From Symbolic to Natural-Language Relations: Rethinking Knowledge Graph Construction in the Era of Large Language Models

Anonymous ACL submission

Abstract

Knowledge graphs (KGs) have commonly been constructed using predefined symbolic relation schemas, typically implemented as categorical relation labels. This design has notable shortcomings: real-world relations are often contextual, nuanced, and sometimes uncertain, and compressing it into discrete relation labels abstracts away critical semantic detail. Nevertheless, symbolic-relation KGs remain widely used because they have been operationally effective and broadly compatible with pre-LLM downstream models and algorithms, in which KG knowledge could be retrieved or encoded into quantified features and embeddings at scale. The emergence of large language models (LLMs) has reshaped how knowledge is created and consumed. LLMs support scalable synthesis of domain facts directly in concise natural language, and prompting-based inference favors context-rich free-form text over quantified representations. This position paper argues that these changes call for rethinking the representation of relations themselves rather than merely using LLMs to populate conventional schemas more efficiently. We therefore advocate moving from symbolic to natural-language relation descriptions, and we propose hybrid design principles that preserve a minimal structural backbone while enabling more flexible and context-sensitive relational representations.

1 Introduction

Knowledge graphs (KGs) are widely used in domain- and task-specific applications to compensate for the lack of specialized knowledge in pre-trained models and other algorithmic systems. When training data or model/algorithm capacity alone is insufficient to capture domain-specific facts and relationships, KGs offer an explicit, structured representation of external knowledge that supports inference and improves performance on knowledge-intensive tasks (Fensel et al., 2020; Hogan et al., 2021; Qu, 2022). A KG is typically represented as a collection of relational triples of the form $[Entity_1, Relation, Entity_2]$, optionally augmented with attributes associated with entities and relations. Entities correspond to real-world concepts (e.g., abstract notions) or instances (e.g., persons, organizations, and events), while relations represent connections between pairs of entities. This structured representation enables KGs to serve as an explicit, machine-recognizable abstraction of domain knowledge that can be integrated with downstream models and algorithms.

Central to the KG paradigm is the design and expressiveness of relations, as they determine what types of knowledge can be represented. The most widely adopted approach to KG construction relies on defining a set of symbolic relation categories with a well-defined schema that specifies how entities should be connected (Hogan et al., 2021; Radevski et al., 2023), such as synonymy, subclass, friend, affiliation, and many others (Fellbaum, 2010; Vrandečić and Krötzsch, 2014; Cao et al., 2020; Steiner et al., 2012).

Although symbolic relation categories capture only simplified abstractions of real-world relationships, they have historically provided a practical and effective paradigm for knowledge extraction and representation (Bordes et al., 2013; Trouillon et al., 2016; Xia et al., 2019). In this framework, relation extraction (Pawar et al., 2017) in KG construction is typically formulated as a classification problem: given a pair of entities co-occurring in text(s) or graph(s), models are used to determine whether a relation exists between them and, if so, to assign it to one of a set of relation types. This formulation aligns well with earlier models that take text/graph features or embeddings as input and symbolic relation labels as output, enabling efficient and accurate construction of KGs from large-scale text corpora or extend existing KGs. The resulting KGs with symbolic relation categories can be

readily used as factual information or encoded into quantified features, embeddings, or graph structures, making them compatible with a wide range of downstream tasks such as question answering, recommendation, clustering, and indexing (Sun et al., 2018; Xiao et al., 2019; Dai et al., 2020; Yani and Krisnadhi, 2021; Ren et al., 2023; Opsahl, 2024).

However, this paradigm also introduces inherent limitations. Since relation types are usually over-simplified, predefined, and discrete (Yates et al., 2007; Niklaus et al., 2018; Pai et al., 2024), the resulting representations tend to abstract away much of the contextual and situational information present in real-world relations and thus restrict the ability to capture nuanced, context-dependent relations (Emirbayer and Goodwin, 1994; Boccaletti et al., 2014). Consequently, many aspects of real-world relationships are either simplified or omitted altogether, revealing a mismatch between the inflexible structure of symbolic relation schemas and labels and the complexity of naturally occurring relational knowledge.

Recent advances in large language models (LLMs) have fundamentally changed the landscape of both knowledge resource construction and their downstream applications. For resource creation (Ding et al., 2024; Tan et al., 2024; Yang et al., 2025), LLMs enable scalable extraction and synthesis of information from massive data, making it feasible to represent knowledge using concise natural-language descriptions rather than numeric or categorical labels. For downstream usage (Baek et al., 2023; Dong et al., 2024; Shu et al., 2024; Wang et al., 2024b), prompting-based inference naturally favors nuanced, context-rich natural language, as such models are designed to interpret and reason over free-form text rather than discrete categorical relation types. As a result, relations between pairs of entities in KG expressed in natural language can be more directly consumed and exploited by LLM-based downstream applications than those represented as traditional symbolic relations.

While recent studies in KG, ontology, or taxonomy construction have increasingly leveraged LLMs, most existing efforts are still within the traditional symbolic paradigm — using LLMs primarily as tools to populate or refine schemas or help symbolic relation prediction (Wei et al., 2023; Fathallah et al., 2024; Liang et al., 2024; Zhu et al., 2024). In contrast, this position paper argues for a more fundamental paradigm shift: Rather than merely employing LLMs to construct conventional KGs with symbolic relations more efficiently, we advocate rethinking the representation of relations themselves. Specifically, we propose moving from symbolic relation categories toward natural-language-based relational representations that better align with the capacity of LLMs. Such a shift enables richer, more flexible, and more context-sensitive representations of knowledge, and is well aligned with the way LLMs are increasingly used in practice.

In this position paper, we first provide a brief review of earlier KG practices and explain why symbolic relations became dominant due to their scalability and compatibility with prior downstream models. We then contrast this tradition with the capabilities and preferences of LLM-centered workflows, highlighting the gap that emerges when complex relational knowledge must be compressed into discrete symbolic labels. Building on this analysis, we articulate conceptual and operational arguments for a paradigm shift from symbolic to natural-language relation representation, and propose a set of principles for hybrid KG designs that preserve structural usability while enabling contextual richness. Finally, we outline concrete research directions for KG construction, refinement, traversal, embedding, and evaluation under this new paradigm.

2 KGs in Prior Literature

2.1 Forms of Relation Representation

From the perspective of relation representation, existing KGs can be broadly categorized into two forms: symbolic relation and OpenIE relation.

