

# LLM-EMPOWERED KNOWLEDGE GRAPH CONSTRUCTION: A SURVEY

**Anonymous authors**

Paper under double-blind review

## ABSTRACT

Knowledge Graphs (KGs) have long served as a fundamental infrastructure for structured knowledge representation and reasoning. With the advent of Large Language Models (LLMs), the construction of KGs has entered a new paradigm—shifting from rule-based and statistical pipelines to language-driven and generative frameworks. This survey provides a comprehensive overview of recent progress in **LLM-empowered knowledge graph construction**, systematically analyzing how LLMs reshape the classical three-layered pipeline of ontology engineering, knowledge extraction, and knowledge fusion.

We first revisit traditional KG methodologies to establish conceptual foundations, and then review emerging LLM-driven approaches from two complementary perspectives: *schema-based* paradigms, which emphasize structure, normalization, and consistency; and *schema-free* paradigms, which highlight flexibility, adaptability, and open discovery. Across each stage, we synthesize representative frameworks, analyze their technical mechanisms, and identify their limitations.

Finally, the survey outlines key trends and future research directions, including KG-based reasoning for LLMs, dynamic knowledge memory for agentic systems, and multimodal KG construction. Through this systematic review, we aim to clarify the evolving interplay between LLMs and knowledge graphs, bridging symbolic knowledge engineering and neural semantic understanding toward the development of adaptive, explainable, and intelligent knowledge systems.

## 1 INTRODUCTION

Knowledge Graphs (KGs) have long served as a cornerstone for representing, integrating, and reasoning over structured knowledge. They provide a unified semantic foundation that underpins a wide range of intelligent applications, such as semantic search, question answering, and scientific discovery. Conventional KG construction pipelines are typically composed of three major components: **ontology engineering**, **knowledge extraction**, and **knowledge fusion**. Despite their success in enabling large-scale knowledge organization, traditional paradigms (e.g., [Zhong et al. \(2023\)](#); [Zhao et al. \(2024\)](#)) continue to face three enduring challenges: (1) *Scalability and data sparsity*, as rule-based and supervised systems often fail to generalize across domains; (2) *Expert dependency and rigidity*, since schema and ontology design require substantial human intervention and lack adaptability; and (3) *Pipeline fragmentation*, where the disjoint handling of construction stages causes cumulative error propagation. These limitations hinder the development of self-evolving, large-scale, and dynamic KGs.

The advent of **Large Language Models (LLMs)** introduces a transformative paradigm for overcoming these bottlenecks. Through large-scale pretraining and emergent generalization capabilities, LLMs enable three key mechanisms: (1) *Generative knowledge modeling*, synthesizing structured representations directly from unstructured text; (2) *Semantic unification*, integrating heterogeneous knowledge sources through natural language grounding; and (3) *Instruction-driven orchestration*, coordinating complex KG construction workflows via prompt-based interaction. Consequently, LLMs are evolving beyond traditional text-processing tools into *cognitive engines* that seamlessly bridge natural language and structured knowledge (e.g., [Zhu et al. \(2024b\)](#); [Zhang & Soh \(2024\)](#)). This evolution marks a paradigm shift from *rule-driven, pipeline-based systems* toward *LLM-driven, unified, and adaptive frameworks*, where knowledge acquisition, organization, and reasoning emerge as interdependent processes within a generative and self-refining ecosystem ([Pan et al., 2024](#)).

In light of these rapid advances, this paper presents a comprehensive **survey** of LLM-driven knowledge graph construction. We systematically review recent research spanning ontology engineering, knowledge extraction, and fusion, analyze emerging methodological paradigms, and highlight open challenges and future directions at the intersection of LLMs and knowledge representation.

The remainder of this paper is organized as follows:

- **Section 2** introduces the foundations of traditional knowledge graph construction, covering ontology engineering, knowledge extraction, and fusion techniques prior to the LLM era.
- **Section 3** reviews LLM-enhanced ontology construction, encompassing both top-down paradigms (LLMs as ontology assistants) and bottom-up paradigms (KGs for LLMs).
- **Section 4** presents LLM-driven knowledge extraction, comparing schema-based and schema-free methodologies.
- **Section 5** discusses LLM-powered knowledge fusion, focusing on schema-level, instance-level, and hybrid frameworks.
- **Section 6** explores future research directions, including KG-based reasoning, dynamic knowledge memory, and multimodal KG construction.

## 2 PRELIMINARIES

The construction of Knowledge Graphs (KGs) traditionally follows a **three-layered pipeline** comprising *ontology engineering*, *knowledge extraction*, and *knowledge fusion*. Prior to the advent of Large Language Models (LLMs), these stages were implemented through rule-based, statistical, and symbolic approaches. This section briefly reviews these conventional methodologies to establish context for the subsequent discussion on LLM-empowered KG construction.

### 2.1 ONTOLOGY ENGINEERING

Ontology Engineering (OE) involves the formal specification of domain concepts, relationships, and constraints. In the pre-LLM era, ontologies were primarily **manually constructed by domain experts**, often supported by semantic web tools such as *Protégé* and guided by established methodologies including *METHONTOLOGY* and *On-To-Knowledge*. These systematic processes emphasized conceptual rigor and logical consistency but required extensive expert intervention.

As summarized by [Zouaq & Nkambou \(2010\)](#), ontology design during this period was characterized by strong human supervision and limited scalability. Subsequent semi-automatic approaches—collectively known as *ontology learning*—sought to derive ontological structures from textual corpora, as reviewed in [Asim et al. \(2018\)](#). However, even advanced frameworks such as *NeOn* struggled with ontology evolution, modular reuse, and dynamic adaptation. As highlighted by [Kotis et al. \(2020\)](#), traditional OE frameworks offered precision and formal soundness but lacked flexibility and efficiency for large-scale or continuously evolving knowledge domains.

### 2.2 KNOWLEDGE EXTRACTION

Knowledge Extraction (KE) aims to identify **entities, relations, and attributes** from unstructured or semi-structured data. Early approaches relied on handcrafted linguistic rules and pattern matching, which provided interpretability but were brittle and domain-specific. The evolution from symbolic and rule-based systems to statistical and neural methods has been systematically summarized in [Pai et al. \(2024\)](#).

The advent of deep learning architectures, such as *BiLSTM-CRF* and *Transformer*-based models, marked a paradigm shift toward data-driven feature learning, as discussed by [Yang et al. \(2022b\)](#). Comprehensive analyses such as [Detroja et al. \(2023\)](#) further categorize supervised, weakly supervised, and unsupervised relation extraction paradigms, emphasizing their dependence on annotated data and limited cross-domain generalization.

In summary, traditional KE methods established the technical foundation for modern extraction pipelines but remained constrained by **data scarcity, weak generalization, and cumulative error propagation**—limitations that motivate the LLM-driven paradigms discussed in later sections.

## 2.3 KNOWLEDGE FUSION

Knowledge Fusion (KF) focuses on integrating heterogeneous knowledge sources into a coherent and consistent graph by resolving issues of duplication, conflict, and heterogeneity. A central sub-task is **entity alignment**, which determines whether entities from different datasets refer to the same real-world object.

