierarchical correspondences indicate limitations in handling semantic relationships beyond concept equivalence. Large ontology inputs may also create context-length and information-selection challenges, while reliance on a closed-source model constrains reproducibility and cross-model comparisons. Future research should examine ontology-aware information selection, hybrid alignment methods, and broader evaluations across different ontologies and language models.

G. Summary

To support cross-study synthesis, the six core studies were analyzed using a common set of analytical criteria, including application domain, research objective, ontology type and design, ontology-LLM interaction, prompting strategy, research method, evaluation evidence, key findings, and limitations. The comparative summary is provided in Appendix A (Table VI). The six core studies were synthesized separately from supporting conceptual references, including Prompt Canvas [15] and MIND-SAFE [9]. These supporting references were used to contextualize structured prompt construction and knowledge grounding, respectively, rather than as direct empirical evidence of ontology-guided PE.

The six core studies can be categorized into four ontology-LLM interaction patterns. First, ontology-guided task execution uses existing semantic knowledge to support downstream applications, including news summarization [11] and construction activity inference [1]. Second, LLM-assisted ontology construction treats ontology development as the primary output objective rather than directly evaluating the effect of ontology-guided prompting [14]. Third, ontology-constrained instance generation employs predefined ontology structures or modeling patterns to guide the production of ontology individuals or A-box axioms [8], [12]. Fourth, prompt-based ontology alignment uses information from existing ontologies to identify semantic correspondences between them [13]. These interaction patterns distinguish ontology as an input knowledge source, a structural constraint, or an output target.

The evidence across these studies is heterogeneous because their research objectives, evaluation methods, and outcome measures differ. Studies focusing on news summarization and construction activity inference assess ontology-informed task performance, although the independent contribution of ontology may be difficult to distinguish from other framework components [1], [11]. Ontology population studies examine the quality of generated ontology instances or assertions [8], [12], while ontology construction research primarily evaluates the generation and organization of domain knowledge [14]. In contrast, ontology alignment research highlights the limitations of prompt-based semantic matching when dealing with complex hierarchical relationships [13]. Therefore, these studies should not be interpreted as providing uniform evidence that ontology integration independently improves LLM performance. Instead, they offer complementary insights into the applications,

evaluation approaches, and limitations of ontology-related prompting, providing a foundation for the capability-oriented discussion in Section V.

V. DISCUSSION

A. Role of Ontology in Prompt Engineering

1) *Knowledge representation*: Knowledge representation provides a foundational analytical perspective for understanding how structured semantic information is incorporated into ontology-related prompting approaches. Across the reviewed studies, its role varies according to the ontology-LLM interaction pattern. In construction activity inference, ontology representations provide structured semantic inputs for prompting [1], while ontology-based news summarization uses domain knowledge to guide content generation [11]. In contrast, ontology construction research focuses on generating structured knowledge representations as outputs [14]. Ontology population studies use predefined semantic structures or modeling patterns to guide instance generation [8], [12]. These examples demonstrate the diverse roles of knowledge representation, but they do not uniformly establish that ontology integration independently improves LLM performance.

In ontology-informed prompting, knowledge representation can provide explicit domain concepts, entities, properties, and relationships for use in task instructions or semantic constraints. In ontology population studies, predefined schemas and modeling patterns may also guide the generation of structured outputs [8], [12]. These applications illustrate how semantic knowledge can be incorporated into LLM-based workflows. However, reductions in ambiguity or hallucination rates, or improvements in output quality, should be attributed to ontology integration only when supported by appropriate controlled comparisons.

Knowledge representation is relevant across the reviewed studies, but its implementation differs according to the ontology-LLM interaction pattern. Some studies incorporate existing semantic structures into task inputs or prompting procedures, whereas others use LLMs to generate or populate ontology-related representations. Accordingly, knowledge representation should not be characterized as a uniformly implemented input layer across all six studies.

Several challenges remain in applying ontology-based knowledge representation to PE. Larger ontology resources may introduce additional information-selection, retrieval, and context-management challenges in ontology-related prompting workflows. For example, selecting relevant concepts and relationships from extensive semantic structures may increase computational overhead or require additional retrieval procedures. Long prompts may also create difficulties in effectively using all supplied information. Although these are plausible limitations, the six reviewed studies do not provide a common evaluation of how ontology size or prompt length affects performance. Their implications for ontology-informed prompting therefore require further empirical investigation. These limitations suggest that future research should explore

more scalable ontology retrieval mechanisms, automated ontology evolution techniques, and systematic evaluation frameworks to better understand the factors influencing the effectiveness of ontology-guided PE across different LLMs and application domains. While knowledge representation establishes the semantic foundation for ontology-guided PE, effective knowledge organization further depends on concept hierarchies, which structure domain knowledge at multiple levels of abstraction.

2) *Concept hierarchy*: Concept hierarchy represents an important ontology capability for organizing domain concepts across different levels of abstraction. Within the reviewed studies, hierarchical structures appear in different forms, including hierarchical activity representations in construction activity inference [1], hierarchical ontology design patterns for A-box extraction [8], and broader–narrower concept relationships in ontology alignment [13]. However, these studies examine hierarchy in different contexts and do not uniformly evaluate its independent contribution to prompting performance. In particular, the difficulties reported in identifying hierarchical correspondences during ontology alignment [13] suggest that the presence of hierarchical information does not necessarily ensure accurate hierarchical reasoning.