Symbolic Relation KGs: Symblic relations come from a predefined schema consisting of discrete relation categories. Each edge in the graph corresponds to a symbolic relation label with well-defined semantics (Fensel et al., 2020; Unni et al., 2022; Zhou et al., 2022). Common examples include linguistic relations (e.g., synonymy-of) (Fellbaum, 2010), taxonomic relations (e.g., child-of/subclass-of) (Vrandečić and Krötzsch, 2014; Han et al., 2020), social or interpersonal relations (e.g., friend-of) (Cao et al., 2020), affiliation or locational relations (e.g., works-at, lives-in) (Steiner et al., 2012), event-centered relations (e.g., participated-in, occurred-at) (Guan et al., 2022), and domain-specific semantic relations (e.g. drug-

treats-disease) (Qu, 2022).

Open Information Extraction (OpenIE) KGs: OpenIE approaches have also been slightly explored for the construction of KGs in prior studies (Martinez-Rodriguez et al., 2018), which represent relations using surface linguistic expressions (e.g., "grew up in", "worked for") , typically verbs or verb phrases extracted directly from text (Radevski et al., 2023). However, OpenIE KGs are less used in downstream tasks as it contains lots of noise and non-canonical relations that are not compatible with domain- or task-specific applications (Romero and Razniewski, 2023). Canonicalizing the open-ended surface linguistic relations into a fixed, normalized set of symbolic labels is often required for domain- or task-specific applications, but has been hardly addressed so far (Niklaus et al., 2018).

2.2 Downstream Usages

The usages of KGs in downstream tasks can be broadly categorized into four groups: (1) factual information retrieval, (2) representation learning where KGs serve as structural supervision for embedding-based models, (3) organization, indexing, and navigation of knowledge artifacts, and (4) new knowledge discovery through link prediction and KG completion.

Factual Information Retrieval: In this usage, KGs function as structured stores of factual information, from which entities, relations, attributes, or subgraphs are retrieved as needed through entity matching or graph traversal. Retrieved information may be presented as answers in fact-based question answering (Bao et al., 2016; Yani and Krisnadhi, 2021), used as auxiliary evidence to augment downstream models (Opsahl, 2024; Zhu et al., 2025c), or converted as additional data for model fine-tuning or training (Tian et al., 2024).

Representation Learning: KGs are used primarily as structural supervision to shape learned representations, most commonly entity embeddings (Dai et al., 2020; He et al., 2020; Ge et al., 2024). Relation structure influences how representations are learned during training, but the explicit graph is typically not consulted at inference time. Downstream tasks such as recommendation (Grad-Gyenge et al., 2017; Sun et al., 2018; Shokrzadeh et al., 2024), ranking (Xiong et al., 2017; Ren et al., 2023), and clustering (Xiao et al., 2019; Gad-Elrab et al., 2020; Li et al., 2024c) are often augmented by KG embeddings.

Organization, Indexing, and Navigation of Knowledge Artifacts: In this usage, the KG is employed as structural backbones for organizing and indexing information resources, such as documents, texts, multimedia, or archival records. It supports efficient indexing, faceted search, browsing, and navigation by providing controlled vocabularies and explicit conceptual relationships, enabling faster search/retrieval and more coherent organization of large knowledge collections (Wenige et al., 2020; Limani et al., 2021).

New Knowledge Discovery via New Link Prediction: In this paradigm, the KG serves as a basis for discovering new or missing links between unlinked entities. Models estimate the plausibility of unobserved links between entities, producing candidate relations or triples that extend the existing graph (Zamini et al., 2022; Cai et al., 2023; Yao et al., 2025a). This usage treats the KG not only as a record of known facts, but as a basis for hypothesizing additional knowledge under the assumptions imposed by the graph's structure and schema.

2.3 Limitations of Existing KGs

While pre-defined relation schemas and the graph structure enable scalability and broad computational compatibility with traditional models and systems as well as LLMs, prior literature in computer science, biomedical science, and social science have long recognized the limitations of the mainstream KGs with symbolic relations from different perspectives.

Rigidity of Predefined Symbolic Relation Schemas: Prior computer science research has discussed the constraints from static and predefined relation, which determine in advance what kinds of facts can be represented, retrieved, organized, or inferred (Yates et al., 2007; Niklaus et al., 2018; Pai et al., 2024). As a result, downstream usages are all bounded by the scope of the relation schema. Relational dimensions that fall outside the predefined categories are systematically omitted, and extending coverage typically requires costly schema redesign and data reprocessing. This rigidity also limits generalization: relations not explicitly anticipated by the schema cannot be easily represented, even when they are salient or frequent in real-world data.

Lack of Canonicalization in OpenIE KGs: OpenIE KGs are the most discussed alternatives to relax the rigidity of symbolic relation schemas by expressing relations as open-ended verb or verb-phrase predicates extracted from text, enabling

broader coverage without predefined ontologies (Martinez-Rodriguez et al., 2018; Pai et al., 2024). However, because these relations remain grounded in surface lexical forms rather than normalized, domain-specific semantics, they introduce severe heterogeneity and sparsity, sacrificing scalability, domain-specificity, and computational compatibility. As a result, they are rarely used in downstream tasks (Niklaus et al., 2018; Romero and Razniewski, 2023).

Ambiguity and Coarse Granularity of Relation Labels: Another fundamental limitation of many existing KGs lies in their insufficient relational granularity for capturing domain-specific distinctions. This issue has been extensively discussed in the biomedical ontology literature, where even seemingly basic relations such as *part-of* exhibit substantial semantic ambiguity. In biomedical contexts, part-of may refer to anatomical composition, participation in a biological process, temporal inclusion, or functional dependency, each of which carries different implications. Collapsing these distinct meanings into a single relation type obscures critical distinctions required for accurate reasoning and integration (Smith et al., 2005; Schulz et al., 2006). Similar limitations arise in other domains. For example, relations such as partner and spouse encode different social and legal semantics, yet are often conflated or mapped to a single coarse-grained relation in schema-based KG construction pipelines (Biswas et al., 2023).

Context Dependence Beyond Symbolic Relation Granularity: A natural response to the limitations of coarse relation schemas is to develop more fine-grained symbolic relation types. However, insights from social science suggest that increasing relational granularity alone does not resolve the fundamental challenge of capturing the meaning and context-dependence of real-world relationships. Many real-world relationships are inherently complex, contextual, nuanced, and uncertain, making them difficult to faithfully capture using a single or a few symbolic labels. Conceptual relations, for example, may be disputed, evolving, or dependent on perspective, often requiring a qualitatively descriptive explanation rather than categorical assignment (Emirbayer and Goodwin, 1994; Boccaletti et al., 2014). Social relationships provide a particularly clear illustration: labels such as friend collapse a wide range of social ties that differ along dimensions such as intimacy, duration, obligation, and context. In some cases, a relationship cannot be cleanly categorized, as it may fall between friendship and romantic involvement. Even attempts to refine such relations using numerical scores or weights remain limited, as scalar values obscure the underlying qualitative distinctions.

