A Survey of Knowledge Graphs, Retrieval-Augmented Generation, and Large Language Models in Conversational AI

www.surveyx.cn

Abstract

This survey explores the integration of Knowledge Graphs (KGs), Retrieval-Augmented Generation, and Large Language Models (LLMs) in advancing Conversational AI. It highlights the transformative impact of these technologies in enhancing factual reasoning, intent classification, and text matching. KGs are crucial for structured information representation, improving AI reasoning capabilities, and addressing LLMs' limitations in recalling accurate information. The survey examines KGs' application in specialized domains like cyber threat intelligence and conversational recommender systems, emphasizing the need for automated report generation and scalable conversational agents. The integration of KGs with LLMs enhances adaptability and scalability, crucial for educational AI tutors and medical applications. Retrieval-Augmented Generation improves text generation by retrieving relevant data, while advanced AI techniques in dialogue systems facilitate interactive communication. The survey underscores challenges in KG construction, integration with LLMs, and retrieval mechanisms, proposing future research directions to refine these integrations and expand their applicability. By providing a comprehensive overview, the survey contributes to advancing AI technologies and their applications across diverse domains.

1 Introduction

1.1 Scope and Significance of the Survey

The convergence of Knowledge Graphs (KGs), Retrieval-Augmented Generation (RAG), and Large Language Models (LLMs) signifies a transformative shift in artificial intelligence (AI), particularly within natural language processing (NLP) and conversational AI. This survey elucidates the integration of these technologies, focusing on pivotal challenges such as factual reasoning, intent classification, and short text matching. KGs are essential for structuring information, thereby enhancing AI systems' reasoning capabilities, which is critical for LLMs that often struggle with information recall. By embedding structured data from KGs, LLMs can markedly improve their reasoning processes [1]. This integration is crucial for addressing the limitations of existing models reliant on unstructured corpora, facilitating the incorporation of factual knowledge into pre-trained language models (PLMs) like BERT [2].

The survey further investigates the application of KGs in specialized domains, such as cyber threat intelligence, where they assist in constructing attack knowledge graphs. This capability is vital for dynamic representation approaches, exemplified by the CKG model, which effectively integrates contextual information and external knowledge to capture the multifaceted meanings of polysemous words. Such integration addresses the shortcomings of existing language representation models, thereby enhancing performance across various NLP tasks [3, 4, 5, 6, 7]. In conversational AI, conversational recommender systems (CRS) leverage multi-turn dialogues to deliver high-quality recommendations, thereby improving customer satisfaction and intent classification.