In ontology-guided PE, concept hierarchies can provide an explicit structure for organizing domain concepts at different levels of abstraction. When incorporated into prompts, these relationships may help distinguish broader and narrower concepts. However, their independent effects on semantic interpretation, prompt consistency, and response reliability remain insufficiently established.

For example, in ontology-related applications, hierarchical structures are addressed through ontology design patterns for assertion extraction [8] and through the evaluation of hierarchical correspondences in ontology alignment [13]. In construction management, hierarchical structures enable systems to infer high-level group activities from lower-level actions and gestures, which improves contextual interpretation [1]. Similar hierarchical organizations are also applied in HTA to enhance semantic standardization, interoperability, and the reusability of conceptual models [12]. These examples illustrate the relevance of hierarchical knowledge organization across different ontology-related tasks, rather than establishing a consistent improvement in prompting performance.

Despite its advantages, existing studies primarily employ manually designed concept hierarchies that remain domain-specific and difficult to generalize across disciplines. A potential practical limitation is that manually maintained hierarchies may become outdated when domain knowledge changes; the effects of this limitation require further evaluation. These limitations highlight the need for more adaptive ontology engineering approaches capable of automatically constructing and updating hierarchical knowledge structures. Future ontology-guided PE should therefore explore dynamic hierarchy construction, automated concept alignment, and adaptive hierarchical reasoning to improve scalability and domain transferability. While concept hierarchies organize concepts into multiple levels

of abstraction, effective PE also depends on explicitly modeling the semantic relationships that connect these concepts.

3) *Semantic relationships*: Semantic relationships provide explicit connections among concepts, entities, and attributes within ontology-based knowledge structures. Their roles vary across the reviewed studies. In construction activity inference, semantic and spatial relationships contribute to the representation of construction scenes [1]. In ontology population, predefined relations or modeling patterns guide the generation of ontology individuals and assertions [8], [12]. In ontology alignment, the identification of equivalence and hierarchical correspondences constitutes the task itself [13]. These applications illustrate the relevance of semantic relationships to ontology-related prompting, but they do not establish that representing such relationships consistently improves LLM interpretation across different tasks.

In ontology-guided PE, semantic relationships provide explicit contextual links between concepts that guide LLMs in understanding domain knowledge. Incorporating ontology-defined relationships into prompts may provide contextual information about entities and their connections. However, improvements in ambiguity reduction, coherence, and interpretability should be claimed only when supported by comparisons that isolate the contribution of those relationships.

Semantic relationships are represented or examined across several of the reviewed studies, but their functions differ by task. In ontology-informed applications, they may provide structured contextual information for prompting, whereas in ontology construction and population studies, they may constitute part of the generated semantic output. Therefore, their presence in an ontology-related workflow should not be taken as evidence that ontology-defined relationships consistently improve prompt construction or LLM performance.

For example, ontology-based semantic relationships are used in manufacturing to organize fragmented technical documentation into interoperable knowledge structures [14]. Similar approaches are employed in construction management to map low-level actions to higher-level semantic interpretations of project activities [1]. In healthcare technology assessment, semantic relationships define interactions among diseases, interventions, and outcomes, thus improving the transparency and reproducibility of conceptual models [12]. Collectively, these studies illustrate different uses of semantic relationships in ontology-related tasks, while their independent effects on LLM accuracy and consistency remain insufficiently established.

Despite these advantages, several challenges remain in applying ontology-based semantic relationships to PE. Existing studies report inconsistencies in relation naming generated by LLMs, which may reduce semantic interoperability across knowledge bases. Moreover, current models still struggle to capture complex semantic relations, particularly broader–narrower or non-equivalence relationships that require deeper contextual understanding. Larger ontologies may increase information-selection and relationship-matching costs, but the magnitude of this overhead and its effects on scalability require direct measurement. These limitations suggest that future

ontology-guided PE should explore more efficient relation alignment, adaptive semantic reasoning, and scalable ontology retrieval mechanisms.

4) *Logical reasoning*: Logical reasoning is conceptually distinct from knowledge representation, concept hierarchy, and semantic relationships because it concerns the derivation of conclusions from available knowledge. However, the reviewed studies provide limited direct evidence of formal ontology-based inference. In construction activity recognition, ontology-informed representations are combined with CoT prompting to infer higher-level activities [1]. This illustrates structured task reasoning, but the contribution of ontology-based reasoning should be distinguished from the effects of CoT and other framework components. In ontology alignment, GPT-4 encounters difficulties with complex semantic correspondences [13], indicating limitations in prompt-based semantic reasoning. Therefore, logical reasoning should be interpreted as a potential capability of ontology-guided PE rather than a consistently demonstrated outcome across the reviewed studies.

In ontology-related prompting, formal logical reasoning could provide an additional mechanism for deriving conclusions from explicit semantic rules and constraints. However, this possibility should be distinguished from multi-step natural-language reasoning generated through prompting strategies such as CoT. The reviewed studies do not establish that incorporating ontology-defined rules consistently improves inference accuracy or logical consistency compared with equivalent prompts without formal ontology reasoning. Although prompting strategies such as CoT can facilitate reasoning, grounding such strategies in structured ontological knowledge may provide additional semantic constraints and reasoning support [1], [12].