3 Paradigm Shift and Opportunities in the LLM Era

The limitations of symbolic relation schemas motivate a reconsideration of how relationships are represented in KGs. Rather than simply refining categorical relation types, we argue for a shift toward representing relations directly in natural language. This shift is enabled by the widespread adoption of LLMs, which prefer contextual text and support downstream reasoning. At the same time, LLMs have also become central to data and knowledge creation workflows, enabling scalable synthesis and evaluation of relational information expressed in natural language.

3.1 Natural Language as Input in LLM-Based Downstream Tasks

Prior to the LLM era, information from KGs was typically incorporated into downstream NLP models such as classifiers, rankers, and scoring functions indirectly, through graph-derived features or embeddings rather than being consumed in natural language form. Consequently, KGs needed to be transformed into quantified representations compatible with these models. The most widely used and effective approaches for this transformation include translation-based models such as TransE and its variants (Bordes et al., 2013), bilinear factorization models such as ComplEx (Trouillon et al., 2016), as well as graph traversal– or random-walk–based feature generation methods (Xia et al., 2019). These approaches typically assume that relations are drawn from a finite, predefined set of categorical types, which serve to index model parameters, define feature spaces, or guide graph operations. Although recent studies have explored incorporating semantic information from text into KG embeddings or representations (Wang et al., 2022; Rao et al., 2024; Zhu et al., 2025b), such efforts are generally designed to augment structure-based representations derived from categorical relations rather than to replace them as the primary relational abstraction. As a result, relations in KGs are commonly constructed in this form to ensure broad compatibility with downstream

models.

In contrast, the LLM era has substantially blurred the boundaries between downstream tasks, as many are now approached through in-context learning via prompting (Dong et al., 2024). Under this paradigm, a single modeling framework can support classification, retrieval, reasoning, or generation, depending primarily on the provided context and instructions. Even tasks that were traditionally formulated as classification with categorical outputs are increasingly reframed as reasoning-oriented problems, in which models take natural language input, generate text-based intermediate reasoning (e.g., chain-of-thought), and ultimately produce textual outputs (Wang et al., 2024d; Diao et al., 2024). Empirical evidence from recent work indicates that such reasoning-oriented strategies, particularly chain-of-thought prompting with rich contextual information, often yield better performance than direct prediction (Wei et al., 2022; Wang et al., 2023). Consequently, LLM-based downstream tasks are no longer tightly coupled to fixed input–output formats and instead favor nuanced, descriptive knowledge provided as context, creating new opportunities for consuming relation knowledge in natural language rather than as simple categorical labels.

Within the KG literature, this shift has led to growing interest in LLM-based inference that leverages KG knowledge expressed in natural language form (Shu et al., 2024; Wang et al., 2024b; Mavromatis and Karypis, 2025). Instead of consuming relational triples in their raw symbolic format or graph-derived structured features, most studies adopt knowledge verbalization strategies that convert KG information into natural language descriptions, which are then incorporated into prompts as contextual input (Baek et al., 2023; Zhang et al., 2024). This KG-to-text conversion is motivated by the observation that LLMs reason more effectively over well-organized natural language representations than over simple symbolic, numerical, or vectorized inputs without verbalization. Consequently, representing relations in natural language within a KG provides a practical path to address the limitations of current representations, while also delivering richer context for downstream LLM reasoning.

3.2 LLM-Assisted Knowledge Resource Creation

Recent work increasingly uses LLMs to create new data and knowledge resources (Tan et al., 2024; Yang et al., 2025). Unlike the pre-LLM era, where reusable resources were typically designed around fixed formats and tightly structured representations to satisfy specific downstream models, LLM-based approaches often prioritize descriptive natural language representations that preserve contextual and semantic detail (Ding et al., 2024; Zheng et al., 2025).

A wide range of recent studies leverage LLMs to extract, synthesize, or articulate core information from complex documents, such as legal texts (Song et al., 2025a; Zhou et al., 2025) and scientific papers (Song et al., 2025b; Zheng et al., 2025). Examples include claim mining (Dmonte et al., 2024; Scirè et al., 2024), idea identification (Kumar et al., 2025; Su et al., 2025), and knowledge organization and summarization (Han et al., 2024; Feng et al., 2025), where models either extract and synthesize factual information or mine more complex hypothesis, ideas, and arguments. In these settings, the goal is not to only store symbolic labels, short textual terms, or numerical features, but also to create knowledge artifacts that retain nuance in natural language.

To enable quality control, many of these approaches combine LLM generation with additional mechanisms. On the generation side, prompts often incorporate explicit rules or constraints to guide content selection and level of abstraction (Cheng et al., 2024; Wang et al., 2024a; Mirbeygi and Beigy, 2025). On the evaluation side, LLM-based judges or filtering strategies are increasingly employed to assess relevance, consistency, or factuality, allowing large-scale construction of natural-language knowledge resources with acceptable reliability (Liu et al., 2024; Li et al., 2024b; Huang et al., 2025). Although such resources are less rigidly structured than traditional datasets, they are well aligned with LLM-centered downstream use, where reasoning over textual context is the dominant mode of computation.

Within the KG literature, LLMs have been increasingly explored as tools for KG construction. However, most existing work remains grounded in the traditional KG paradigm, using LLMs primarily to assist with ontology design and refinement, entity and relation extraction, or knowledge

completion via link prediction (Wei et al., 2023; Fathallah et al., 2024; Liang et al., 2024; Zhu et al., 2024). In these approaches, LLMs largely function as extractors or classifiers, while the resulting KGs continue to rely on predefined symbolic relation schemas. A small number of studies have begun to enrich entity representations with more nuanced natural-language descriptions (Zhu et al., 2025a), yet analogous advances for relation representation remain largely unexplored.

4 Principles for Structure–Expressiveness Trade-offs

Despite the expanded flexibility introduced by LLMs, fundamental challenges in leveraging KGs remain. In particular, LLMs are unable to ingest an entire KG within their context window, both due to input token limits and because excessively long contexts incur higher computational cost and may degrade performance (Du et al., 2025; Li et al., 2025b). As a result, Retrieval-Augmented Generation (RAG) is still the most common way to use LLM (Arslan et al., 2024), which first requires to identify and retrieve the most relevant subsets of knowledge for a given downstream task. Consequently, effective KG utilization in the LLM era still relies on mechanisms for efficient information retrieval, often supported by well-defined schemas and graph-based algorithms or embeddings that commonly take a fixed set of symbolic relations as graph edges.