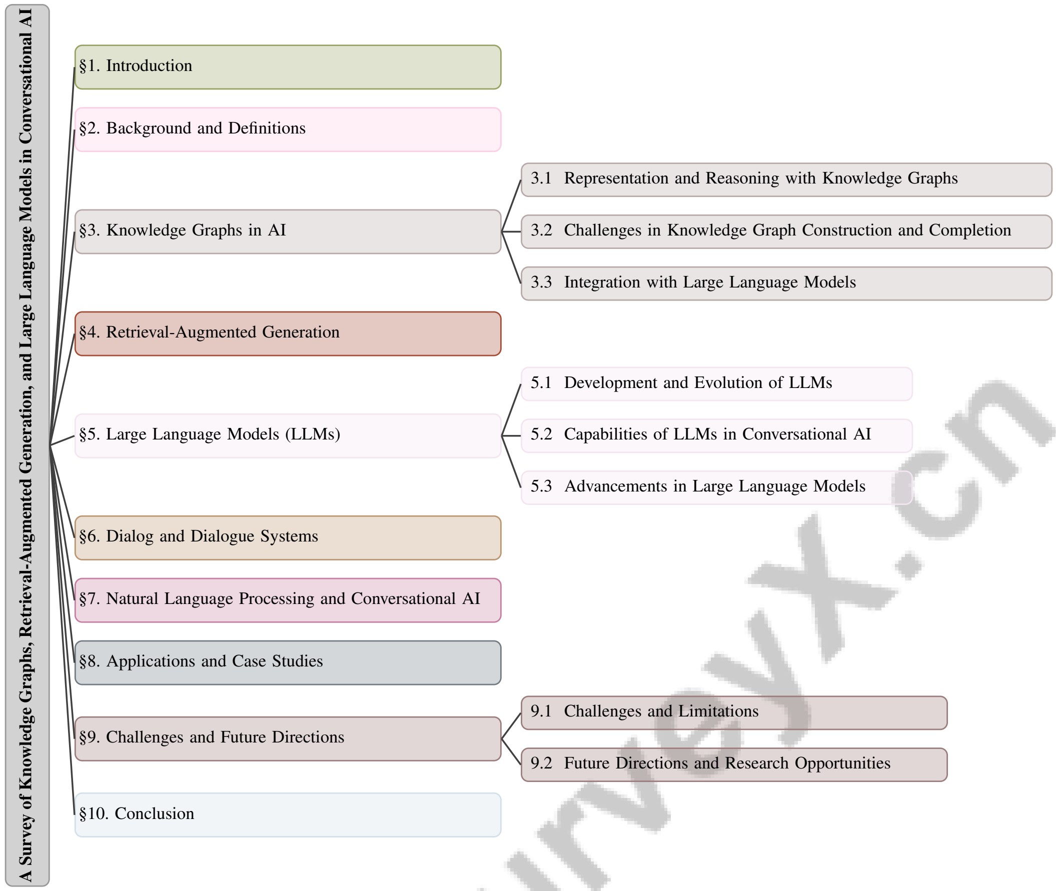

Figure 1: chapter structure

The survey emphasizes the necessity for automating intelligence report generation through advanced narrative construction techniques, streamlining traditionally manual processes. It also underscores the importance of efficient information retrieval in specialized domains, as exemplified by the AI Research Navigator platform, which merges classical keyword search with neural retrieval to enhance literature discovery. Additionally, the survey advocates for the development of distributed multi-agent platforms that support scalable conversational agents, facilitating effective interactions and data management across applications [8, 9]. These advancements are critical for personalizing AI applications and integrating deep contextualized knowledge representations, ultimately improving AI systems' effectiveness across diverse fields. By providing a comprehensive overview of the current state and potential of integrating KGs, RAG, and LLMs, this survey significantly contributes to the advancement of AI technologies and their applications.

1.2 Motivation and Relevance

This survey is motivated by the need to bridge the gap between traditional AI systems and the evolving demands of NLP and conversational AI. Current AI architectures often face scalability and adaptability challenges, which restrict their ability to incorporate new capabilities without extensive regression testing [10]. This work seeks to address these issues by exploring the integration of LLMs and KGs to enhance AI systems' adaptability and scalability, thereby improving conversational question answering systems' efficacy in managing incomplete follow-up questions [11].

A key motivation is the advancement of conversational recommendation systems (CRS), which rely on KGs yet often lack the integration of contextual dialogue information [12]. This survey examines how LLMs can enhance concept representation and tackle challenges in textual information construction and adaptation for recommendation tasks, particularly in educational AI tutors, where maintaining accuracy across diverse academic disciplines is essential [6]. This addresses prevalent issues such as information hallucination [13].

In the medical domain, there is increasing interest in utilizing LLMs for applications that require integrating domain-specific knowledge into neural generative models to improve medical conversations essential for accurate diagnosis and treatment recommendations [14]. The survey also highlights the need for specialized tools that leverage recent advancements in AI and NLP to enhance scholarly literature search and organization, addressing the complexities of navigating intricate graphical interfaces.

Moreover, the exploration of KGs and multi-modal learning provides insights into the evolving MMKG research landscape, addressing existing research divides [15]. The motivation extends to the necessity of dialogue systems that generate responses grounded in factual information from KGs [16] and managing the overwhelming growth of unstructured content in regulatory domains [17].

The survey underscores the importance of enhancing emergency decision-making processes through the integration of KGs, RAG, and LLMs [18], as well as the necessity for mixed-initiative interactions in emotional support systems [19]. By examining these facets, the survey illustrates its relevance in advancing AI technologies and improving conversational systems' effectiveness across diverse applications.

1.3 Structure of the Survey

This survey is systematically organized to provide a comprehensive exploration of the integration and application of KGs, RAG, and LLMs within the field of Conversational AI. The paper is structured into several key sections, each addressing distinct aspects of these technologies and their interplay.

The survey begins with an **Introduction**, which discusses the scope, significance, motivation, and relevance of the survey, laying a foundational understanding of the convergence of these technologies and their importance in advancing AI systems.

Following the introduction, the **Background and Definitions** section delves into core concepts underpinning the survey, offering detailed explanations of KGs, RAG, LLMs, dialog, NLP, conversational AI, and dialogue systems, elucidating their individual roles and interconnections within the AI landscape.

The third section, **Knowledge Graphs in AI**, examines KGs' role in representing structured information and enhancing reasoning capabilities. It further explores challenges in constructing and completing KGs and their integration with LLMs.

The paper then transitions to a detailed examination of **Retrieval-Augmented Generation (RAG)**, emphasizing its critical role in enhancing text generation by integrating relevant external data sources, such as KGs, to improve the accuracy and relevance of generated responses. This section scrutinizes challenges faced by LLMs in producing reliable outputs for knowledge-intensive queries and highlights innovative frameworks like Mindful-RAG and TRACE, which address common pitfalls in knowledge retrieval and multi-hop reasoning. Additionally, it discusses RAG's application in narrative construction for intelligence report generation, showcasing its potential to streamline complex event information synthesis into coherent narratives [20, 8, 21, 22]. Current methodologies, applications, and challenges in integrating retrieval mechanisms with generative models are discussed.