Within the proposed capability-oriented framework, logical reasoning refers to the potential derivation of additional conclusions through formal inference over represented knowledge. It is analytically distinct from knowledge representation, concept hierarchy, and semantic relationships, although these capabilities may provide information required by reasoning mechanisms. This distinction does not imply that all ontology-related prompting approaches implement formal reasoning or that the four capabilities form a universally validated progression.

For example, construction activity inference combines ontology-informed representations with sequential CoT prompting [1]. This constitutes structured task reasoning rather than direct evidence of formal ontology rule execution. In HTA applications, predefined ontology structures guide the generation of model-related individuals [12]. Together, these examples illustrate structured knowledge use in prompting, rather than direct evidence of formal ontology-based inference.

Despite these advantages, several challenges remain. First, current LLMs continue to struggle with complex logical inference, particularly when reasoning involves non-equivalence relationships or sophisticated domain assumptions. Second, lengthy reasoning chains may lead to the “lost-in-the-middle” phenomenon, in which models fail to effectively utilize

information located in the middle of long-context prompts, thereby reducing reasoning consistency. Finally, the stochastic nature of LLMs limits the reproducibility of reasoning outcomes and often requires expert validation to ensure semantic correctness and factual reliability. These limitations indicate that future ontology-guided PE should explore more robust reasoning strategies, adaptive ontology-aware inference mechanisms, and hybrid neuro-symbolic reasoning frameworks to improve reasoning accuracy, scalability, and reproducibility.

5) *Capability-oriented synthesis*: The reviewed studies can be interpreted through four complementary ontology knowledge capabilities: knowledge representation, concept hierarchy, semantic relationships, and logical reasoning. These capabilities provide an analytical framework for examining how ontology-related knowledge is represented, organized, connected, and potentially used for inference. Rather than constituting a universally validated sequence of mechanisms, they represent complementary conceptual dimensions whose empirical support varies across the reviewed studies. As summarized in Table III, each capability has a distinct potential role in supporting prompt construction and reasoning.

TABLE III. CONCEPTUAL FUNCTIONS AND POTENTIAL CONTRIBUTIONS OF FOUR ONTOLOGY KNOWLEDGE CAPABILITIES

Capability	Primary Function	Contribution to Prompt Engineering	Current Limitation	Future Direction
Knowledge Representation	Represent domain knowledge	Provides structured semantic knowledge	Ontology currency and maintenance requirements	Ontology evolution
Concept Hierarchy	Organize concepts	Organizes semantic knowledge across abstraction levels	Hierarchy design and adaptation requirements	Dynamic hierarchy construction
Semantic Relationships	Connect concepts	Connects concepts through contextual dependencies	Potentially incomplete or inconsistent relationships	Relation discovery and alignment
Logical Reasoning	Infer new knowledge	May support formal inference when appropriate reasoning mechanisms are available	Limited evidence of formal inference in reviewed studies	Neuro-symbolic reasoning

Specifically, knowledge representation provides structured domain knowledge that can serve as a semantic foundation for prompting. Concept hierarchy organizes knowledge across different levels of abstraction, while semantic relationships establish explicit connections among concepts. Logical reasoning, when supported by appropriate inference mechanisms, can enable the derivation of additional knowledge from these structures. These capabilities may operate

individually or in combination, depending on the task and ontology-LLM interaction pattern. Their coordinated use represents a potential approach to ontology-guided PE rather than an empirically established requirement across all reviewed studies.

This capability-oriented synthesis provides an integrated analytical perspective for understanding the different roles of ontology in PE. When considered alongside the study-by-capability evidence matrix, it also highlights differences between conceptual capabilities, their implementation in existing studies, and the available empirical evidence. These distinctions provide a basis for identifying research gaps and developing more targeted directions for future ontology-guided PE research.

Table III summarizes the conceptual functions and potential contributions of the four ontology knowledge capabilities. To distinguish these conceptual functions from the evidence available in the reviewed studies, Table IV presents a study-by-capability evidence matrix covering the six core studies. Each capability is classified as explicitly examined (E), inferred from the described methodology (I), or not evidenced (N). These categories describe the extent to which a capability is represented or examined and do not indicate that it has independently been demonstrated to improve LLM performance. For logical reasoning, an I rating indicates conceptual relevance inferred from task-level reasoning, not evidence that a formal ontology reasoner was implemented or evaluated.

TABLE IV. STUDY-BY-CAPABILITY EVIDENCE MATRIX FOR THE SIX CORE STUDIES

Study	Knowledge Representation	Concept Hierarchy	Semantic Relationships	Logical Reasoning
[14] Ontology Construction	E	I	E	N
[11] News Summarization	E	I	I	N
[1] Construction Activities	E	E	E	I
[12] HTA Ontology Population	E	I	E	N
[8] A-box Extraction	E	E	E	N
[13] Ontology Alignment	E	E	E	I

The matrix indicates that knowledge representation is widely represented across the reviewed studies, whereas concept hierarchy and semantic relationships are examined in more task-specific ways. Direct evidence of formal ontology-based logical inference remains limited. In particular, structured task reasoning, such as ontology-informed CoT prompting, should not be equated with the execution or evaluation of formal ontology reasoning rules. These distinctions suggest that the four capabilities are useful as an analytical framework, but their empirical support varies across the selected studies.