This requirement highlights a persistent tension among **expressiveness**, referring to the level of detail in relational descriptions, **generalizability**, referring to the range of domains or tasks to which a KG can be adapted, and **structural regularity**, which typically relies on abstracted symbolic relation types. While natural language relations allow richer and more nuanced representations of relational knowledge, purely descriptive relations substantially complicate graph traversal, indexing, and retrieval. By contrast, categorical or symbolic relations continue to provide an efficient and effective means of organizing global graph structure and supporting path-based reasoning and relevance filtering.

These challenges suggest that the LLM era does not eliminate the need for structure in KGs, but instead reshapes how structure and language should interact within a knowledge representation framework. Rather than replacing symbolic relations entirely, natural-language relations must be designed in a way that preserves structural usability while enabling richer semantic expression.

Motivated by this trade-off, we outline a set of guiding principles for constructing natural-language-relation–based KGs that balance structural tractability with linguistic flexibility.

4.1 Implicit, Task-Aware Schemas Remain Necessary

Although natural-language relations relax the rigidity imposed by fixed symbolic schemas, some form of schema remains essential for effective use. Rather than being fully open-ended, relation generation and representation should be guided by implicit, task-aware schemas derived from clearly defined goals, domains, and downstream requirements. In particular, relations should be constrained to an appropriate semantic scope: introducing a large number of heterogeneous or cross-domain relation types often increases ambiguity and noise, ultimately hindering retrieval and reasoning rather than improving expressiveness. At the same time, schemas should not be overly explicit or rigid, as fixed ontologies can limit generalization across tasks and domains.

In practice, effective relation synthesis, especially when integrating knowledge from multiple sources, requires an explicit specification of the intended goal, scope, and domain of the resulting relations. When prompting LLMs to synthesize knowledge materials into concise relational descriptions, constraints or guiding rules are often necessary to control granularity, scope, and relevance. Such implicit, scoped schemas provide structural guidance without enforcing a rigid ontology, thereby balancing flexibility, generalization, and practical usability.

4.2 Natural-Language Relations Should Support Secondary Symbolic Inference

Closely related to the need for implicit, task-aware schemas, natural-language relation descriptions should not be treated as isolated language artifacts. Instead, they should be designed to support secondary inference into symbolic relations when required. In downstream settings that rely on symbolic representations, such as structured querying, aggregation, or graph-based retrieval, natural-language relations should be readily classifiable or mappable to existing symbolic categories.

From this perspective, natural-language relations are not intended to replace symbolic relations with an incompatible representation, but rather to serve as a richer carrier of relational information that remains compatible with symbolic abstraction when needed. This compatibility enables natural-language relations to preserve fine-grained semantic content while still supporting structure-dependent operations, thereby bridging flexible language-based representation and traditional symbolic reasoning.

4.3 Hybrid Use of Symbolic Relations and Natural-Language Descriptions for KG Construction

A hybrid design that combines a small set of symbolic relations with natural-language relation descriptions provides a practical compromise between structural stability and representational flexibility. Symbolic relations establish an initial global backbone for the graph, enabling efficient organization, traversal, and retrieval, while natural-language descriptions capture nuanced, contextual, or evolving aspects of relationships that are difficult to encode in fixed categorical form. For example, two entities may be connected by a symbolic friend relation, accompanied by a natural-language description that specifies the nature and context.

In this hybrid setting, the symbolic relation schema is best designed to be generic, domain-agnostic, and minimally committal, serving primarily as a guiding structural scaffold rather than a complete semantic specification. Such symbolic relations provide coarse-grained relational anchors that support graph-level operations, while leaving fine-grained semantics to be expressed in natural language. Importantly, because rich relational meaning is preserved in textual form, symbolic relations can be more easily revised, refined, or restructured, such as being split, merged, or redefined, based on insights derived from natural-language descriptions. This design enables the KG to evolve without requiring disruptive changes to its global structure, supporting both scalability and long-term adaptability. This hybrid design directly reflects the first two principles: symbolic relations provide an implicit, task-aware structural schema, while natural-language descriptions preserve sufficient semantic detail to enable secondary inference into symbolic relations when needed.

5 Future Research Directions

While the principles outlined above provide high-level guidance for integrating natural-language relations into KGs, translating these principles into practical systems raises a number of open challenges and realistic considerations. Constructing and maintaining KGs that balance structural tractability, domain and task generalizability, and contextual nuance requires coordinated progress across data acquisition, representation, retrieval, inference, and evaluation. In this section, we outline several promising directions for future research aimed at overcoming these challenges and advancing the development of KGs grounded in natural-language relations.

5.1 Constructing Natural-Language Relation KGs under Heterogeneous Sources

A fundamental challenge in constructing natural-language-relation KGs from raw text is the prevalence of heterogeneous and conflicting information across sources. Classical computer science literature typically addresses this problem under the frameworks of data fusion, knowledge fusion, or truth discovery, where conflicting facts are reconciled by estimating source reliability or latent truth, often with the goal of selecting the most "correct" value rather than preserving disagreement (Wang et al., 2024c). While effective for many objective settings, such approaches are less suitable for domains such as history or social narratives, where conflicting accounts may all be plausible and should be retained rather than forcibly resolved. In the context of natural-language-relation KGs, a key research question is how to synthesize and preserve conflicting relational descriptions in a concise yet faithful manner (Hofer et al., 2024), leveraging LLMs' reasoning capabilities without collapsing uncertainty.

5.2 LLM-Driven Refinement of Existing KGs

Beyond constructing KGs from raw text, an alternative and complementary direction is to start from existing KGs and leverage additional data sources, or even the internal knowledge and reasoning capabilities of LLMs, to enrich symbolic relations with more nuanced natural-language descriptions. For example, recent studies have explored distilling relational knowledge from LLMs in natural language for KG completion, while still operating within a symbolic relation paradigm and targeting

symbolic relation prediction (Li et al., 2024a; Yao et al., 2025b).

However, enriching existing KGs with natural-language relations introduces additional challenges when the underlying symbolic schema itself is limited or misaligned with the newly added semantic content. In such cases, KG modification cannot be restricted to appending information on top of fixed relation types, but may require simultaneous refinement of the relation schema, including revising, splitting, or merging symbolic relations based on richer relational evidence. This raises open research questions about how to coordinate LLM-driven semantic enrichment with controlled schema evolution, while preserving structural consistency and downstream usability.