Classical approaches relied on lexical and structural similarity measures, as reviewed in Zeng et al. (2021). The introduction of representation learning enabled embedding-based techniques that align entities within shared vector spaces, improving scalability and automation, as surveyed by Zhu et al. (2024a). Domain-specific applications, such as Yang et al. (2022a), demonstrate multi-feature fusion strategies combining structural, attribute, and relational similarities. Other graph-level models, such as Liu et al. (2022), further integrate semantic cues to enhance alignment robustness.

Despite these advancements, traditional fusion pipelines continue to struggle with **semantic heterogeneity, large-scale integration, and dynamic knowledge updating**—challenges that contemporary LLM-based fusion frameworks are increasingly designed to address.

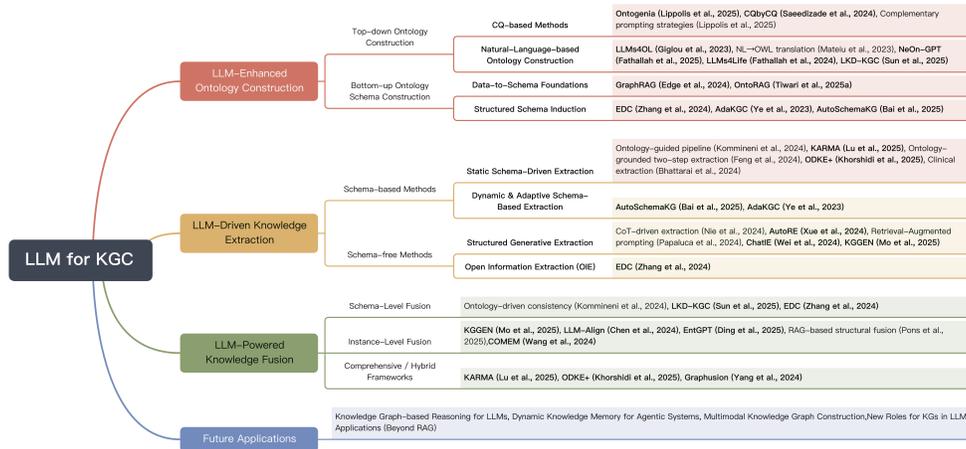


Figure 1: Taxonomy of LLM for KGC

## 3 LLM-ENHANCED ONTOLOGY CONSTRUCTION

The integration of Large Language Models (LLMs) has introduced a fundamental paradigm shift in Ontology Engineering (OE) and, by extension, Knowledge Graph (KG) construction. Current research generally follows two complementary directions: a **top-down** approach, which leverages LLMs as intelligent assistants for formal ontology modeling, and a **bottom-up** approach, which employs ontology construction to enhance the reasoning and representation capabilities of LLMs themselves.

### 3.1 TOP-DOWN ONTOLOGY CONSTRUCTION: LLMs AS ONTOLOGY ASSISTANTS

The top-down paradigm extends the traditions of the Semantic Web and Knowledge Engineering, emphasizing ontology development guided by predefined semantic requirements. Within this framework, LLMs serve as advanced co-modelers that assist human experts in translating natural language specifications—such as competency questions (CQs), user stories, or domain descriptions—into formal ontologies, typically represented in OWL or related standards. This paradigm prioritizes conceptual abstraction, the precise definition of relations, and structured semantic representation to ensure that subsequent knowledge extraction and instance population adhere to well-defined logical constraints.

### 3.1.1 COMPETENCY QUESTION (CQ)-BASED ONTOLOGY CONSTRUCTION

CQ-based methods represent a requirements-driven pathway toward automated ontology modeling. In this setting, LLMs parse CQs or user stories to identify, categorize, and formalize domain-specific concepts, attributes, and relationships.

A pioneering framework, **Ontogenia** (Lippolis et al., 2025a), introduced the use of Metacognitive Prompting for ontology generation, enabling the model to perform self-reflection and structural correction during synthesis. By incorporating Ontology Design Patterns (ODPs), Ontogenia improves both the consistency and complexity of generated ontologies. Similarly, the **CQbyCQ** framework (Saeedizade & Blomqvist, 2024) demonstrated that LLMs can directly translate CQs and user stories into OWL-compliant schemas, effectively automating the transition from requirements to structured ontological models.

Building on these advances, Lippolis et al. (2025b) proposed two complementary prompting strategies: a “memoryless” approach for modular construction and a reflective iterative method inspired by Ontogenia. Empirical evaluations revealed that LLMs can autonomously identify classes, object properties, and data properties, while generating corresponding logical axioms with consistency comparable to that of junior human modelers. Collectively, these studies have led to semi-automated ontology construction pipelines encompassing the entire lifecycle—from CQ formulation and validation to ontology instantiation—with human experts intervening only at critical checkpoints. Through this evolution, LLMs have transitioned from passive analytical tools to active modeling collaborators in ontology design.

### 3.1.2 NATURAL LANGUAGE-BASED ONTOLOGY CONSTRUCTION

Beyond CQ-driven paradigms, natural language-based ontology construction seeks to induce semantic schemas directly from unstructured or semi-structured text corpora, eliminating the dependency on explicitly formulated questions. The goal is to enable LLMs to autonomously uncover conceptual hierarchies and relational patterns from natural language, achieving a direct mapping from textual descriptions to formal logical representations.

Foundational work in this domain—including Saeedizade & Blomqvist (2024) and Lippolis et al. (2025b)—systematically evaluated GPT-4’s performance and confirmed that its outputs approach the quality of novice human modelers, thereby validating the feasibility of “intelligent ontology assistants.” The **LLMs4OL** framework (Giglou et al., 2023) further verified LLMs’ capacity for concept identification, relation extraction, and semantic pattern induction in general-purpose domains. Likewise, Mateiu & Groza (2023) demonstrated the use of fine-tuned models to directly translate natural language into OWL axioms within established ontology editors such as Protégé.

Recent systems such as **NeOn-GPT** (Fathallah et al., 2025) and **LLMs4Life** (Fathallah et al., 2024) have advanced this direction by introducing end-to-end, prompt-driven workflows that integrate ontology reuse and adaptive refinement to construct deep, coherent ontological structures in complex scientific domains (e.g., life sciences). Meanwhile, lightweight frameworks such as **LKD-KGC** (Sun et al., 2025) enable rapid schema induction for open-domain knowledge graphs by clustering entity types extracted from document summaries.

In summary, top-down research on LLM-assisted ontology construction emphasizes semantic consistency, structural completeness, and human–AI collaboration, marking a significant evolution of traditional knowledge engineering toward more intelligent, language-driven paradigms.

## 3.2 BOTTOM-UP ONTOLOGY SCHEMA CONSTRUCTION: KGs FOR LLMs

The bottom-up methodology has gained increasing attention as a response to paradigm shifts introduced by the era of Large Language Models (LLMs), particularly within Retrieval-Augmented Generation (RAG) frameworks. In this paradigm, the knowledge graph is no longer viewed merely as a static repository of structured knowledge for human interpretation. Instead, it serves as a dynamic infrastructure that provides factual grounding and structured memory for LLMs. Consequently, research focus has shifted from manually designing ontological hierarchies to **automatically inducing schemas** from unstructured or semi-structured data. This evolution can be delineated through three interrelated stages of progress.