The focus then shifts to **Large Language Models (LLMs)**, discussing their development, capabilities, and recent advancements, particularly in the context of conversational AI. This section analyzes their strengths and limitations, offering insights into their evolution and future directions.

Subsequently, the survey explores **Dialog and Dialogue Systems**, highlighting the technologies and methodologies involved in their creation. The evolution and current state of AI-driven search systems are examined, emphasizing the integration of advanced AI techniques such as neural retrieval and semantic search to enhance interactive communication. Systems like the AI Research Navigator and a conversational search platform utilizing KGs aim to improve scholarly literature discovery and organization by addressing vocabulary differences and data navigation complexities. The architecture and functional components of these systems are detailed, alongside performance evaluations demonstrating their effectiveness compared to traditional search methods [23, 9].

AI-generated, for reference only.

The role of **Natural Language Processing and Conversational AI** is then discussed, emphasizing advancements in NLP techniques that enable machines to comprehend and respond to human language. This section highlights the utilization of contextual and semantic information to enhance user interactions.

To provide practical insights, the **Applications and Case Studies** section presents real-world examples of successful technology implementations, analyzing the impact of integrating KGs, RAG, and LLMs across various domains, including recommendation systems, question answering systems, and educational and knowledge management systems.

The survey concludes with a discussion on **Challenges and Future Directions**, identifying current limitations and exploring potential future research opportunities. This section aims to address challenges faced in these technologies' integration and application, guiding future advancements in the field.

The **Conclusion** synthesizes the principal findings of the research, emphasizing the critical role of advanced technologies such as neural search platforms, KGs, and narrative construction in propelling AI forward. It underscores the significant potential for future research and development, particularly in enhancing scientific literature analysis and management, improving decision-making in high-stakes domains, and automating intelligence report generation, thereby addressing the challenges posed by the rapid expansion of AI research and applications [24, 25, 8, 9, 26].The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Background and Definitions

Knowledge Graphs (KGs), Retrieval-Augmented Generation (RAG), Large Language Models (LLMs), Dialog, Natural Language Processing (NLP), Conversational AI, and Dialogue Systems are fundamental components of contemporary AI systems, each contributing unique enhancements. KGs serve as structured frameworks for representing information, enabling AI to transform unstructured data into interconnected entities that facilitate reasoning and context-awareness. The TEGKE method exemplifies the extraction of rich structural and semantic information through multi-hop topic KGs, crucial for applications like topic-to-essay generation [21]. Despite the capabilities of Pre-trained Language Models (PLMs), they often struggle with world knowledge acquisition due to their reliance on large-scale corpora, highlighting the role of KGs in providing structured context [1].

RAG enhances text generation by retrieving pertinent data from extensive databases, addressing LLMs' challenges in recalling and applying factual knowledge. This approach is particularly beneficial for generating knowledge-grounded content, thereby improving the accuracy and relevance of AI-generated responses. The integration of dynamic KGs into conversational systems to manage dialogue states demonstrates this method's effectiveness, especially in mixed-initiative dialogue systems designed to enhance emotional support interactions [19].

LLMs, such as BERT and GPT, have revolutionized NLP by enabling human-like text generation. However, they often encounter difficulties in recalling factual knowledge and understanding complex terminology, particularly in specialized fields like medicine [27]. The incorporation of KGs with LLMs enhances these models' reasoning capabilities, which is essential for applications requiring precise information extraction and reasoning, such as multi-type medical text classification [28].

Dialog and Dialogue Systems are central to Conversational AI, facilitating interactive communication between humans and machines. These systems employ advanced NLP techniques to comprehend and respond to human language, promoting natural and effective interactions. A significant challenge in their development is accurately recognizing spoken utterances and annotating them with machine-understandable meanings linked to KGs [29]. The synergy between KGs and LLMs enriches dialogue systems by enhancing their understanding of polysemous words and context-dependent meanings, thereby improving the overall dialogue experience.

NLP underpins Conversational AI, encompassing a wide range of techniques that enable machines to process and understand human language. It plays a vital role in tasks such as fact verification, addressing misinformation challenges, and reducing the labor-intensive nature of traditional methods.

AI-generated, for reference only.

The construction of KGs from domain-specific data and their application in extracting task-based information provide timely insights, underscoring NLP's critical role in AI applications [17].

The integration of these technologies within the AI landscape highlights their collective potential to enhance AI systems' capabilities. By combining KGs, RAG, and LLMs, AI systems can achieve more accurate, context-aware, and reliable outcomes, which is crucial for overcoming the limitations of existing models and enabling innovative applications across various domains [18].