5.3 Contextual Relation–Based Traversal and Reasoning

Most existing KG traversal methods rely primarily on graph structure, such as neighborhood expansion, path-based reasoning, or graph embedding similarity, treating relations as fixed symbolic labels or parameters (Zhu et al., 2025b). In natural-language-relation KGs, however, the reasoning capabilities of LLMs enable traversal decisions to be made dynamically based on the semantic content of relations and the evolving context of the task. Rather than treating traversal as a fixed graph operation, LLMs can iteratively select relevant nodes, relations, or subgraphs by reasoning over natural-language descriptions at each step. This opens the possibility of multi-stage reasoning pipelines, where LLMs are repeatedly invoked to guide retrieval, assess relevance, and refine the search space over the KG.

5.4 KG Embeddings

Another promising research direction is the development of joint structural–semantic embedding methods for KGs with natural-language relations. Prior work has incorporated semantic information into KG embeddings by leveraging relation names, textual descriptions of entities, or auxiliary text encoders, demonstrating improvements over purely symbolic representations (Wang et al., 2017; Yao et al., 2019; Cao et al., 2024). However, the relation semantics considered in these approaches are typically simple, static, and limited in scope, and are largely treated as auxiliary signals rather than primary relational representations. In contrast, the hybrid design advocated in Principle 3, where a small set of generic symbolic relations provides global structure and natural-language descriptions, opens new opportunities for embedding models that jointly capture structural regularities and nuanced semantic content.

5.5 KG Evaluation

Evaluating KGs with natural-language relations presents a significant challenge because relations are no longer fixed symbolic types or simple lexical surfaces, but synthesized descriptions produced by LLMs. As a result, traditional KG metrics such as precision and recall over triples are insufficient, as they fail to capture key qualities of natural-language relations like factual faithfulness, core information coverage, and conciseness. Recent work has therefore explored LLM-as-a-judge frameworks that assess generated text using rubric-based prompting (Liu et al., 2023; Es et al., 2024; Li et al., 2025a). Adapting such judge-based approaches to KG evaluation introduces new questions around consistency, bias, and scalability, motivating evaluation protocols tailored to natural-language-relation KGs.

6 Conclusion

KGs have long relied on symbolic relation schemas to enable scalable construction and broad computational use. The rise of LLMs reshapes this foundation by making natural language the primary interface for downstream inference. This paper argues that treating relations as natural-language descriptions, supported by implicit and hybrid structural constraints, provides a practical way to align KGs with current LLM usage paradigms while preserving nuanced relational information. Rather than viewing LLMs only as more efficient tools for populating conventional schemas, we advocate rethinking the relations themselves as contextual language descriptions that remain supported by a minimal symbolic backbone for retrieval and traversal. Building on this perspective, we further advocate the development of new methods and empirical research practices under this paradigm.

7 Limitations

This position paper has three main limitations.

First, we focus primarily on the paradigm shift for relations in KGs, while leaving analogous changes for entities less examined. In the LLM era, entity extraction and presentation also require

substantial reconsideration. By addressing only the relational side, our discussion captures a partial view of the broader KG transformation. Future research should jointly rethink entity and relation representations so that both components align with prompting-based inference and LLM-centered knowledge usage.

Second, our arguments are developed largely around textual KGs constructed from language corpora. KGs may also be grounded in images, audio, or other modalities, yet this work does not explore such multi-modal settings. Relations expressed across heterogeneous modalities introduce additional challenges that differ from purely text-based scenarios. Consequently, the conclusions here should be interpreted within the limits of language-centered KG construction rather than as a general account of all KG forms.

Finally, this is a position paper that provides conceptual analysis based on a synthesis of prior literature, without presenting new large-scale experiments. The discussion therefore remains high-level and partly theoretical, and practical effectiveness of natural-language relations is not empirically validated here. The proposed principles serve as guidance rather than tested protocols, and implementing them in real systems will require dedicated methodological research and systematic evaluation.

References

Muhammad Arslan, Hussam Ghanem, Saba Munawar, and Christophe Cruz. 2024. A survey on rag with llms. *Procedia computer science*, 246:3781–3790.

Jinheon Baek, Alham Fikri Aji, and Amir Saffari. 2023. Knowledge-augmented language model prompting for zero-shot knowledge graph question answering. In *Proceedings of the 1st Workshop on Natural Language Reasoning and Structured Explanations (NLRSE)*, pages 78–106.

Junwei Bao, Nan Duan, Zhao Yan, Ming Zhou, and Tiejun Zhao. 2016. Constraint-based question answering with knowledge graph. In *Proceedings of COLING 2016, the 26th international conference on computational linguistics: technical papers*, pages 2503–2514.

Russa Biswas, Lucie-Aimée Kaffee, Michael Cochez, Stefania Dumbrava, Theis E Jendal, Matteo Lissandrini, Vanessa Lopez, Eneldo Loza Mencía, Heiko Paulheim, Harald Sack, and 1 others. 2023. Knowledge graph embeddings: open challenges and opportunities. *Transactions on Graph Data and Knowledge*, 1(1):4–1.

Stefano Boccaletti, Ginestra Bianconi, Regino Criado, Charo I Del Genio, Jesús Gómez-Gardenes, Miguel Romance, Irene Sendina-Nadal, Zhen Wang, and Massimiliano Zanin. 2014. The structure and dynamics of multilayer networks. *Physics reports*, 544(1):1–122.

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating embeddings for modeling multi-relational data. *Advances in neural information processing systems*, 26.

Borui Cai, Yong Xiang, Longxiang Gao, He Zhang, Yunfeng Li, and Jianxin Li. 2023. Temporal knowledge graph completion: a survey. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 6545–6553.

Jiahang Cao, Jinyuan Fang, Zaiqiao Meng, and Shangsong Liang. 2024. Knowledge graph embedding: A survey from the perspective of representation spaces. *ACM Computing Surveys*, 56(6):1–42.

Lei Cao, Huijun Zhang, and Ling Feng. 2020. Building and using personal knowledge graph to improve suicidal ideation detection on social media. *IEEE Transactions on Multimedia*, 24:87–102.

Kewei Cheng, Nesreen K Ahmed, Theodore L Willke, and Yizhou Sun. 2024. Structure guided prompt: Instructing large language model in multi-step reasoning by exploring graph structure of the text. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 9407–9430.

Yuanfei Dai, Shiping Wang, Neal N Xiong, and Wenzhong Guo. 2020. A survey on knowledge graph embedding: Approaches, applications and benchmarks. *Electronics*, 9(5):750.

Shizhe Diao, Pengcheng Wang, Yong Lin, Rui Pan, Xiang Liu, and Tong Zhang. 2024. Active prompting with chain-of-thought for large language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1330–1350.

Bosheng Ding, Chengwei Qin, Ruochen Zhao, Tianze Luo, Xinze Li, Guizhen Chen, Wenhan Xia, Junjie Hu, Luu Anh Tuan, and Shafiq Joty. 2024. Data augmentation using llms: Data perspectives, learning paradigms and challenges. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 1679–1705.