Early studies such as **GraphRAG** (Edge et al., 2024) and **OntoRAG** (Tiwari et al., 2025) established the foundation for data-driven ontology construction. These approaches first generate instance-level graphs from raw text via open information extraction, and then abstract ontological concepts and relations through clustering and generalization. This “data-to-schema” process transforms empirical knowledge into reusable conceptual structures, illustrating how instance-rich corpora can give rise to ontological blueprints.

Building upon this foundation, the **EDC (Extract–Define–Canonicalize)** framework (Zhang & Soh, 2024) advanced the pipeline into a three-stage process consisting of open extraction, semantic definition, and schema normalization. It enables the alignment of automatically induced schemas with existing ontologies, or the creation of new ones when predefined structures are absent. Extending this adaptability, **AdaKGC** (Ye et al., 2023) addressed the challenge of dynamic schema evolution, allowing models to incorporate novel relations and entity types without retraining. Collectively, these advances shift the focus from static schema construction toward **continuous schema adaptation** within evolving knowledge environments.

More recent efforts have transitioned beyond algorithmic prototypes toward deployable knowledge systems. For example, **AutoSchemaKG** (Bai et al., 2025) integrates schema-based and schema-free paradigms within a unified architecture, supporting the real-time generation and evolution of enterprise-scale knowledge graphs. In this stage, KGs operate as a form of *external knowledge memory* for LLMs—prioritizing factual coverage, scalability, and maintainability over purely semantic completeness. This transformation marks a pragmatic reorientation of ontology construction, emphasizing its service to LLM reasoning and interpretability in knowledge-intensive applications.

In summary, bottom-up ontology schema construction redefines the interplay between LLMs and knowledge engineering. The focus evolves from “**LLMs for Ontology Engineering**” to “**Ontologies and KGs for LLMs**”. Whereas the *top-down* trajectory emphasizes semantic modeling, logical consistency, and expert-guided alignment—positioning LLMs as intelligent assistants in ontology design—the *bottom-up* trajectory prioritizes automatic extraction, schema induction, and dynamic evolution. This progression advances toward self-updating, interpretable, and scalable knowledge ecosystems that strengthen the grounding and reasoning capabilities of LLMs.

## 4 LLM-DRIVEN KNOWLEDGE EXTRACTION

Through a systematic examination of recent advances, it becomes evident that methodologies for Large Language Model (LLM)-driven knowledge extraction have evolved along two principal paradigms: **schema-based** extraction, which operates under explicit structural guidance, and **schema-free** extraction, which transcends the limitations of predefined templates. The former emphasizes normalization, structural consistency, and semantic alignment, while the latter prioritizes adaptability, openness, and exploratory discovery. Together, these paradigms delineate the conceptual landscape of contemporary research in LLM-based knowledge extraction.

### 4.1 SCHEMA-BASED METHODS

The central principle of schema-based extraction lies in its reliance on an explicit knowledge schema that provides both structural guidance and semantic constraints for the extraction process. Within this paradigm, the research trajectory demonstrates a clear evolution—from the use of static ontological blueprints toward adaptive and dynamically evolving schema frameworks.

#### 4.1.1 STATIC SCHEMA-DRIVEN EXTRACTION

Early studies in LLM-driven knowledge extraction predominantly employed **predefined, static schemas** that rigidly constrained the extraction process. In this paradigm, the ontology functions as a fixed semantic backbone, directing the LLM to populate the knowledge base under strict structural supervision. The progression of this research line can be broadly characterized by three developmental stages.

Initial efforts, such as Kommineni et al. (2024), utilized fully predefined ontological structures to ensure precision and interpretability. In their pipeline, the LLM first generates Competency Questions (CQs) to delineate the knowledge scope, constructs the corresponding ontology (TBox), and

subsequently performs ABox population under explicit schema supervision—achieving high consistency but limited flexibility. Similarly, the **KARMA** framework (Lu & Wang, 2025) adopts a multi-agent architecture, in which each agent executes schema-guided extraction tasks to guarantee accurate entity normalization and relation classification within a fixed ontological boundary.

Building on these rigid frameworks, subsequent studies sought to enhance modularity and reusability through staged prompting. For instance, Feng et al. (2024) proposed a two-step “ontology-grounded extraction” approach: first generating a domain-specific ontology directly from text, and then leveraging it as a directive prompt for RDF triple extraction. This approach strengthens structural alignment while maintaining partial adaptability.

More recent developments introduce localized flexibility within otherwise static frameworks. **ODKE+** (Khorshidi et al., 2025) proposes *ontology snippets*—dynamically selected ontology subsets—to construct context-aware prompts tailored to specific entities, thus enabling limited schema adaptation at runtime. Likewise, Bhattarai et al. (2024) utilize the medical ontology UMLS to dynamically generate task-specific prompts for clinical information extraction. Although these methods introduce local adaptability, they remain bounded by pre-existing macro-schemas, representing a transitional phase toward schema dynamism.

In summary, static schema-driven extraction forms the foundational paradigm of LLM-assisted knowledge extraction, emphasizing **precision, logical consistency, and interpretability**. However, its dependence on rigid ontological templates restricts scalability and cross-domain generalization. The progression from fixed schema control to selective, context-aware schema prompting marks the field’s gradual shift toward more adaptive, data-responsive frameworks.

#### 4.1.2 DYNAMIC AND ADAPTIVE SCHEMA-BASED EXTRACTION

Recent approaches reconceptualize the schema as a **dynamic, evolving component** of the extraction process rather than a fixed template—a shift from schema *guiding* extraction to schema *co-evolving* with it.

**AutoSchemaKG** (Bai et al., 2025) exemplifies this trend by inducing schemas from large-scale corpora via unsupervised clustering and relation discovery. It employs multi-stage prompts tailored to different relation types, enabling the schema to evolve iteratively with extracted content and improving open-domain scalability. Building on this idea, **AdaKGC** (Ye et al., 2023) tackles *schema drift* through two mechanisms: the **Schema-Enriched Prefix Instruction (SPI)** for context-aware prompting and the **Schema-Constrained Dynamic Decoding (SDD)** for schema adaptation without retraining.

Together, these methods enable **adaptive schema learning**, bridging symbolic structure and data-driven flexibility. They lay the groundwork for continual, self-updating knowledge graph construction where extraction and schema evolution progress synergistically.

## 4.2 SCHEMA-FREE METHODS

In contrast to paradigms that depend on externally defined blueprints, **schema-free extraction** aims to acquire structured knowledge directly from unstructured text without relying on any predefined ontology or relation schema. The central idea is to leverage Large Language Models (LLMs) as autonomous extractors capable of identifying entities and relations through advanced prompt engineering, instruction tuning, and self-organizing reasoning. The evolution of this paradigm unfolds along two major trajectories: *structured generative extraction* and *open information extraction*.