B. Ontology-Enhanced Prompt Engineering Frameworks

The Prompt Canvas [15] provides a conceptual reference for organizing prompt components and prompting techniques within a structured framework. Its primary emphasis is on prompt formulation and organization rather than the formal representation of domain knowledge. In this review, the Prompt Canvas is therefore considered a supporting framework for conceptual comparison, rather than direct empirical evidence of ontology-guided PE.

Based on the capability-oriented synthesis, this review proposes a conceptual mapping between four ontology knowledge capabilities and structured prompt construction. Knowledge representation may provide formalized domain concepts and information for prompt components, while concept hierarchy may organize such knowledge across levels of abstraction. Semantic relationships may specify connections among entities, tasks, and domain concepts, whereas logical reasoning may support inference when appropriate reasoning mechanisms are available. This mapping represents an analytical proposition developed in this review and has not been systematically implemented or empirically validated across the six core studies.

From this conceptual perspective, ontology and the Prompt Canvas may serve complementary purposes. The Prompt Canvas provides a structure for organizing prompt components, whereas ontology can provide formally represented domain knowledge and semantic constraints. Their potential integration could support the development of more semantically grounded prompts. However, the reviewed studies do not establish a standardized implementation of this integration, and its effectiveness remains to be evaluated through controlled comparisons.

The proposed integration should therefore be regarded as a direction for future research rather than an empirically validated framework. Its feasibility and effectiveness require evaluation across different application domains, LLMs, ontology designs, and prompting strategies. Such evaluations should consider not only task performance but also ontology construction and maintenance costs, semantic consistency, scalability, and the need for expert validation.

It is also important to distinguish the contribution of ontology from other techniques that may be combined with prompt engineering. Ontology provides formalized domain knowledge and semantic constraints, whereas RAG retrieves external information to support generation. CoT prompting structures intermediate reasoning steps, while fine-tuning modifies model parameters through additional training. These techniques may be complementary, but their contributions should not be treated as interchangeable. In particular, improvements observed in systems combining ontology with retrieval or reasoning strategies cannot be attributed solely to ontology without appropriate ablation studies. Future evaluations should therefore compare ontology-based and non-ontology-based configurations while controlling for retrieval methods, prompting strategies, model selection, and other relevant components.

C. Comparative Analysis of Existing Studies

Building upon the capability-oriented synthesis and the study-by-capability evidence matrix presented above, the six core studies are further compared in terms of ontology–LLM interaction, prompting strategies, research methods, and evaluation evidence.

First, the six core studies differ in their research objectives and ontology–LLM interaction patterns. In news summarization [11] and construction activity inference [1], existing ontology knowledge is used to support downstream tasks. In ontology construction [14], ontology development is the primary output objective. In ontology population [8], [12], predefined semantic structures guide the generation of instances or assertions, while ontology alignment [13] focuses on identifying correspondences between existing ontologies. These differences indicate that the studies investigate distinct applications of ontology and LLMs rather than a single, uniformly evaluated ontology-guided prompting approach.

Second, the studies employ different ontology resources and prompting strategies according to their task requirements. Ontology-informed prompts are used in news summarization [11], while construction activity inference combines ontology-based semantic representations with CoT prompting [1]. Ontology population studies use predefined schemas or ontology design patterns to guide instance and assertion generation [8], [12]. Ontology construction [14] and alignment [13] instead use prompting to produce or identify structured semantic knowledge. Although RAG, CoT, and fine-tuning are relevant techniques in the broader LLM literature, they should not be treated as interchangeable with ontology integration or assumed to be employed across all six studies. The diversity of approaches also limits the extent to which their reported outcomes can be attributed to a common prompting mechanism.

Third, the studies differ in their research designs, evaluation objectives, and outcome measures. News summarization [11] and construction activity inference [1] evaluate performance on specific downstream tasks, whereas ontology population studies [8], [12] focus on the generation of structured semantic instances or assertions. Ontology construction [14] examines the production of ontology-related knowledge, while ontology alignment [13] evaluates semantic correspondences against a reference alignment using precision, recall, and F1-score. Consequently, the reported results address different evaluation objectives and cannot be directly aggregated into a common measure of ontology-guided PE effectiveness. Moreover, where ontology is combined with additional prompting or system components, the available comparisons may not isolate its independent contribution.

The evaluation measures reported across the six studies assess different outcomes and therefore should not be treated as directly comparable indicators of ontology effectiveness. Information extraction measures assess the correctness of extracted concepts, relations, or assertions, while summarization metrics primarily assess similarity to reference summaries and may not fully capture factual consistency. Expert assessments of ontology population examine domain-specific qualities such as relevance, completeness, and adequacy, whereas precision, recall, and F1-score in ontology alignment evaluate

correspondence matching against reference alignments. These differences require closer consideration of what each evaluation measure can establish about the role of ontology in the respective task.

Despite their methodological differences, the studies reveal several recurring limitations. Their findings are closely tied to particular application domains, ontology resources, model configurations, and evaluation settings, which limits cross-study comparability and broader generalization. Ontology construction, adaptation, and validation may also require domain expertise and ongoing maintenance, although these costs are not systematically evaluated across the six studies. Furthermore, the reviewed research provides limited controlled evidence for separating the effects of ontology-based knowledge from those of prompting strategies and other system components. These limitations highlight the need for more transparent evaluation designs, broader cross-domain comparisons, and systematic investigation of ontology-related implementation costs.