In recent years, the integration of Knowledge Graphs (KGs) within artificial intelligence (AI) has garnered significant attention due to their potential to enhance reasoning and dialogue systems. As illustrated in Figure 2, this figure highlights the multifaceted applications of KGs, emphasizing their contributions to improving AI reasoning, contextual understanding, and accuracy. Furthermore, it addresses the challenges associated with the construction and completion of these graphs, as well as their integration with Large Language Models (LLMs). Such insights are crucial for advancing the field of AI and understanding the complexities involved in leveraging KGs effectively.

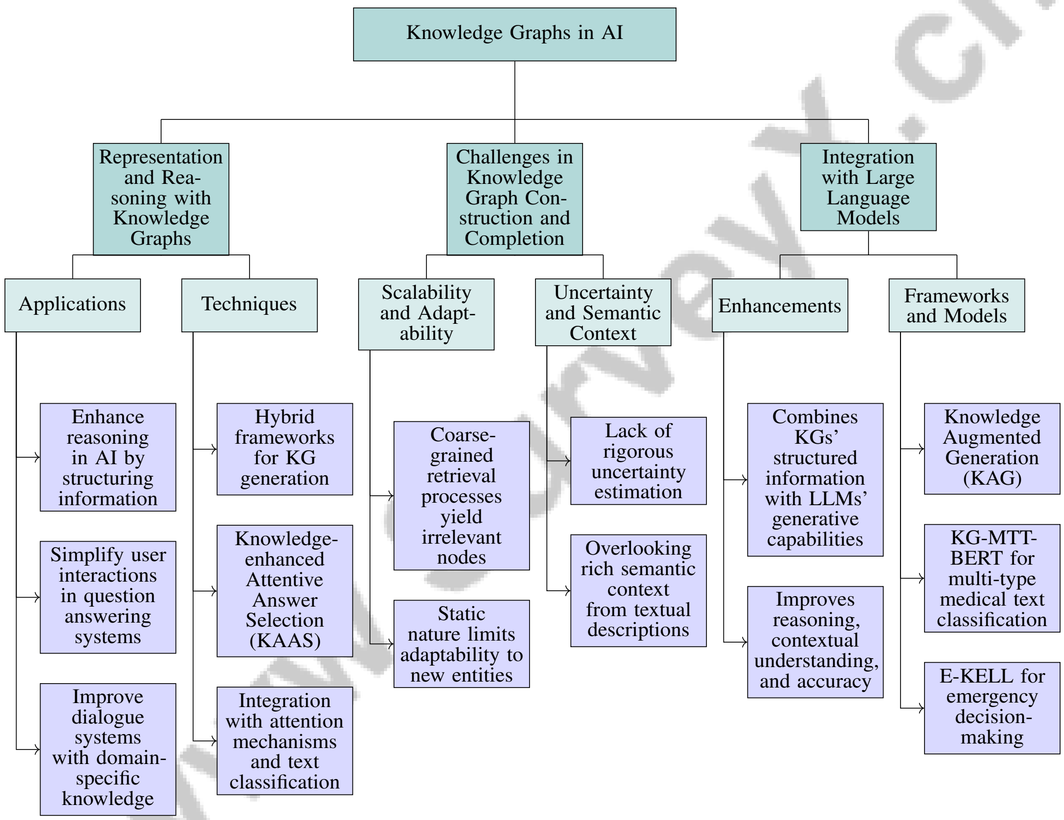

Figure 2: This figure illustrates the role of Knowledge Graphs (KGs) in AI, highlighting their applications in enhancing reasoning and dialogue systems, challenges in construction and completion, and integration with Large Language Models (LLMs) to improve AI reasoning, contextual understanding, and accuracy.

3 Knowledge Graphs in AI

3.1 Representation and Reasoning with Knowledge Graphs

Knowledge Graphs (KGs) are pivotal in AI for structuring information to enhance reasoning. They transform unstructured data into interconnected entities and relationships, forming a foundation for complex reasoning across AI applications. As illustrated in Figure 3, the role of Knowledge Graphs in AI encompasses various applications, construction methods, and integration techniques across different domains. Hybrid frameworks that combine supervised and unsupervised methods have advanced KG generation, leading to more precise triples, thereby increasing their utility [26]. In question answering, KGs simplify user interactions with complex datasets, like biological information, without requiring users to understand semantic technologies [30]. The Knowledge-enhanced Attentive Answer Selection (KAAS) model exemplifies how KGs refine answer selection, highlighting their role in enhancing AI decision-making [31].

KGs are constructed from diverse sources, such as textbooks, where methodologies extract triples from glossaries and text to capture domain-specific knowledge [24]. This is particularly beneficial in educational contexts demanding precise knowledge representation. In emotional support systems, KGs enhance mixed-initiative interactions by incorporating mental health knowledge, improving conversation quality and support [19]. This underscores KGs' potential to enrich dialogue systems with domain-specific knowledge for more meaningful interactions.