### 4.2.1 STRUCTURED GENERATIVE EXTRACTION

The first trajectory, **structured generative extraction**, focuses on prompting LLMs to construct an *implicit* or *on-the-fly* schema during generation. Although no external ontology is provided, structured reasoning patterns and generative templates guide the model toward consistent and coherent knowledge generation.

Early studies, such as Nie et al. (2024), integrated the extraction process with **Chain-of-Thought (CoT)** prompting, encouraging stepwise reasoning for entity and relation identification. This approach demonstrated that explicit schemas can be effectively replaced by reasoning-driven organiza-

324 tion. Building on this insight, **AutoRE** (Xue et al., 2024) introduced an RHF (Relation–Head–Facts)  
 325 pipeline via instruction fine-tuning, enabling the model to internalize relational regularities and im-  
 326 prove coherence and scalability across documents.

327 Subsequent works further enhanced schema-free extraction by incorporating retrieval and interac-  
 328 tivity. For instance, Papaluca et al. (2024) proposed a **Retrieval-Augmented** prompting framework  
 329 that dynamically enriches the context window with semantically related exemplars, thereby improv-  
 330 ing factual precision. In parallel, **ChatIE** (Wei et al., 2024) reformulated extraction as a **multi-**  
 331 **turn dialogue** process, wherein the model iteratively refines entity and relation candidates through  
 332 chained question answering. Similarly, **KGGEN** (Mo et al., 2025) decomposed extraction into two  
 333 sequential LLM invocations—first detecting entities, then generating relations—to reduce cognitive  
 334 load and mitigate error propagation.

335 Collectively, these studies reveal that even without explicit schemas, LLMs can internalize latent re-  
 336 lational structures through guided reasoning, modular prompting, and interactive refinement—laying  
 337 the groundwork for open-ended and self-organizing knowledge generation.

#### 339 4.2.2 OPEN INFORMATION EXTRACTION (OIE)

341 **Open Information Extraction (OIE)** extends structured generative methods toward schema-free  
 342 extraction, aiming to discover all possible entity–relation–object triples from text without relying on  
 343 predefined types. The **EDC** framework (Zhang & Soh, 2024) exemplifies this paradigm: its *Extract*  
 344 stage uses few-shot prompting to generate comprehensive natural-language triples, producing a raw  
 345 open knowledge graph that later undergoes definition and canonicalization. OIE prioritizes coverage  
 346 and discovery over structural regularity. When combined with schema induction or canonicalization,  
 347 it bridges unstructured text and emergent ontological organization—completing the continuum from  
 348 schema-free to schema-generative knowledge construction.

## 349 5 LLM-POWERED KNOWLEDGE FUSION

351 An examination of recent advances and pioneering studies indicates that methodologies leveraging  
 352 Large Language Models (LLMs) for knowledge fusion predominantly address challenges at two  
 353 fundamental levels: (1) constructing a unified and normalized knowledge *skeleton* at the **schema**  
 354 **layer**, and (2) integrating and aligning the specific knowledge *flesh* at the **instance layer**. According  
 355 to this distinction, existing approaches can be categorized into three major classes: **schema-level**  
 356 **fusion**, **instance-level fusion**, and **hybrid frameworks** that integrate both.

### 358 5.1 SCHEMA-LEVEL FUSION

360 Schema-level fusion unifies the **structural backbone** of knowledge graphs—concepts, entity types,  
 361 relations, and attributes—into a coherent and semantically consistent schema. By aligning heteroge-  
 362 neous elements, it ensures that all knowledge adheres to a unified conceptual specification. Research  
 363 in this area has evolved through three major phases.

364 **(1) Ontology-driven consistency.** Early work relied on explicit ontologies as global constraints.  
 365 For instance, Kommineni et al. (2024) enforced alignment between extracted triples and predefined  
 366 ontological definitions, achieving high semantic consistency but limited flexibility across domains.

367 **(2) Data-driven unification.** To overcome this rigidity, **LKD-KGC** (Sun et al., 2025) introduced  
 368 adaptive, embedding-based schema integration. It automatically extracts and merges equivalent en-  
 369 tity types via *vector clustering* and *LLM-based deduplication*, allowing schema alignment to emerge  
 370 dynamically from data.

371 **(3) LLM-enabled canonicalization.** Recent approaches such as **EDC** (Zhang & Soh, 2024) extend  
 372 fusion toward semantic canonicalization. By prompting LLMs to generate natural language defini-  
 373 tions of schema components and comparing them via vector similarity, this method supports both  
 374 self-alignment and cross-schema mapping with greater automation and semantic precision.

375 In summary, schema-level fusion has progressed from **ontology-driven** to **data-driven** to **LLM-**  
 376 **enabled** paradigms—marking a shift from rigid rule-based alignment toward flexible, semantics-  
 377 oriented fusion mediated by LLM reasoning.

378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431

## 5.2 INSTANCE-LEVEL FUSION

Instance-level fusion integrates concrete knowledge instances by addressing **entity alignment, disambiguation, deduplication, and conflict resolution**. Its goal is to reconcile heterogeneous or redundant entities to maintain a coherent and semantically precise knowledge graph. Recent work reflects a clear evolution—from heuristic clustering to structured, reasoning-based frameworks.

Early studies such as *KGGEN* (Mo et al., 2025) employed iterative LLM-guided clustering to merge equivalent entities and relations beyond surface matching. The framework performs progressive triple extraction followed by semantic grouping, revealing the potential of LLMs to aggregate semantically related entities through implicit reasoning rather than explicit rules. Later, *LLM-Align* (Chen et al., 2024) and *EntGPT* (Ding et al., 2025) reframed alignment as a *contextual reasoning* task, using multi-step prompting to enhance semantic discrimination. *LLM-Align* treats alignment as a constrained multiple-choice problem, while *EntGPT* introduces a two-phase refinement pipeline that first generates candidate entities and then applies targeted reasoning for final selection, significantly improving alignment precision. More recent efforts incorporate structural and retrieval cues—e.g., Pons et al. (2025) leverage **RAG**-based fusion to exploit class–subclass hierarchies and entity descriptions for zero-shot disambiguation. This integration of graph topology enables more robust reasoning about unseen or ambiguous entities. Efficiency has also improved through hierarchical designs such as **COMEM** (Wang et al., 2024), which combines lightweight filtering with fine-grained reasoning. By cascading smaller and larger LLMs in a multi-stage pipeline, it achieves substantial efficiency gains while maintaining high semantic accuracy in large-scale fusion tasks.

Overall, LLMs have evolved from simple matchers to **adaptive reasoning agents** that integrate contextual, structural, and retrieved signals for scalable, self-correcting fusion—paving the way toward autonomous knowledge graph construction.

## 5.3 COMPREHENSIVE AND HYBRID FRAMEWORKS

Comprehensive and hybrid frameworks unify **schema-level** and **instance-level** fusion within a single, end-to-end workflow, moving beyond traditional modular pipelines toward integrated, prompt-driven architectures.