These differences in ontology–LLM interaction, prompting strategies, evaluation evidence, and implementation requirements inform the research gaps and future directions discussed in the following sections.

D. Research Gaps

Building upon the capability-oriented synthesis, study-by-capability evidence matrix, and comparative analysis of the six core studies, this review identifies five interrelated but analytically distinct research gaps: theoretical, methodological, empirical, knowledge-related, and practical. These gaps concern the conceptual interpretation of ontology-guided PE, the design and evaluation of prompting approaches, the generalizability of reported findings, the mechanisms underlying ontology-related contributions, and the practical requirements of ontology development and deployment. The gaps are derived from the scope and limitations of the selected literature and should not be interpreted as an exhaustive assessment of the entire research field.

1) *Theoretical gap*: The six core studies adopt different assumptions about the role of ontology in LLM-based tasks, including semantic grounding, ontology construction, ontology population, and alignment [1], [8], [11]–[14]. However, these approaches are not organized around a shared conceptual account of how ontology–LLM interaction patterns relate to prompt construction and task objectives. The four ontology knowledge capabilities synthesized in this review provide one analytical perspective, but their relationships should not be interpreted as a universally validated theoretical model. Further research is needed to develop and critically examine conceptual frameworks that distinguish ontology as an input knowledge source, a structural constraint, or an output target.

2) *Methodological gap*: The six core studies employ heterogeneous ontology representations, prompting procedures, baselines, and evaluation measures, making methodological comparison difficult [1], [8], [11]–[14]. In particular, the studies do not establish a common evaluation protocol for distinguishing the contribution of ontology-based knowledge

from that of other prompting or system components. Although structured approaches such as the Prompt Canvas [15] offer a conceptual reference for organizing prompt elements, their integration with ontology capabilities remains an unvalidated proposition in this review. Future research should develop transparent reporting practices, comparable baselines, and controlled evaluation procedures for ontology-related prompting approaches.

3) *Empirical gap*: The six core studies evaluate different ontology-related tasks, including news summarization [11], construction activity inference [1], ontology construction [14], ontology population [8], [12], and ontology alignment [13]. Because their evaluation objectives, datasets, model configurations, and outcome measures differ, the available findings provide limited evidence for the transferability of particular ontology-guided prompting strategies across tasks and domains. Further research should examine whether reported benefits can be reproduced across multiple LLMs, ontology resources, application domains, and evaluation settings using comparable experimental protocols.

4) *Knowledge gap*: Although the reviewed studies involve different ontology knowledge capabilities, the independent contribution of these capabilities to LLM-based task performance remains insufficiently examined. As indicated by the study-by-capability evidence matrix, knowledge representation, concept hierarchy, semantic relationships, and logical reasoning are not equally or directly evaluated across the six studies. Furthermore, ontology-informed prompting is sometimes combined with other techniques, such as CoT prompting [1], making causal attribution difficult. Future research should investigate the effects of individual ontology capabilities through controlled comparisons and capability-specific ablation studies, while distinguishing formal ontology reasoning from LLM-generated reasoning.

5) *Practical-knowledge gap*: The practical requirements of developing, validating, updating, and deploying ontology-related prompting approaches remain insufficiently evaluated in the selected studies. Depending on the application, ontology construction and maintenance may require domain expertise, expert review, and additional computational resources. Moreover, outdated, incomplete, or inconsistent ontology knowledge may introduce semantic errors or reduce the reliability of generated outputs. Future research should systematically assess ontology maintenance effort, expert involvement, computational overhead, and robustness to ontology quality issues alongside task-level performance.

6) *Summary*: In summary, the six core studies reveal five interrelated research gaps concerning conceptual foundations, methodological comparability, empirical generalizability, capability-specific evidence, and practical implementation requirements. As summarized in Table V, each gap reflects a distinct limitation identified through the narrative synthesis and suggests a corresponding direction for future investigation.

Overall, the six core studies illustrate diverse applications of ontology-related prompting, while also revealing important

differences in research objectives, evaluation evidence, and practical requirements. The five identified gaps provide a structured basis for considering future research priorities without assuming that the effectiveness of ontology integration has been uniformly established. The following section outlines research directions aimed at improving conceptual clarity, methodological rigor, empirical validation, capability-specific understanding, and practical applicability.

TABLE V. SUMMARY OF THE IDENTIFIED RESEARCH GAPS AND FUTURE RESEARCH DIRECTIONS

Gap Type	Description	Evidence from Literature	Future Direction
Theoretical Gap	Limited shared conceptual account of ontology–LLM interaction	Different ontology roles across construction, population, task execution, and alignment [1], [8], [11]–[14]	Develop and examine conceptual models distinguishing ontology inputs, constraints, and outputs
Methodological Gap	Limited comparability of prompting and evaluation procedures	Heterogeneous prompts, baselines, and evaluation measures across the six studies	Establish transparent reporting, controlled baselines, and comparable evaluation protocols
Empirical Gap	Limited evidence of cross-task and cross-domain generalizability	Different task objectives and evaluation settings across the six studies	Conduct multi-model, multi-domain, and reproducibility studies
Knowledge Gap	Insufficient evidence isolating individual ontology capabilities	Uneven capability coverage in the study-by-capability matrix; ontology and CoT combined in [1]	Conduct capability-specific ablation studies and isolate ontology contributions
Practical-Knowledge Gap	Limited assessment of ontology maintenance, costs, and robustness	The selected studies do not provide a comparable, systematic assessment of deployment requirements	Evaluate expert effort, maintenance cost, scalability, and ontology quality failures

E. Research Opportunities and Future Research Agenda

The comparative analysis and five research gaps identified in this review suggest several directions for future ontology-guided prompt engineering research. Rather than assuming that ontology integration consistently improves LLM performance, future investigations should examine the conditions under which different ontology–LLM interaction patterns and knowledge capabilities contribute to specific task outcomes. The following research agenda links conceptual development, methodological rigor, empirical generalizability, capability-specific evaluation, and practical implementation to the five gaps identified in the preceding section.