The integration of KGs with advanced attention mechanisms in text classification tasks allows for better utilization of prior knowledge, focusing on relevant concepts while minimizing irrelevant data [32]. Graph-augmented learning to rank models demonstrate KGs' utility in optimizing query responses by selecting the most relevant sub-KGs [33]. The synergy between KGs and pre-trained language models through Knowledge-Enhanced Entity Representation Learning (KERL) enhances entity representation by encoding textual descriptions and leveraging structured knowledge [34]. This is critical for improving AI reasoning capabilities.

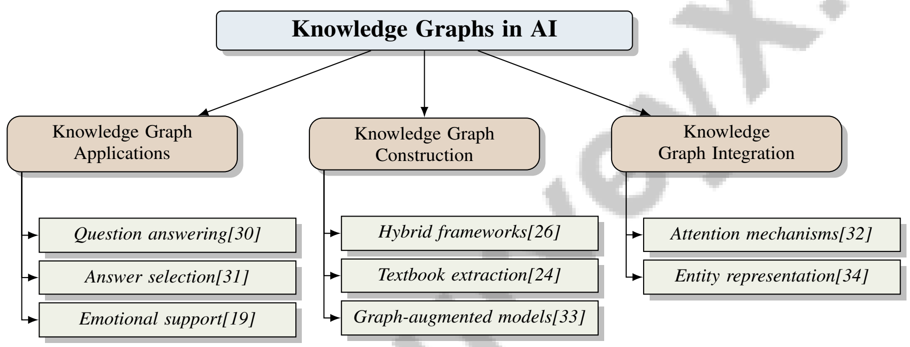

Figure 3: This figure illustrates the role of Knowledge Graphs in AI, categorizing their applications, construction methods, and integration techniques in various domains.

3.2 Challenges in Knowledge Graph Construction and Completion

Constructing and completing Knowledge Graphs (KGs) involves significant challenges that affect their scalability and effectiveness. Coarse-grained retrieval processes often yield knowledge subgraphs (KSGs) with irrelevant nodes, complicating answer identification [33]. The static nature of KGs limits adaptability to new entities and relations, constraining performance [35]. Current methodologies often overlook the rich semantic context from textual descriptions, impeding the construction of comprehensive KGs [34]. Assessing Large Language Models (LLMs) for completing missing information in static KGs highlights difficulties in deducing absent entities and relationships [36].

Integrating uncertainty quantification into KG frameworks, especially those combined with LLMs, presents challenges due to complex architectures. Current frameworks lack rigorous uncertainty estimation, limiting their reliable application in high-stakes environments. Addressing this is crucial for enhancing KG-LLM systems' trustworthiness and effectiveness, as shown by the UAG (Uncertainty Aware Knowledge-Graph Reasoning) framework, which incorporates uncertainty quantification with theoretical guarantees on prediction accuracy [3, 37, 38]. Overcoming these challenges is vital for advancing KG construction and completion, enabling more effective AI support across domains.

3.3 Integration with Large Language Models

Integrating Knowledge Graphs (KGs) with Large Language Models (LLMs) significantly advances AI, enhancing reasoning, contextual understanding, and accuracy. This synergy combines KGs' structured information with LLMs' generative capabilities, improving performance across fields. Frameworks like Knowledge Augmented Generation (KAG) and Knowledge Graph-Retrieval Augmented Generation (KG-RAG) address reasoning gaps and enhance information retrieval and reasoning processes, enabling sophisticated applications such as domain-specific QA systems [39, 40, 3].

The KELM framework integrates structured knowledge from KGs into pre-trained language models during fine-tuning, enhancing world knowledge understanding [1]. The KG-MTT-BERT model extends BERT for multi-type medical text classification by incorporating a medical knowledge graph, underscoring KGs' importance in providing structured context that enhances AI-generated responses' relevance and precision [28].

The E-KELL framework illustrates KGs' integration with LLMs, enhancing emergency decision-making by utilizing structured knowledge [18]. In essay generation, the TEGKE method shows how external KG knowledge can enhance generated content's accuracy and relevance [21]. The MKGformer model, a hybrid transformer for multimodal knowledge graph completion, integrates visual and textual data, illustrating KGs' versatility in enriching LLMs with multimodal information [41].

Baek's method introduces augmenting knowledge directly in LLMs' input using KG-retrieved facts, enhancing models' knowledge retrieval and application capabilities [2]. A benchmark to assess LLMs' suitability in leveraging external KG knowledge through natural language prompts emphasizes structured knowledge's importance in refining LLM evaluation and application [36].