The *KARMA* framework (Lu & Wang, 2025) exemplifies a multi-agent design where specialized agents collaboratively handle schema alignment, conflict resolution, and quality evaluation, achieving scalability and global consistency. Building on this, *ODKE+* (Khorshidi et al., 2025) employs an ontology-guided workflow coupling schema supervision with instance-level corroboration for improved semantic fidelity. More recently, *Graphusion* (Yang et al., 2024) introduces a unified, prompt-based paradigm that performs all fusion subtasks—alignment, consolidation, and inference—within a single generative cycle.

Together, these frameworks signal a shift toward **integrated, adaptive, and generative** fusion systems, marking a crucial step toward autonomous, self-evolving knowledge graphs capable of continuous construction and reasoning in LLM-driven ecosystems.

## 6 FUTURE APPLICATIONS

Research at the intersection of Large Language Models (LLMs) and Knowledge Graphs (KGs) is progressively advancing toward deeper intelligent interaction and greater autonomy in knowledge representation and reasoning. In light of these developments, several promising directions are emerging for future exploration.

### 6.1 KNOWLEDGE GRAPH-BASED REASONING FOR LLMs

Future work is expected to further integrate structured KGs into the reasoning mechanisms of LLMs, enhancing their logical consistency, causal inference, and interpretability. This research direction signifies not only an improvement in reasoning capabilities but also a conceptual transition from *knowledge construction* to *knowledge-driven reasoning*. High-quality, well-structured KGs will provide the foundation for explainable and verifiable model inference. Existing studies such as *KG-*

432 *based Random-Walk Reasoning* (Kim et al., 2024) and *KG-RAR* (Wu et al., 2025) have demonstrated  
433 the potential of this paradigm. A crucial complementary challenge, however, lies in how enhanced  
434 reasoning abilities can in turn support more robust and automated KG construction—forming a self-  
435 improving, virtuous cycle between knowledge building and reasoning.

## 437 6.2 DYNAMIC KNOWLEDGE MEMORY FOR AGENTIC SYSTEMS

438  
439 Achieving **autonomy** in LLM-powered agents requires overcoming the limits of finite context win-  
440 dows through **persistent, structured memory**. Recent architectures envision the **knowledge graph**  
441 (**KG**) as a *dynamic memory substrate*, evolving continuously with agent interactions rather than  
442 storing static histories. Frameworks such as **A-MEM** (Xu et al., 2025) model memory as intercon-  
443 nected “notes” enriched with contextual metadata, enabling continual reorganization and growth.  
444 Similarly, **Zep** (Rasmussen et al., 2025) employs a **temporal knowledge graph (TKG)** to manage  
445 fact validity and support time-aware reasoning and updates. These advances highlight dynamic KGs  
446 as long-term, interpretable memory systems that enable **continuous learning, multi-agent collab-**  
447 **oration, and self-reflective reasoning**. Future work will focus on improving scalability, temporal  
448 coherence, and multimodal integration for fully autonomous, knowledge-grounded agents.

## 449 6.3 MULTIMODAL KNOWLEDGE GRAPH CONSTRUCTION

450  
451 Multimodal Knowledge Graph (MMKG) construction aims to integrate heterogeneous modalities—such as text, images, audio, and video—into unified, structured representations that enable  
452 richer reasoning and cross-modal alignment. Representative work includes **VaLiK** (Vision-align-  
453 to-Language integrated KG) (Liu et al., 2025), which cascades pretrained Vision-Language Models  
454 (VLMs) to translate visual features into textual form, followed by a cross-modal verification module  
455 to filter noise and assemble MMKGs. This process achieves entity–image linkage without manual  
456 annotation. Beyond structure, representation learning methods such as **KG-MRI** (Lu et al., 2024)  
457 employ multimodal embeddings with contrastive objectives to align heterogeneous modalities into  
458 coherent semantic spaces. Key challenges remain in modality heterogeneity, alignment noise, scal-  
459 ability, and robustness under missing or imbalanced modalities. As LLMs and VLMs continue to  
460 co-evolve, MMKGs will become a cornerstone for bridging perceptual input and symbolic reasoning  
461 across modalities.

## 463 6.4 NEW ROLES FOR KGs IN LLM APPLICATIONS: BEYOND RAG

464  
465 Beyond their use as retrieval backbones in RAG systems, Knowledge Graphs (KGs) are increasingly  
466 envisioned as a **cognitive middle layer** bridging raw input and LLM reasoning. In this paradigm,  
467 KGs provide a structured scaffold for **querying, planning, and decision-making**, enabling more  
468 interpretable and grounded generation. Recent studies illustrate this shift. **CogER** (Bing et al.,  
469 2023) formulates recommendation as cognition-aware KG reasoning, integrating intuitive and path-  
470 based inference for explainability. In the biomedical domain, **PKG-LLM** (Sarabadani et al., 2025)  
471 employs domain KGs for knowledge augmentation and predictive modeling in mental health diag-  
472 nosis. Together, these approaches treat the KG as an **interactive reasoning substrate**, promising  
473 more robust and explainable generation in domains such as science, code, and healthcare.

## 474 7 CONCLUSION

475  
476 This survey presents a comprehensive overview of how Large Language Models (LLMs) are trans-  
477 forming Knowledge Graph (KG) construction across ontology engineering, knowledge extraction,  
478 and knowledge fusion. LLMs shift the paradigm from rule-based and modular pipelines toward  
479 unified, adaptive, and generative frameworks. Across these stages, three trends emerge: (1) the  
480 evolution from *static schemas* to *dynamic induction*, (2) the integration of *pipeline modularity* into  
481 *generative unification*, and (3) the transition from *symbolic rigidity* to *semantic adaptability*. To-  
482 gether, these shifts redefine KGs as **living, cognitive infrastructures** that blend language under-  
483 standing with structured reasoning. Despite remarkable progress, challenges remain in scalability,  
484 reliability, and continual adaptation. Future advances in *prompt design, multimodal integration, and*  
485 *knowledge-grounded reasoning* will be key to realizing **autonomous and explainable knowledge-**  
**centric AI systems**.