First, to address the theoretical gap, future research should develop and critically examine conceptual models that distinguish the different roles of ontology in LLM-based prompting, including ontology as an input knowledge source, a semantic constraint, or an output target. Structured prompt organization approaches such as the Prompt Canvas [15] may serve as conceptual references for investigating how ontology knowledge capabilities relate to prompt components. However, the mapping proposed in this review remains unvalidated and should be treated as a research proposition rather than an established framework. Future studies could examine whether the proposed conceptual distinctions adequately describe different ontology–LLM interaction patterns across tasks and whether they support more consistent framework design.

Second, to address the methodological gap, future studies should establish transparent and reproducible evaluation protocols for ontology-related prompting approaches. These protocols should document ontology resources, prompt templates, model configurations, retrieval procedures, baselines, and evaluation measures. Controlled comparisons could include a non-ontology prompt baseline, an ontology-informed prompt using the same model and task, and additional configurations combining ontology with retrieval or reasoning strategies where applicable. Ablation studies should vary one component at a time to distinguish the effects of ontology from those of CoT prompting, RAG, and other system components. Such designs would improve the interpretability and comparability of reported results.

Third, to address the empirical gap, future research should evaluate ontology-related prompting approaches across multiple LLMs, application domains, datasets, and ontology resources. Such evaluations should examine whether the effects observed under controlled conditions can be reproduced across different tasks and model configurations. Where feasible, studies should include both within-domain and cross-domain testing, report variability across repeated runs, and assess performance using task-appropriate measures. These comparisons would help determine the conditions under which ontology-informed prompting provides transferable benefits, produces no measurable improvement, or introduces additional limitations.

Fourth, to address the knowledge gap, future research should investigate the individual and combined contributions of the four ontology knowledge capabilities identified in this review: knowledge representation, concept hierarchy, semantic relationships, and logical reasoning. Capability-specific ablation studies could compare prompts containing different levels of ontology information while controlling for model configuration, task requirements, and other prompting strategies. In particular, studies should distinguish the use of structured semantic knowledge from the execution of formal ontology reasoning procedures. Such evaluations could clarify which capabilities contribute to particular task outcomes, whether their effects interact, and under what conditions additional ontology structure provides measurable value.

One testable hypothesis is that adding task-relevant ontology relationships to concept-only prompts improves relation-dependent task accuracy and reduces semantic errors. This hypothesis could be evaluated using matched prompts with and

without those relationships while holding the LLM version, source information, prompting strategy, and generation settings constant. Prompt length should be matched where feasible, and performance should be assessed against expert-validated reference data across repeated runs.

Finally, to address the practical-knowledge gap, future studies should examine the costs and operational requirements of ontology construction, validation, maintenance, and deployment. Automated and semi-automated ontology development may reduce some forms of manual effort, but their benefits should be evaluated against semantic quality, expert correction requirements, computational overhead, and long-term maintenance costs. Robustness evaluations should also investigate how incomplete, outdated, or inconsistent ontology knowledge affects generated outputs. A second hypothesis is that incomplete or incorrect ontology relationships reduce output quality relative to a validated ontology. Controlled perturbations could vary the proportion of missing or incorrect relationships while keeping other conditions fixed, with evaluation based on task accuracy and semantic error rates. Such studies could provide a more balanced assessment of the practical value and limitations of ontology-related prompting approaches.

Together, these five research directions translate the identified gaps into a set of conceptual, methodological, empirical, capability-specific, and practical research priorities. By emphasizing controlled comparisons, explicit baselines, reproducibility, and implementation costs, the proposed agenda offers testable directions for future investigation without assuming that the benefits of ontology integration have already been established across tasks and domains.

VI. CONCLUSION AND FUTURE WORK

This narrative literature review examined six core studies published between 2024 and 2026 that investigate different interactions between ontologies and large language models in prompting-related tasks. The synthesis distinguishes approaches in which ontology serves as an input knowledge source or semantic constraint, an output target, or both an input structure and an output representation. These interaction patterns are reflected in applications including news summarization, construction activity inference, ontology construction, ontology population, and ontology alignment. Rather than demonstrating a uniform improvement in prompt effectiveness, the reviewed studies illustrate diverse roles of ontology in LLM-based tasks, with findings that depend on their respective objectives, methods, and evaluation settings.