Integrating Knowledge Graphs with Large Language Models significantly enhances AI systems by improving reasoning capabilities, accuracy, and contextual awareness, as evidenced by advancements in Natural Language Processing tasks. This integration allows AI models to access and leverage organized world knowledge through fact triples from Knowledge Graphs, enabling dynamic extraction of relevant information. Consequently, the performance of these systems is markedly improved, demonstrated by substantial gains in various applications, including text classification and semantic query processing, where the combination of structured knowledge and advanced language understanding leads to more effective outcomes [3, 42, 43, 44, 26]. By leveraging structured knowledge within KGs, LLMs can overcome inherent limitations, facilitating more advanced and contextually aware AI applications across diverse domains. This integration is crucial for enhancing the depth and accuracy of AI reasoning capabilities, enabling more meaningful interactions across various applications.

4 Retrieval-Augmented Generation

4.1 Enhancing Retrieval-Augmented Generation

Enhancing Retrieval-Augmented Generation (RAG) involves integrating advanced techniques to refine both information retrieval and content generation, thereby improving the contextual relevance and accuracy of AI outputs. The SubgraphRAG model exemplifies this by utilizing a lightweight multilayer perceptron with a parallel triple-scoring mechanism, which facilitates efficient subgraph retrieval and encodes structural distances to enhance retrieval efficacy [45]. These innovations ensure that retrieved data is pertinent and contextually suitable for specific tasks.

The Knowledge Augmented Generation (KAG) framework further illustrates advancements in RAG by leveraging Knowledge Graphs (KGs) and vector retrieval to generate coherent responses in specialized domains such as law and medicine. This approach integrates structural and semantic knowledge, significantly enhancing the precision and relevance of AI-generated outputs [39].

Additionally, the KIC-GPT model combines large language models (LLMs) with triple-based Knowledge Graph Completion (KGC) techniques, enabling the simultaneous use of structural and semantic knowledge without incurring additional training costs [46]. This integration enriches RAG systems' content generation capabilities, fostering more effective AI applications.

The MKGformer model introduces a novel architecture that combines coarse-grained prefix-guided interaction with fine-grained correlation-aware fusion, improving multimodal data fusion and enhancing contextually aware output generation [41].

Furthermore, the UAG framework incorporates multi-step reasoning with conformal prediction, offering theoretical guarantees on prediction sets and managing error propagation through an error rate control module [38]. This ensures that RAG systems produce reliable content, which is crucial for high-stakes applications.

The MEG model showcases the effectiveness of conditioning LLM responses on factual knowledge from KGs, thereby enhancing contextual understanding and accuracy [27]. This targeted knowledge injection is vital for generating precise and contextually relevant responses, particularly in information-sensitive domains.

These advancements highlight the transformative potential of merging structured knowledge with sophisticated retrieval techniques and LLMs. Innovations such as Knowledge Graph-enhanced Retrieval-Augmented Generation (KG-RAG) and TRACE markedly improve the accuracy, relevance, and contextual awareness of AI-generated content. By utilizing structured knowledge representation and reasoning chains, these methods address critical challenges like factual accuracy and coherence in instructional contexts, ultimately enhancing user outcomes and engagement across diverse fields, including personalized education and complex question answering [20, 9, 13].

5 Large Language Models (LLMs)

5.1 Development and Evolution of LLMs

The development of Large Language Models (LLMs) has been marked by substantial progress in their text comprehension and generation capabilities, particularly through the integration of external knowledge sources like Knowledge Graphs (KGs). Models such as RELATIONLM and CoCoLM exemplify this trend by utilizing relational memory from KGs to enhance text generation and understanding of complex commonsense, respectively [47, 48]. Addressing behavioral inconsistencies remains essential, as they affect LLM reliability in downstream tasks [49]. The GraphEval framework addresses this by focusing on hallucination detection in closed domains, ensuring contextually grounded outputs [50]. Furthermore, challenges in Knowledge Graph Question Answering (KGQA), such as data sparsity and complex query interpretation, necessitate advanced inference capabilities [51].

Innovative methods, such as Large Language Models-guided Dynamic Adaptation (LLM-DA), enable LLMs to dynamically update temporal logical rules based on historical data, enhancing adaptability [52]. The SyntheT2C framework improves LLM performance on tasks like Text2Cypher by generating synthetic datasets for fine-tuning [53]. The KG-Copy method demonstrates the effectiveness of merging factual knowledge with conversational context to improve response quality [16], while KGs have significantly enhanced citation generation metrics, such as METEOR and ROUGE-1 scores [54].

The evolution of LLMs is significantly driven by advancements in integrating structured knowledge sources, which enhance their ability to deliver accurate and contextually relevant responses. Frameworks like Mindful-RAG address behavioral inconsistencies by targeting specific knowledge retrieval failures, improving memory retention, information incorporation, and domain-specific accuracy [55, 22, 56]. These innovations are crucial for expanding the applicability and effectiveness of LLMs across diverse fields, including education and healthcare.