## REFERENCES

- 486  
487  
488 Muhammad Nabeel Asim, Muhammad Wasim, Muhammad Usman Ghani Khan, Waqar Mah-  
489 mood, and Hafiza Mahnoor Abbasi. A survey of ontology learning techniques and applications.  
490 *Database*, 2018:bay101, 2018.
- 491  
492 Jiaxin Bai, Wei Fan, Qi Hu, Qing Zong, Chunyang Li, Hong Ting Tsang, Hongyu Luo, Yauwai  
493 Yim, Haoyu Huang, Xiao Zhou, Feng Qin, Tianshi Zheng, Xi Peng, Xin Yao, Huiwen Yang,  
494 Leijie Wu, Yi Ji, Gong Zhang, Renhai Chen, and Yangqiu Song. AutoSchemaKG: Autonomous  
495 Knowledge Graph Construction through Dynamic Schema Induction from Web-Scale Corpora.  
496 *arXiv preprint arXiv:2505.23628*, aug 2025. doi: 10.48550/arXiv.2505.23628. URL <http://arxiv.org/abs/2505.23628>.
- 497  
498 Kriti Bhattarai, Inez Y. Oh, Zachary B. Abrams, and Albert M. Lai. Document-level Clinical Entity  
499 and Relation extraction via Knowledge Base-Guided Generation. In *Proceedings of the 23rd*  
500 *Workshop on Biomedical Natural Language Processing*, pp. 318–327, Bangkok, Thailand, 2024.  
501 Association for Computational Linguistics. doi: 10.18653/v1/2024.bionlp-1.24. URL <https://aclanthology.org/2024.bionlp-1.24>.
- 502  
503 Qingyu Bing, Qiannan Zhu, and Zhicheng Dou. Cognition-aware Knowledge Graph Reasoning for  
504 Explainable Recommendation. In *Proceedings of the Sixteenth ACM International Conference on*  
505 *Web Search and Data Mining*, pp. 402–410, Singapore, Singapore, feb 2023. ACM. doi: 10.1145/  
506 3539597.3570391. URL <https://dl.acm.org/doi/10.1145/3539597.3570391>.
- 507  
508 Xuan Chen, Tong Lu, and Zhichun Wang. LLM-Align: Utilizing Large Language Models for Entity  
509 Alignment in Knowledge Graphs. *arXiv preprint arXiv:2412.04690*, dec 2024. doi: 10.48550/  
510 arXiv.2412.04690. URL <http://arxiv.org/abs/2412.04690>.
- 511  
512 Kartik Detroja, CK Bhensdadia, and Brijesh S Bhatt. A survey on relation extraction. *Intelligent*  
513 *Systems with Applications*, 19:200244, 2023.
- 514  
515 Yifan Ding, Amrit Poudel, Qingkai Zeng, Tim Weninger, Balaji Veeramani, and Sanmitra Bhat-  
516 tacharya. EntGPT: Entity Linking with Generative Large Language Models. *arXiv preprint*  
517 *arXiv:2402.06738*, may 2025. doi: 10.48550/arXiv.2402.06738. URL <http://arxiv.org/abs/2402.06738>.
- 518  
519 Darren Edge, Ha Trinh, Newman Cheng, Joshua Bradley, Alex Chao, Apurva Mody, Steven Truitt,  
520 and Jonathan Larson. From Local to Global: A Graph RAG Approach to Query-Focused Sum-  
521 marization. *arXiv preprint arXiv:2404.16130*, apr 2024. doi: 10.48550/arXiv.2404.16130. URL  
522 <http://arxiv.org/abs/2404.16130>.
- 523  
524 Nadeen Fathallah, Steffen Staab, and Alsayed Algergawy. LLMs4Life: Large Language Models for  
525 Ontology Learning in Life Sciences. *arXiv preprint arXiv:2412.02035*, dec 2024. doi: 10.48550/  
526 arXiv.2412.02035. URL <http://arxiv.org/abs/2412.02035>.
- 527  
528 Nadeen Fathallah, Arunav Das, Stefano De Giorgis, Andrea Poltronieri, Peter Haase, and Liubov  
529 Kovriguina. NeOn-GPT: A Large Language Model-Powered Pipeline for Ontology Learning.  
530 In Albert Meroño Peñuela, Oscar Corcho, Paul Groth, Elena Simperl, Valentina Tamma, An-  
531 drea Giovanni Nuzzolese, Maria Poveda-Villalón, Marta Sabou, Valentina Presutti, Irene Celino,  
532 Artem Revenko, Joe Raad, Bruno Sartini, and Pasquale Lisena (eds.), *The Semantic Web: ESWC*  
533 *2024 Satellite Events*, volume 15344, pp. 36–50. Springer Nature Switzerland, Cham, 2025.  
534 doi: 10.1007/978-3-031-78952-6\_4. URL [https://link.springer.com/10.1007/978-3-031-78952-6\\_4](https://link.springer.com/10.1007/978-3-031-78952-6_4).
- 535  
536 Xiaohan Feng, Xixin Wu, and Helen Meng. Ontology-grounded Automatic Knowledge Graph Con-  
537 struction by LLM under Wikidata schema. *arXiv preprint arXiv:2412.20942*, dec 2024. doi:  
538 10.48550/arXiv.2412.20942. URL <http://arxiv.org/abs/2412.20942>.
- 539  
540 Hamed Babaei Giglou, Jennifer D’Souza, and Sören Auer. LLMs4OL: Large Language Models  
541 for Ontology Learning. *arXiv preprint arXiv:2307.16648*, aug 2023. doi: 10.48550/arXiv.2307.  
542 16648. URL <http://arxiv.org/abs/2307.16648>.