Alphaeus Dmonte, Roland Oruche, Marcos Zampieri, Prasad Calyam, and Isabelle Augenstein. 2024. Claim verification in the age of large language models: A survey. *arXiv preprint arXiv:2408.14317*.

Qingxiu Dong, Lei Li, Damai Dai, Ce Zheng, Jingyuan Ma, Rui Li, Heming Xia, Jingjing Xu, Zhiyong Wu, Baobao Chang, and 1 others. 2024. A survey on

in-context learning. In *Proceedings of the 2024 conference on empirical methods in natural language processing*, pages 1107–1128.

Yufeng Du, Minyang Tian, Srikanth Ronanki, Subendhu Rongali, Sravan Babu Bodapati, Aram Galstyan, Azton Wells, Roy Schwartz, Eliu A Huerta, and Hao Peng. 2025. Context length alone hurts LLM performance despite perfect retrieval. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 23281–23298. Association for Computational Linguistics.

Mustafa Emirbayer and Jeff Goodwin. 1994. Network analysis, culture, and the problem of agency. *American journal of sociology*, 99(6):1411–1454.

Shahul Es, Jithin James, Luis Espinosa Anke, and Steven Schockaert. 2024. Ragas: Automated evaluation of retrieval augmented generation. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, pages 150–158.

Nadeen Fathallah, Arunav Das, Stefano De Giorgis, Andrea Poltronieri, Peter Haase, and Liubov Kovriguina. 2024. Neon-gpt: a large language model-powered pipeline for ontology learning. In *European Semantic Web Conference*, pages 36–50. Springer.

Christiane Fellbaum. 2010. Wordnet. In *Theory and applications of ontology: computer applications*, pages 231–243. Springer.

Yu Feng, Wenkang An, Hao Wang, and Zhen Yin. 2025. Enhancing scientific literature summarization via contrastive learning and chain-of-thought prompting. *Scientometrics*, 130(8):4773–4799.

Dieter Fensel, Umutcan Simsek, Kevin Angele, Elwin Huaman, Elias Kärle, Oleksandra Panasiuk, Ioan Toma, Jürgen Umbrich, and Alexander Wahler. 2020. *Knowledge graphs*. Springer.

Mohamed H Gad-Elrab, Daria Stepanova, Trung-Kien Tran, Heike Adel, and Gerhard Weikum. 2020. Excut: Explainable embedding-based clustering over knowledge graphs. In *International Semantic Web Conference*, pages 218–237. Springer.

Xiou Ge, Yun Cheng Wang, Bin Wang, C-C Jay Kuo, and 1 others. 2024. Knowledge graph embedding: An overview. *APSIPA Transactions on Signal and Information Processing*, 13(1).

László Grad-Gyenge, Attila Kiss, and Peter Filzmoser. 2017. Graph embedding based recommendation techniques on the knowledge graph. In *Adjunct publication of the 25th conference on user modeling, adaptation and personalization*, pages 354–359.

Saiping Guan, Xueqi Cheng, Long Bai, Fujun Zhang, Zixuan Li, Yutao Zeng, Xiaolong Jin, and Jiafeng Guo. 2022. What is event knowledge graph: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 35(7):7569–7589.

Binglan Han, Teo Susnjak, and Anuradha Mathrani. 2024. Automating systematic literature reviews with retrieval-augmented generation: a comprehensive overview. *Applied Sciences*, 14(19):9103.

Kanyao Han, Pingjing Yang, Shubhanshu Mishra, and Jana Diesner. 2020. Wikicssh: extracting computer science subject headings from wikipedia. In *International Conference on Theory and Practice of Digital Libraries*, pages 207–218. Springer.

Qizhen He, Liang Wu, Yida Yin, and Heming Cai. 2020. Knowledge-graph augmented word representations for named entity recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 7919–7926.

Marvin Hofer, Daniel Obraczka, Alieh Saeedi, Hanna Köpcke, and Erhard Rahm. 2024. Construction of knowledge graphs: Current state and challenges. *Information*, 15(8).

Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard De Melo, Claudio Gutierrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, and 1 others. 2021. Knowledge graphs. *ACM Computing Surveys (Csur)*, 54(4):1–37.

Hui Huang, Xingyuan Bu, Hongli Zhou, Yingqi Qu, Jing Liu, Muyun Yang, Bing Xu, and Tiejun Zhao. 2025. An empirical study of llm-as-a-judge for llm evaluation: Fine-tuned judge model is not a general substitute for gpt-4. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 5880–5895.

Sandeep Kumar, Tirthankar Ghosal, Vinayak Goyal, and Asif Ekbal. 2025. Can large language models unlock novel scientific research ideas? In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 33551–33575.

Dawei Li, Bohan Jiang, Liangjie Huang, Alimohammad Beigi, Chengshuai Zhao, Zhen Tan, Amrita Bhattacharjee, Yuxuan Jiang, Canyu Chen, Tianhao Wu, and 1 others. 2025a. From generation to judgment: Opportunities and challenges of llm-as-a-judge. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 2757–2791.

Dawei Li, Zhen Tan, Tianlong Chen, and Huan Liu. 2024a. Contextualization distillation from large language model for knowledge graph completion. In *Findings of the Association for Computational Linguistics: EACL 2024*, pages 458–477.

Haitao Li, Qian Dong, Junjie Chen, Huixue Su, Yujia Zhou, Qingyao Ai, Ziyi Ye, and Yiqun Liu. 2024b. Llms-as-judges: a comprehensive survey on llm-based evaluation methods. *arXiv preprint arXiv:2412.05579*.

Nan Li, Zhihao Yang, Jian Wang, and Hongfei Lin. 2024c. Drug–target interaction prediction using knowledge graph embedding. *Iscience*, 27(6).

Tianle Li, Ge Zhang, Quy Duc Do, Xiang Yue, and Wenhu Chen. 2025b. Long-context LLMs struggle with long in-context learning. *Transactions on Machine Learning Research*.

Xinxin Liang, Zuoxu Wang, Mingrui Li, and Zhijie Yan. 2024. A survey of llm-augmented knowledge graph construction and application in complex product design. *Procedia CIRP*, 128:870–875.

Fidan Limani, Atif Latif, and Klaus Tochtermann. 2021. Scholarly artifacts knowledge graph: use cases for digital libraries. In *Research Conference on Metadata and Semantics Research*, pages 245–256. Springer.

Yang Liu, Dan Iter, Yichong Xu, Shuohang Wang, Ruochen Xu, and Chenguang Zhu. 2023. G-eval: Nlg evaluation using gpt-4 with better human alignment. *arXiv preprint arXiv:2303.16634*.