A central contribution of this review is the capability-oriented synthesis of four ontology knowledge capabilities: knowledge representation, concept hierarchy, semantic relationships, and logical reasoning. These capabilities provide an analytical framework for examining how ontology-related knowledge is represented, organized, connected, and potentially used for inference. The study-by-capability evidence matrix further distinguishes capabilities explicitly examined in the reviewed studies from those inferred from the described approaches or not evidenced. This synthesis highlights variation in the available evidence and should be interpreted as a conceptual analytical perspective rather than an empirically

validated model of how ontology improves prompting performance.

The review identifies five interrelated research gaps: theoretical, methodological, empirical, knowledge-related, and practical. These gaps concern the conceptual interpretation of ontology–LLM interactions, the comparability of prompting and evaluation procedures, the generalizability of reported findings, the independent contributions of ontology knowledge capabilities, and the practical requirements of ontology development and maintenance. In particular, the heterogeneous evaluation settings and limited capability-specific comparisons make it difficult to isolate the effects of ontology from those of other prompting or system components. The identified gaps reflect limitations in the selected literature rather than an exhaustive assessment of the entire research field.

Future research should investigate ontology-related prompting through controlled comparisons, capability-specific ablation studies, cross-model and cross-domain evaluations, and assessments of ontology maintenance costs and semantic robustness. Structured prompt organization approaches such as the Prompt Canvas [15] may also provide conceptual references for exploring the integration of prompt components with ontology-based knowledge. However, the proposed mapping between Prompt Canvas elements and ontology capabilities has not been empirically validated and should be evaluated independently before claims about its effectiveness, reusability, or generalizability can be established.

Overall, this review contributes a narrative synthesis of six heterogeneous ontology–LLM studies, a capability-oriented analytical framework, and a research agenda grounded in five identified gaps. By distinguishing ontology interaction patterns, levels of capability-specific evidence, and unvalidated conceptual propositions, the review provides a structured basis for future investigations into the role of ontology in prompt engineering. Further empirical research is required to determine when, how, and to what extent ontology-related approaches contribute to LLM-based task performance.

FUNDING

This work was funded by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia, [Grant No: KFU266114].

ACKNOWLEDGMENT

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia.

REFERENCES

- [1] C. Zeng, T. Hartmann, and L. Ma, "Ontology-based prompting with large language models for inferring construction activities from construction images," *Advanced Engineering Informatics*, vol. 69, Art. no. 103869, 2026, doi: 10.1016/j.aei.2025.103869.
- [2] F. Algobaei and E. Alzain, "Prompt engineering for non-native English learners: A generative AI approach to personalised language feedback," *Social Sciences & Humanities Open*, vol. 13, Art. no. 102341, 2026, doi: 10.1016/j.ssho.2025.102341.
- [3] J. Park and S. Choo, "Generative AI prompt engineering for educators: Practical strategies," *Journal of Special Education Technology*, vol. 40, no. 3, pp. 411–417, 2025, doi: 10.1177/01626434241298954.
- [4] J. White et al., "A prompt pattern catalog to enhance prompt engineering with ChatGPT," *arXiv preprint arXiv:2302.11382*, 2023, doi: 10.48550/arXiv.2302.11382.
- [5] L. S. Lo, "The CLEAR path: A framework for enhancing information literacy through prompt engineering," *The Journal of Academic Librarianship*, vol. 49, no. 4, Art. no. 102720, 2023, doi: 10.1016/j.jacalib.2023.102720.
- [6] A. Gao, "Prompt engineering for large language models," SSRN working paper, 2023. [Online]. Available: <https://ssrn.com/abstract=4504303>.
- [7] L. Henrickson and A. Meroño-Peñuela, "Prompting meaning: A hermeneutic approach to optimising prompt engineering with ChatGPT," *AI & Society*, vol. 40, pp. 903–918, 2025, doi: 10.1007/s00146-023-01752-8.
- [8] R. Shyama, I. Wilson, A. Ginige, and J. Goonetillake, "Pattern-based prompting for accurate extraction of ontology assertional (A-box) axioms using LLMs for ontology population," *Procedia Computer Science*, vol. 270, pp. 1438–1447, 2025, doi: 10.1016/j.procs.2025.09.265.
- [9] S. Boit and R. Patil, "A prompt engineering framework for large language model-based mental health chatbots: Conceptual framework," *JMIR Mental Health*, vol. 12, Art. no. e75078, 2025, doi: 10.2196/75078.
- [10] V. Liu and L. B. Chilton, "Design guidelines for prompt engineering text-to-image generative models," in *Proc. CHI Conf. Human Factors in Computing Systems (CHI '22)*, 2022, pp. 1–23, doi: 10.1145/3491102.3501825.
- [11] A. R. S. Silva and Y. H. P. P. Priyadarshana, "Ontology-based prompt tuning for news article summarization," *Frontiers in Artificial Intelligence*, vol. 8, Art. no. 1520144, 2025, doi: 10.3389/frai.2025.1520144.
- [12] E. González-González, I. Castilla-Rodríguez, and J. A. Dorta-Hernández, "Ontology-driven generation of parameters for health technology assessment models: A prompt engineering study," *International Journal of Technology Assessment in Health Care*, vol. 42, no. 1, Art. no. e47, 2026, doi: 10.1017/S0266462326103754.
- [13] G. Macilenti, A. Stellato, and M. Fiorelli, "Prompting is not all you need: Evaluating GPT-4 performance on a real-world ontology alignment use case," *Procedia Computer Science*, vol. 246, pp. 1289–1298, 2024, doi: 10.1016/j.procs.2024.09.557.
- [14] X. Liu et al., "From prompt to graph: Comparing LLM-based information extraction strategies in domain-specific ontology development," *Procedia Computer Science*, vol. 277, pp. 50–60, 2026, doi: 10.1016/j.procs.2026.02.046.
- [15] M. Hewing and V. Leinhos, "The Prompt Canvas: A literature-based practitioner guide for creating effective prompts in large language models," *arXiv preprint arXiv:2412.05127*, 2024. [Online]. Available: <https://arxiv.org/abs/2412.05127>.
- [16] A. Basheer, "The art and science of writing narrative reviews," *International Journal of Advanced Medical and Health Research*, vol. 9, no. 2, pp. 124–126, 2022, doi: 10.4103/ijamr.ijamr_234_22.
- [17] H. Snyder, "Literature review as a research methodology: An overview and guidelines," *Journal of Business Research*, vol. 104, pp. 333–339, 2019, doi: 10.1016/j.jbusres.2019.07.039.
- [18] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a literature review," *MIS Quarterly*, vol. 26, no. 2, pp. xiii–xxiii, 2002, doi: 10.2307/4132319.
- [19] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and the PRISMA Group, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Medicine*, vol. 6, no. 7, Art. no. e1000097, 2009, doi: 10.1371/journal.pmed.1000097.
- [20] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, Art. no. n71, 2021, doi: 10.1136/bmj.n71.
- [21] R. J. Torracco, "Writing integrative literature reviews: Using the past and present to explore the future," *Human Resource Development Review*, vol. 15, no. 4, pp. 404–428, 2016, doi: 10.1177/1534484316671606.