- 540 Samira Khorshidi, Azadeh Nikfarjam, Suprita Shankar, Yisi Sang, Yash Govind, Hyun Jang, Ali  
541 Kaskari, Alexis McClimans, Mohamed Soliman, Vishnu Konda, Ahmed Fakhry, and Xiaoguang  
542 Qi. ODKE+: Ontology-Guided Open-Domain Knowledge Extraction with LLMs. *arXiv preprint*  
543 *arXiv:2509.04696*, sep 2025. doi: 10.48550/arXiv.2509.04696. URL [http://arxiv.org/  
544 abs/2509.04696](http://arxiv.org/abs/2509.04696).
- 545 Yejin Kim, Eojin Kang, Juae Kim, and H. Howie Huang. Causal Reasoning in Large Language  
546 Models: A Knowledge Graph Approach. *arXiv preprint arXiv:2410.11588*, oct 2024. doi: 10.  
547 48550/arXiv.2410.11588. URL <http://arxiv.org/abs/2410.11588>.
- 548 Vamsi Krishna Kommineni, Birgitta König-Ries, and Sheeba Samuel. Towards the automation of  
549 knowledge graph construction using large language models. *Journal Name*, 2024.
- 550 Konstantinos I. Kotis, George A. Vouros, and Dimitris Spiliotopoulos. Ontology engineer-  
551 ing methodologies for the evolution of living and reused ontologies: status, trends, findings  
552 and recommendations. *The Knowledge Engineering Review*, 35:e4, 2020. doi: 10.1017/  
553 S0269888920000065.
- 554 Anna Sofia Lippolis, Miguel Ceriani, Sara Zuppiroli, and Andrea Giovanni Nuzzolese. Ontoge-  
555 nia: Ontology Generation with Metacognitive Prompting in Large Language Models. In Al-  
556 bert Meroño Peñuela, Oscar Corcho, Paul Groth, Elena Simperl, Valentina Tamma, Andrea Gio-  
557 vanni Nuzzolese, Maria Poveda-Villalón, Marta Sabou, Valentina Presutti, Irene Celino, Artem  
558 Revenko, Joe Raad, Bruno Sartini, and Pasquale Lisena (eds.), *The Semantic Web: ESWC*  
559 *2024 Satellite Events*, volume 15344, pp. 259–265. Springer Nature Switzerland, Cham, 2025a.  
560 doi: 10.1007/978-3-031-78952-6\_38. URL [https://link.springer.com/10.1007/  
561 978-3-031-78952-6\\_38](https://link.springer.com/10.1007/978-3-031-78952-6_38).
- 562 Anna Sofia Lippolis, Mohammad Javad Saeezade, Robin Keskisärkkä, Sara Zuppiroli, Miguel  
563 Ceriani, Aldo Gangemi, Eva Blomqvist, and Andrea Giovanni Nuzzolese. Ontology Generation  
564 using Large Language Models. *arXiv preprint arXiv:2503.05388*, mar 2025b. doi: 10.48550/  
565 arXiv.2503.05388. URL <http://arxiv.org/abs/2503.05388>.
- 566 Junming Liu, Siyuan Meng, Yanting Gao, Song Mao, Pinlong Cai, Guohang Yan, Yirong Chen,  
567 Zilin Bian, Ding Wang, and Botian Shi. Aligning Vision to Language: Annotation-Free  
568 Multimodal Knowledge Graph Construction for Enhanced LLMs Reasoning. *arXiv preprint*  
569 *arXiv:2503.12972*, jul 2025. doi: 10.48550/arXiv.2503.12972. URL [http://arxiv.org/  
570 abs/2503.12972](http://arxiv.org/abs/2503.12972).
- 571 Shuang Liu, Man Xu, Yufeng Qin, and Niko Lukač. Knowledge graph alignment network with  
572 node-level strong fusion. *Applied Sciences*, 12(19), 2022. doi: 10.3390/app12199434. URL  
573 <https://www.mdpi.com/2076-3417/12/19/9434>.
- 574 Yuxing Lu and Jinzhao Wang. KARMA: Leveraging Multi-Agent LLMs for Automated Knowledge  
575 Graph Enrichment. *arXiv preprint arXiv:2502.06472*, feb 2025. doi: 10.48550/arXiv.2502.06472.  
576 URL <http://arxiv.org/abs/2502.06472>.
- 577 Yuxing Lu, Weichen Zhao, Nan Sun, and Jinzhao Wang. Enhancing multimodal knowledge graph  
578 representation learning through triple contrastive learning. In Kate Larson (ed.), *Proceedings of*  
579 *the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pp. 5963–  
580 5971. International Joint Conferences on Artificial Intelligence Organization, aug 2024. doi:  
581 10.24963/ijcai.2024/659. URL <https://doi.org/10.24963/ijcai.2024/659>. Main  
582 Track.
- 583 Patricia Mateiu and Adrian Groza. Ontology engineering with Large Language Models. *arXiv*  
584 *preprint arXiv:2307.16699*, jul 2023. doi: 10.48550/arXiv.2307.16699. URL [http://arxiv.  
585 org/abs/2307.16699](http://arxiv.org/abs/2307.16699).
- 586 Belinda Mo, Kyssen Yu, Joshua Kazdan, Proud Mpala, Lisa Yu, Chris Cundy, Charilaos Kanat-  
587 soulis, and Sanmi Koyejo. KGen: Extracting Knowledge Graphs from Plain Text with Lan-  
588 guage Models. *arXiv preprint arXiv:2502.09956*, feb 2025. doi: 10.48550/arXiv.2502.09956.  
589 URL <http://arxiv.org/abs/2502.09956>.

- 594 Jixuan Nie, Xia Hou, Wenfeng Song, Xuan Wang, Xinyu Zhang, Xingliang Jin, Shuoze Zhang,  
595 and Jiaqi Shi. Knowledge graph efficient construction: Embedding chain-of-thought into llms.  
596 *Proceedings of the VLDB Endowment*. ISSN, 2150:8097, 2024.  
597
- 598 Liu Pai, Wenyang Gao, Wenjie Dong, Lin Ai, Ziwei Gong, Songfang Huang, Li Zongsheng,  
599 Ehsan Hoque, Julia Hirschberg, and Yue Zhang. A survey on open information extraction  
600 from rule-based model to large language model. In Yaser Al-Onaizan, Mohit Bansal, and  
601 Yun-Nung Chen (eds.), *Findings of the Association for Computational Linguistics: EMNLP*  
602 *2024*, pp. 9586–9608, Miami, Florida, USA, 2024. Association for Computational Linguistics.  
603 doi: 10.18653/v1/2024.findings-emnlp.560. URL [https://aclanthology.org/2024.  
604 findings-emnlp.560/](https://aclanthology.org/2024.findings-emnlp.560/).
- 605 Shirui Pan, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, and Xindong Wu. Unifying large  
606 language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data*  
607 *Engineering*, 36(7):3580–3599, 2024.
- 608 Andrea Papaluca, Daniel Krefl, Sergio Rodríguez Méndez, Artem Lensky, and Hanna Suominen.  
609 Zero- and Few-Shots Knowledge Graph Triplet Extraction with Large Language Models. In  
610 *Proceedings of the 1st Workshop on Knowledge Graphs and Large Language Models (KaLLM*  
611 *2024)*, pp. 12–23, Bangkok, Thailand, 2024. Association for Computational Linguistics. doi:  
612 10.18653/v1/2024.kallm-1.2. URL <https://aclanthology.org/2024.kallm-1.2>.
- 613 Gerard Pons, Besim Bilalli, and Anna Queralt. Knowledge Graphs for Enhancing Large Language  
614 Models in Entity Disambiguation. *arXiv preprint arXiv:2505.02737*, pp. 162–179, 2025. doi:  
615 10.1007/978-3-031-77844-5\_9. URL <http://arxiv.org/abs/2505.02737>.
- 616 Preston Rasmussen, Pavlo Paliychuk, Travis Beauvais, Jack Ryan, and Daniel Chalef. Zep: A  
617 Temporal Knowledge Graph Architecture for Agent Memory. *arXiv preprint arXiv:2501.13956*,  
618 jan 2025. doi: 10.48550/arXiv.2501.13956. URL <http://arxiv.org/abs/2501.13956>.
- 619 Mohammad Javad Saeezade and Eva Blomqvist. Navigating Ontology Development with  
620 Large Language Models. In Albert Meroño Peñuela, Anastasia Dimou, Raphaël Troncy,  
621 Olaf Hartig, Maribel Acosta, Mehwish Alam, Heiko Paulheim, and Pasquale Lisena (eds.),  
622 *The Semantic Web*, volume 14664, pp. 143–161. Springer Nature Switzerland, Cham, 2024.  
623 doi: 10.1007/978-3-031-60626-7\_8. URL [https://link.springer.com/10.1007/  
625 978-3-031-60626-7\\_8](https://link.springer.com/10.1007/<br/>624 978-3-031-60626-7_8).
- 626 Ali Sarabadani, Hadis Taherinia, Niloufar Ghadiri, Ehsan Karimi Shahmarvandi, and Ramin  
627 Mousa. PKG-LLM: A Framework for Predicting GAD and MDD Using Knowledge Graphs  
628 and Large Language Models in Cognitive Neuroscience. *Preprints*, feb 2025. doi: 10.  
629 20944/preprints202502.0982.v1. URL [https://www.preprints.org/manuscript/  
631 202502.0982/v1](https://www.preprints.org/manuscript/<br/>630 202502.0982/v1).
- 632 Jiaqi Sun, Shiyong Qian, Zhangchi Han, Wei Li, Zelin Qian, Dingyu Yang, Jian Cao, and Guangtao  
633 Xue. LKD-KGC: Domain-Specific KG Construction via LLM-driven Knowledge Dependency  
634 Parsing. *arXiv preprint arXiv:2505.24163*, may 2025. doi: 10.48550/arXiv.2505.24163. URL  
635 <http://arxiv.org/abs/2505.24163>.
- 636 Yash Tiwari, Owais Ahmad Lone, and Mayukha Pal. OntoRAG: Enhancing Question-Answering  
637 through Automated Ontology Derivation from Unstructured Knowledge Bases. *arXiv preprint*  
638 *arXiv:2506.00664*, may 2025. doi: 10.48550/arXiv.2506.00664. URL [http://arxiv.org/  
640 abs/2506.00664](http://arxiv.org/<br/>639 abs/2506.00664).
- 641 Tianshu Wang, Xiaoyang Chen, Hongyu Lin, Xuanang Chen, Xianpei Han, Hao Wang, Zhenyu  
642 Zeng, and Le Sun. Match, compare, or select? an investigation of large language models for  
643 entity matching. *arXiv preprint arXiv:2405.16884*, 2024.  
644
- 645 Xiang Wei, Xingyu Cui, Ning Cheng, Xiaobin Wang, Xin Zhang, Shen Huang, Pengjun Xie, Jinan  
646 Xu, Yufeng Chen, Meishan Zhang, Yong Jiang, and Wenjuan Han. ChatIE: Zero-Shot Information  
647 Extraction via Chatting with ChatGPT. *arXiv preprint arXiv:2302.10205*, may 2024. doi: 10.  
48550/arXiv.2302.10205. URL <http://arxiv.org/abs/2302.10205>.