5.2 Capabilities of LLMs in Conversational AI

LLMs have significantly transformed conversational AI by enhancing the generation and comprehension of human-like text, thereby improving interaction quality in dialogue systems. The integration of structured knowledge from Knowledge Graphs (KGs) is crucial for generating coherent and contextually accurate responses. The KERL framework, for instance, uses intrinsic information from entity descriptions to boost recommendation accuracy in conversational systems [34]. Similarly, the JAKET framework enhances reasoning capabilities by integrating knowledge with language understanding, improving comprehension of world knowledge, entities, and relations [35].

LLMs manage diverse user inputs and tailor interactions using knowledge graphs and contextual information for personalized conversations, enhancing user engagement and satisfaction in applications like educational chatbots and conversational recommendation systems [23, 36, 12, 57]. Relational memory integration further enhances coherence and logical consistency in generated text, improving the overall dialogue experience.

In practical applications, LLMs generate factually accurate answers by augmenting inputs with relevant external knowledge. Despite challenges in generating semantically correct SPARQL queries, structured knowledge integration significantly enhances LLMs' ability to process complex language data. Techniques like Ontology-based Query Check (OBQC) and LLM Repair have improved question-answering accuracy from 16

LLMs' extensive capabilities in conversational AI support advanced intent classification, personalized interactions, and improved reasoning abilities. Recent research demonstrates their effectiveness in knowledge graph-related tasks, such as knowledge graph completion, even in zero- or few-shot scenarios. However, challenges like generating hallucinated responses necessitate careful prompt design for accuracy. By integrating LLMs with structured knowledge graphs and human mentorship, their potential can be harnessed while mitigating associated risks [36, 57], advancing the effectiveness and reliability of conversational systems across various applications.

5.3 Advancements in Large Language Models

Recent advancements in Large Language Models (LLMs) focus on enhancing accuracy, consistency, and structured knowledge integration. Deterministic checks with ontology semantics for query repair significantly improve accuracy compared to traditional methods by leveraging structured ontological knowledge [58]. Fine-tuning processes have improved by identifying Frequency Shock and Range Shift as detrimental effects, with model mixing and mixture fine-tuning offering more stable processes that preserve factual knowledge [59].

Addressing biases within LLMs is critical, with strategies refactoring harmful stereotypes into structured knowledge graphs and employing adversarial techniques to induce bias [60]. This highlights the potential of knowledge graphs in mitigating bias and underscores the importance of adversarial strategies in refining LLM outputs. The LLM-KG-Bench framework represents a significant advancement in LLM evaluation by incorporating automated procedures and configurable task sizing, providing insights into performance across various tasks [61].

Consistency testing has been enhanced through methods like KG-LLM, which evaluates LLM comprehension of concept hierarchies using knowledge graphs. Automated test case generation provides a robust mechanism for assessing LLM consistency, ensuring logical coherence in outputs [62]. The KONTEST framework generates test cases aimed at uncovering LLM inconsistencies, advancing automated consistency testing [49].

These advancements underscore efforts to enhance LLM capabilities through structured knowledge integration, refined fine-tuning processes, and robust evaluation frameworks. By addressing challenges such as accuracy, bias, and consistency, recent innovations leverage techniques like ontology-based query checks and knowledge-augmented generation to improve LLM reliability and effectiveness across various domains. For instance, integrating knowledge graphs has improved question-answering accuracy from 16

6 Dialog and Dialogue Systems

6.1 Advanced AI Techniques in Dialogue Systems

Advanced AI techniques have significantly enhanced dialogue systems, enabling more natural and context-aware interactions. Structured knowledge paths allow systems to dynamically shift topics while maintaining coherence, as demonstrated by Wu et al.'s benchmark, which combines knowledge planning with dialogue generation to improve user engagement [63]. Knowledge Graphs (KGs) are pivotal in understanding context and semantics, enhancing dialogue systems' ability to generate accurate and relevant responses by capturing complex entity relationships. Transforming natural language queries into formal queries facilitates the retrieval of current information from KGs, thereby improving conversational AI and question-answering systems. Integrating KGs into response generation has notably improved knowledge groundedness, allowing both goal-oriented and non-goal-oriented systems to produce factually enriched dialogues [64, 65, 16, 66].

Retrieval-Augmented Generation (RAG) techniques further advance dialogue capabilities by enabling large language models (LLMs) to access and integrate external information dynamically, improving response relevance and accuracy. Challenges such as irrelevant information retrieval and question intent discernment are addressed by frameworks like Mindful-RAG and TRACE, which emphasize intent-based retrieval and knowledge-grounded reasoning chains, enhancing performance in knowledge-intensive tasks [20, 22, 67].