[22] A. H. M. Nasir, E. M. S. Salwana, and M. N. Ahmad, "Integrating information governance as the foundation for information dashboard

design ontology," Journal of Information Technology Education: Innovations in Practice, vol. 24, Art. no. 001, 2025, doi: 10.28945/5432.

TABLE VI. APPENDIX A. COMPARATIVE SUMMARY OF SIX CORE ONTOLOGY-LLM STUDIES

Ref.	Application Domain	Research Objective	Ontology Name	Ontology-LLM Interaction Pattern	Prompting Strategy	Research Method	Evaluation / Evidence	Key Finding	Limitation
[14]	Manufacturing	Semi-automate domain-specific ontology construction	Casting domain ontology	Target knowledge representation for ontology construction; predefined categories support extraction	Pre-trained prompting, in-context learning, and fine-tuning	Comparative experiment	Comparative evaluation of information extraction approaches	Fine-tuning performed best for concept and relationship extraction in the evaluated casting-domain setting	Does not isolate the effect of ontology-guided prompting on LLM reasoning performance; limited dataset and domain
[11]	News Summarization	Improve domain-specific news summarization using ontology-derived prompt information	Sports news domain ontology	Input knowledge source for semantic mapping and prompt generation	Ontology-based prompt tuning with T5-based summarization	Experimental study	Comparison with baseline summarization models using automatic metrics	The complete framework reportedly outperformed selected baselines on evaluated summarization metrics	Sports-news-specific evaluation; ontology quality and maintenance; limited causal attribution and metric validity
[1]	Construction	Infer construction activities from images using ontology-guided LLM prompting	ConSE construction ontology	Input semantic knowledge source; ConSE-based representation of construction entities and spatial relationships	Ontology-derived prompt representations combined with CoT and in-context examples	Comparative / ablation experiments	Comparative and ablation experiments using GPT-based models	The study evaluates ontology-integrated prompting for construction activity inference and examines the contributions of framework components	Dependence on visual feature extraction; small dataset; ontology development and maintenance; model-dependent performance
[12]	Healthcare	Support semi-automated ontology population for HTA model parameters	OSDi formal ontology	Existing OSDi schema as input constraint; generated ontology individuals as output	Ontology-constrained prompting with different prompt configurations and serialization formats	Experimental study	Expert assessment of consistency, relevance, completeness, and adequacy; statistical analysis of output quality and hallucinations	Prompting configurations and ontology representations were evaluated for their effects on generated instance quality	Expert validation requirements; complex ontology formats; limited HTA cases; operational costs not fully established
[8]	Agriculture	Extract A-box axioms from unstructured text for ontology population	Ontology Design Patterns (ODPs)	ODPs as input modeling guidance; extracted A-box axioms as output	Pattern-based prompting using hierarchical, binary-relation, and attribute-value patterns	Experimental study	Agricultural case study with expert-annotated reference data and comparisons involving few-shot prompting	Pattern-based prompting reportedly improved assertion extraction and reduced hallucinations in the evaluated setting	Manual ODP selection; pattern-by-pattern processing; agricultural-domain evaluation; dependence on suitable patterns

[13]	Ontology Alignment	Evaluate GPT-4-based prompting for real-world ontology alignment	EuroVoc–TESEO ontologies	Two existing ontologies as task inputs; semantic correspondences as output	Zero-shot and few-shot prompting using ontology concept information	Evaluation study	Reference alignment comparison using precision, recall, and F1-score	GPT-4 reportedly performed better on equivalence matching than on complex hierarchical correspondences	Semantic complexity; large ontology inputs; GPT-4 dependence; limited cross-model generalizability
------	--------------------	--	--------------------------	--	---	------------------	--	--	--