- 648 Wenjie Wu, Yongcheng Jing, Yingjie Wang, Wenbin Hu, and Dacheng Tao. Graph-Augmented Reasoning: Evolving Step-by-Step Knowledge Graph Retrieval for LLM Reasoning. *arXiv preprint arXiv:2503.01642*, mar 2025. doi: 10.48550/arXiv.2503.01642. URL <http://arxiv.org/abs/2503.01642>.
- 649  
650  
651  
652
- 653 Wujiang Xu, Zujie Liang, Kai Mei, Hang Gao, Juntao Tan, and Yongfeng Zhang. A-MEM: Agentic Memory for LLM Agents. *arXiv preprint arXiv:2502.12110*, oct 2025. doi: 10.48550/arXiv.2502.12110. URL <http://arxiv.org/abs/2502.12110>.
- 654  
655
- 656 Lilong Xue, Dan Zhang, Yuxiao Dong, and Jie Tang. AutoRE: Document-Level Relation Extraction with Large Language Models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pp. 211–220, Bangkok, Thailand, 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.acl-demos.20. URL <https://aclanthology.org/2024.acl-demos.20>.
- 657  
658  
659  
660
- 661 Linyao Yang, Chen Lv, Xiao Wang, Ji Qiao, Weiping Ding, Jun Zhang, and Fei-Yue Wang. Collective entity alignment for knowledge fusion of power grid dispatching knowledge graphs. *IEEE/CAA Journal of Automatica Sinica*, 9(11):1990–2004, 2022a. doi: 10.1109/JAS.2022.105947.
- 662  
663  
664  
665
- 666 Rui Yang, Boming Yang, Sixun Ouyang, Tianwei She, Aosong Feng, Yuang Jiang, Freddy Lecue, Jinghui Lu, and Irene Li. Graphusion: Leveraging Large Language Models for Scientific Knowledge Graph Fusion and Construction in NLP Education. *arXiv preprint arXiv:2407.10794*, jul 2024. doi: 10.48550/arXiv.2407.10794. URL <http://arxiv.org/abs/2407.10794>.
- 667  
668  
669
- 670 Yang Yang, Zhilei Wu, Yuexiang Yang, Shuangshuang Lian, Fengjie Guo, and Zhiwei Wang. A survey of information extraction based on deep learning. *Applied Sciences*, 12(19):9691, 2022b.
- 671  
672
- 673 Hongbin Ye, Honghao Gui, Xin Xu, Xi Chen, Huajun Chen, and Ningyu Zhang. Schema-adaptable Knowledge Graph Construction. *arXiv preprint arXiv:2305.08703*, nov 2023. doi: 10.48550/arXiv.2305.08703. URL <http://arxiv.org/abs/2305.08703>.
- 674  
675
- 676 Kaisheng Zeng, Chengjiang Li, Lei Hou, Juanzi Li, and Ling Feng. A comprehensive survey of entity alignment for knowledge graphs. *AI Open*, 2:1–13, 2021. doi: <https://doi.org/10.1016/j.aiopen.2021.02.002>. URL <https://www.sciencedirect.com/science/article/pii/S2666651021000036>.
- 677  
678  
679
- 680 Bowen Zhang and Harold Soh. Extract, Define, Canonicalize: An LLM-based Framework for Knowledge Graph Construction. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 9820–9836, Miami, Florida, USA, 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.emnlp-main.548. URL <https://aclanthology.org/2024.emnlp-main.548>.
- 681  
682  
683  
684  
685
- 686 Zhigang Zhao, Xiong Luo, Maojian Chen, and Ling Ma. A survey of knowledge graph construction using machine learning. *CMES-Computer Modeling in Engineering & Sciences*, 139(1), 2024.
- 687
- 688 Lingfeng Zhong, Jia Wu, Qian Li, Hao Peng, and Xindong Wu. A comprehensive survey on automatic knowledge graph construction. *ACM Computing Surveys*, 56(4):1–62, 2023.
- 689  
690
- 691 Beibei Zhu, Ruolin Wang, Junyi Wang, Fei Shao, and Kerun Wang. A survey: knowledge graph entity alignment research based on graph embedding. *Artificial Intelligence Review*, 57(9):229, 2024a.
- 692  
693
- 694 Yuqi Zhu, Xiaohan Wang, Jing Chen, Shuofei Qiao, Yixin Ou, Yunzhi Yao, Shumin Deng, Huajun Chen, and Ningyu Zhang. Llms for knowledge graph construction and reasoning: Recent capabilities and future opportunities. *World Wide Web*, 27(5):58, 2024b.
- 695  
696
- 697 Amal Zouaq and Roger Nkambou. A Survey of Domain Ontology Engineering: Methods and Tools. In Roger Nkambou, Jacqueline Bourdeau, and Riichiro Mizoguchi (eds.), *Advances in Intelligent Tutoring Systems*, volume 308, pp. 103–119. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. doi: 10.1007/978-3-642-14363-2\_6. URL [http://link.springer.com/10.1007/978-3-642-14363-2\\_6](http://link.springer.com/10.1007/978-3-642-14363-2_6).
- 698  
699  
700  
701