Yuxuan Liu, Tianchi Yang, Shaohan Huang, Zihan Zhang, Haizhen Huang, Furu Wei, Weiwei Deng, Feng Sun, and Qi Zhang. 2024. Calibrating llm-based evaluator. In *Proceedings of the 2024 joint international conference on computational linguistics, language resources and evaluation (lrec-coling 2024)*, pages 2638–2656.

Jose L Martinez-Rodriguez, Ivan López-Arévalo, and Ana B Rios-Alvarado. 2018. Openie-based approach for knowledge graph construction from text. *Expert Systems with Applications*, 113:339–355.

Costas Mavromatis and George Karypis. 2025. Gnn-rag: Graph neural retrieval for efficient large language model reasoning on knowledge graphs. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 16682–16699.

Mohaddeseh Mirbeygi and Hamid Beigy. 2025. Prompt guided diffusion for controllable text generation. In *Proceedings of the Tenth Workshop on Noisy and User-generated Text*, pages 78–84.

Christina Niklaus, Matthias Cetto, André Freitas, and Siegfried Handschuh. 2018. A survey on open information extraction. *arXiv preprint arXiv:1806.05599*.

Tobias Aanderaa Opsahl. 2024. Fact or fiction? improving fact verification with knowledge graphs through simplified subgraph retrievals. In *Proceedings of the Seventh Fact Extraction and VERification Workshop (FEVER)*, pages 307–316.

Liu Pai, Wenyang Gao, Wenjie Dong, Lin Ai, Ziwei Gong, Songfang Huang, Li Zongsheng, Ehsan Hoque, Julia Hirschberg, and Yue Zhang. 2024. A survey on open information extraction from rule-based model to large language model. *Findings of the association for computational linguistics: EMNLP 2024*, pages 9586–9608.

Sachin Pawar, Girish K Palshikar, and Pushpak Bhattacharyya. 2017. Relation extraction: A survey. *arXiv preprint arXiv:1712.05191*.

Jia Qu. 2022. A review on the application of knowledge graph technology in the medical field. *Scientific Programming*, 2022(1):3212370.

Gorjan Radevski, Kiril Gashteovski, Chia-Chien Hung, Carolin Lawrence, and Goran Glavaš. 2023. Linking surface facts to large-scale knowledge graphs. *arXiv preprint arXiv:2310.14909*.

Qiang Rao, Tiejun Wang, Xiaoran Guo, Kaijie Wang, and Yue Yan. 2024. Knowledge graph completion using a pre-trained language model based on categorical information and multi-layer residual attention. *Applied Sciences*, 14(11):4453.

Hongyu Ren, Ali Mousavi, Anil Pacaci, Shihabur R Chowdhury, Jason Mohoney, Ihab F Ilyas, Yunyao Li, and Theodoros Rekatsinas. 2023. Fact ranking over large-scale knowledge graphs with reasoning embedding models. *Data Engineering*, page 124.

Julien Romero and Simon Razniewski. 2023. Mapping and cleaning open commonsense knowledge bases with generative translation. In *International Semantic Web Conference*, pages 368–387. Springer.

Stefan Schulz, Anand Kumar, and Thomas Bittner. 2006. Biomedical ontologies: what part-of is and isn't. *Journal of Biomedical Informatics*, 39(3):350–361.

Alessandro Scirè, Karim Ghonim, and Roberto Navigli. 2024. Fenice: Factuality evaluation of summarization based on natural language inference and claim extraction. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 14148–14161.

Zeinab Shokrzadeh, Mohammad-Reza Feizi-Derakhshi, Mohammad-Ali Balafar, and Jamshid Bagherzadeh Mohasefi. 2024. Knowledge graph-based recommendation system enhanced by neural collaborative filtering and knowledge graph embedding. *Ain Shams Engineering Journal*, 15(1):102263.

Dong Shu, Tianle Chen, Mingyu Jin, Chong Zhang, Mengnan Du, and Yongfeng Zhang. 2024. Knowledge graph large language model (kg-llm) for link prediction. *Proceedings of Machine Learning Research*, 260(62).

Barry Smith, Werner Ceusters, Bert Klagges, Jacob Köhler, Anand Kumar, Jane Lomax, Chris Mungall, Fabian Neuhaus, Alan L Rector, and Cornelius Rosse. 2005. Relations in biomedical ontologies. *Genome biology*, 6(5):R46.

Yumei Song, Yongbin Qin, Ruizhang Huang, Yanping Chen, and Chuan Lin. 2025a. Legal text summarization via judicial syllogism with large language models. *Journal of King Saud University Computer and Information Sciences*, 37(5):111.

Zihan Song, Gyo-Yeob Hwang, Xin Zhang, Shan Huang, and Byung-Kwon Park. 2025b. A scientific-article key-insight extraction system based on multi-actor of fine-tuned open-source large language models. *Scientific Reports*, 15(1):1608.

Thomas Steiner, Ruben Verborgh, Raphaël Troncy, Joaquim Gabarro, and Rik Van de Walle. 2012. Adding realtime coverage to the google knowledge graph. In *ISWC (Posters & Demos)*.

Haoyang Su, Renqi Chen, Shixiang Tang, Zhenfei Yin, Xinzhe Zheng, Jinzhe Li, Biqing Qi, Qi Wu, Hui Li, Wanli Ouyang, and 1 others. 2025. Many heads are better than one: Improved scientific idea generation by a llm-based multi-agent system. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 28201–28240.

Zhu Sun, Jie Yang, Jie Zhang, Alessandro Bozzon, Long-Kai Huang, and Chi Xu. 2018. Recurrent knowledge graph embedding for effective recommendation. In *Proceedings of the 12th ACM conference on recommender systems*, pages 297–305.

Zhen Tan, Dawei Li, Song Wang, Alimohammad Beigi, Bohan Jiang, Amrita Bhattacharjee, Mansooreh Karami, Jundong Li, Lu Cheng, and Huan Liu. 2024. Large language models for data annotation and synthesis: A survey. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 930–957.

Shiyu Tian, Yangyang Luo, Tianze Xu, Caixia Yuan, Huixing Jiang, Chen Wei, and Xiaojie Wang. 2024. Kg-adapter: Enabling knowledge graph integration in large language models through parameter-efficient fine-tuning. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 3813–3828.

Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. 2016. Complex embeddings for simple link prediction. In *International conference on machine learning*, pages 2071–2080. PMLR.

Deepak R Unni, Sierra AT Moxon, Michael Bada, Matthew Brush, Richard Bruskiewich, J Harry Caufield, Paul A Clemons, Vlado Dancik, Michel Dumontier, Karamarie Fecho, and 1 others. 2022. Biolink model: A universal schema for knowledge graphs in clinical, biomedical, and translational science. *Clinical and translational science*, 15(8):1848–1855.

Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Communications of the ACM*, 57(10):78–85.

Boshi Wang, Sewon Min, Xiang Deng, Jiaming Shen, You Wu, Luke Zettlemoyer, and Huan Sun. 2023. Towards understanding chain-of-thought prompting: An empirical study of what matters. In *Proceedings of the 61st annual meeting of the association for computational linguistics (volume 1: Long papers)*, pages 2717–2739.

Jianing Wang, Qiushi Sun, Xiang Li, and Ming Gao. 2024a. Boosting language models reasoning with chain-of-knowledge prompting. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4958–4981.

Liang Wang, Wei Zhao, Zhuoyu Wei, and Jingming Liu. 2022. Simkgc: Simple contrastive knowledge graph completion with pre-trained language models. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4281–4294.

Quan Wang, Zhendong Mao, Bin Wang, and Li Guo. 2017. Knowledge graph embedding: A survey of approaches and applications. *IEEE transactions on knowledge and data engineering*, 29(12):2724–2743.

Shanshan Wang, Chun Zhang, and Ning Zhang. 2024b. Mgsa: multi-granularity graph structure attention for knowledge graph-to-text generation. In *2024 12th International Conference on Information Systems and Computing Technology (ISCTech)*, pages 1–5. IEEE.

Shuang Wang, He Zhang, Quan Z Sheng, Xiaoping Li, Zhu Sun, Taotao Cai, Wei Emma Zhang, Jian Yang, and Qing Gao. 2024c. A survey on truth discovery: Concepts, methods, applications, and opportunities. *IEEE Transactions on Big Data*, 11(2):314–332.

Yajing Wang, Jiazhen Hu, and Zongwei Luo. 2024d. Cot-tree: An efficient prompting strategy for large language model classification reasoning. In *2024 IEEE International Conference on Big Data (BigData)*, pages 5450–5458. IEEE.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, and 1 others. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.

Yanbin Wei, Qiushi Huang, Yu Zhang, and James Kwok. 2023. Kicgpt: Large language model with knowledge in context for knowledge graph completion. In *Findings of the association for computational linguistics: EMNLP 2023*, pages 8667–8683.

Lisa Wenige, Claus Stadler, Simon Bin, Lorenz Bühmann, Kurt Junghanns, and Michael Martin. 2020. Automatic subject indexing with knowledge graphs. In *LASCAR Workshop at the Extended Semantic Web Conference (ESWC)*.

Feng Xia, Jiaying Liu, Hansong Nie, Yonghao Fu, Liangtian Wan, and Xiangjie Kong. 2019. Random walks: A review of algorithms and applications. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 4(2):95–107.

Han Xiao, Yidong Chen, and Xiaodong Shi. 2019. Knowledge graph embedding based on multi-view clustering framework. *IEEE Transactions on Knowledge and Data Engineering*, 33(2):585–596.

Chenyan Xiong, Russell Power, and Jamie Callan. 2017. Explicit semantic ranking for academic search via knowledge graph embedding. In *Proceedings of the 26th international conference on world wide web*, pages 1271–1279.

Wenli Yang, Lilian Some, Michael Bain, and Byeong Kang. 2025. A comprehensive survey on integrating large language models with knowledge-based methods. *Knowledge-Based Systems*, page 113503.

Mohammad Yani and Adila Alfa Krisnadhi. 2021. Challenges, techniques, and trends of simple knowledge graph question answering: a survey. *Information*, 12(7):271.

Liang Yao, Chengsheng Mao, and Yuan Luo. 2019. Kg-bert: Bert for knowledge graph completion. *arXiv preprint arXiv:1909.03193*.

Liang Yao, Jiazhen Peng, Chengsheng Mao, and Yuan Luo. 2025a. Exploring large language models for knowledge graph completion. In *ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5.

Liang Yao, Jiazhen Peng, Chengsheng Mao, and Yuan Luo. 2025b. Exploring large language models for knowledge graph completion. In *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE.

Alexander Yates, Michele Banko, Matthew Broadhead, Michael J Cafarella, Oren Etzioni, and Stephen Soderland. 2007. Textrunner: open information extraction on the web. In *Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*, pages 25–26.

Mohamad Zamini, Hassan Reza, and Minou Rabiei. 2022. A review of knowledge graph completion. *Information*, 13(8):396.

Qinggang Zhang, Junnan Dong, Hao Chen, Daochen Zha, Zailiang Yu, and Xiao Huang. 2024. Knowgpt: knowledge graph based prompting for large language models. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, pages 6052–6080.

Tianshi Zheng, Zheye Deng, Hong Ting Tsang, Weiqi Wang, Jiaxin Bai, Zihao Wang, and Yangqiu Song. 2025. From automation to autonomy: A survey on large language models in scientific discovery. *arXiv preprint arXiv:2505.13259*.

Dongzhuoran Zhou, Baifan Zhou, Zhuoxun Zheng, Ahmet Soylu, Ognjen Savkovic, Egor V Kostylev, and Evgeny Kharlamov. 2022. Schere: Schema reshaping for enhancing knowledge graph construction. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 5074–5078.

Zhi Zhou, Kun-Yang Yu, Shi-Yu Tian, Xiao-Wen Yang, Jiang-Xin Shi, Pengxiao Song, Yi-Xuan Jin, Lan-Zhe Guo, and Yu-Feng Li. 2025. Lawgpt: Knowledge-guided data generation and its application to legal llm. *arXiv preprint arXiv:2502.06572*.

Kun Zhu, Lizi Liao, Yuxuan Gu, Lei Huang, Xiaocheng Feng, and Bing Qin. 2025a. Context-aware hierarchical taxonomy generation for scientific papers via llm-guided multi-aspect clustering. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 15627–15645.

Quntao Zhu, Mengfan Li, Yuanjun Gao, Yao Wan, Xuanhua Shi, and Hai Jin. 2025b. Text-augmented long-term relation dependency learning for knowledge graph representation. *High-Confidence Computing*, page 100315.

Xiangrong Zhu, Yuexiang Xie, Yi Liu, Yaliang Li, and Wei Hu. 2025c. Knowledge graph-guided retrieval augmented generation. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 8912–8924.

Yuqi Zhu, Xiaohan Wang, Jing Chen, Shuofei Qiao, Yixin Ou, Yunzhi Yao, Shumin Deng, Huajun Chen, and Ningyu Zhang. 2024. Llms for knowledge graph construction and reasoning: Recent capabilities and future opportunities. *World Wide Web*, 27(5):58.