LLMs have been instrumental in advancing dialogue systems by generating human-like text and enhancing dialogue fluency. By integrating structured knowledge from KGs, these systems not only generate coherent responses but also provide domain-specific insights, improving tasks requiring accurate factual knowledge and nuanced understanding, such as academic research and professional applications. Techniques like Knowledge Augmented Generation (KAG) and KG-enhanced Query Processing (KGQP) optimize interactions between LLMs and KGs, leading to precise information retrieval and reasoning capabilities [3, 39, 68, 69].

These advanced AI techniques, including knowledge planning, KGs, RAG, and LLMs, have transformed dialogue systems into sophisticated tools capable of complex interactions. Innovations like the TEGKE model, which uses comprehensive knowledge for essay generation, and AI Research Navigator, which employs neural retrieval and KGs for literature discovery, signify a leap in conversational AI. Systems leveraging KGs in conversational search address traditional search challenges, enabling efficient navigation of complex data [23, 21, 9, 11].

6.2 Application-Specific Dialogue Systems

Application-specific dialogue systems are tailored to particular domains, integrating specialized domain knowledge with AI techniques like natural language processing and machine learning for precise interactions. By employing entity extraction and semantic role labeling, these systems create structured knowledge graphs that enhance data management and provide timely insights [17, 3, 9, 26].

In healthcare, dialogue systems enhance communication by integrating medical knowledge graphs, improving understanding of complex terminology and supporting patient care [27]. In education, systems create adaptive learning environments by leveraging educational knowledge graphs, offering personalized tutoring and support across academic disciplines [13]. Customer service applications benefit from RAG techniques, which integrate external knowledge sources for timely and accurate responses, while frameworks like Mindful-RAG and TRACE improve response relevance through intent-based retrieval [20, 22, 13].

In the legal field, dialogue systems assist with document retrieval and analysis by integrating legal knowledge graphs, ensuring responses are grounded in accurate information [39]. These systems illustrate the potential of customized AI solutions to enhance interactions across diverse fields, improving engagement and providing tailored insights. From educational support chatbots to scholarly searches, these systems foster effective exchanges between users and technology by integrating domain-specific knowledge and advanced AI techniques, ensuring contextually relevant and accurate responses [9, 70, 23, 57, 71].

7 Natural Language Processing and Conversational AI

7.1 Role of Contextual Information

Contextual information plays a crucial role in enhancing Natural Language Processing (NLP) and Conversational AI, facilitating the generation of coherent, relevant, and context-aware responses. Deep contextualized knowledge representations, such as CoLAKE, integrate language with a unified word-knowledge graph structure, enabling conversational agents to better discern user intent and preferences. This capability is further enhanced by conversational recommendation systems that utilize contextual embeddings to align item metadata with dialogue context, thereby improving response quality [72, 23, 12, 42]. The integration of contextual information allows systems to navigate the complexities of human language, addressing nuances like polysemy and idiomatic expressions crucial for maintaining dialogue flow.

Knowledge Graphs (KGs) are instrumental in enriching dialogue systems with structured knowledge, facilitating the dynamic integration of relevant contextual data to enhance the accuracy and relevance of generated responses. For instance, KGs can provide background on entities discussed in a conversation, enabling systems to generate informed replies [19]. Retrieval-Augmented Generation (RAG) techniques further enhance contextual information use by allowing systems to dynamically

AI-generated, for reference only.

retrieve and incorporate external data, particularly advantageous in scenarios requiring up-to-date or specialized information [39].

Large Language Models (LLMs) such as BERT and GPT leverage extensive contextual data during training to produce coherent and contextually aware responses, thereby enhancing dialogue naturalness and fluency [27]. Moreover, advanced NLP techniques enable systems to manage dialogue states and track user intents over multi-turn interactions. By maintaining a thorough understanding of conversational context, these systems deliver more relevant and personalized responses, improving user satisfaction and engagement [35].

Effectively utilizing contextual information in NLP for Conversational AI is essential for developing systems capable of engaging in meaningful, contextually aware interactions. By integrating structured knowledge repositories like KGs with advanced retrieval mechanisms and LLMs, these systems enhance their ability to comprehend and respond to the intricacies of human language, ultimately improving semantic understanding and query accuracy. This integration fosters the development of sophisticated conversational agents adept at managing complex information, as evidenced in applications such as scholarly research and domain-specific question answering [3, 23, 22].

7.2 Semantic Information and Comprehension

Semantic information is fundamental to understanding and generating responses in NLP and conversational AI systems, enabling the interpretation of meanings within context to enhance output relevance and accuracy. Integrating semantic information into AI models significantly improves comprehension abilities, allowing for nuanced understanding and enriching response meaningfulness. This is particularly evident in systems utilizing Knowledge Graphs and Large Language Models, which facilitate precise query processing and context-aware retrieval, leading to more effective interactions in applications like academic research and personalized education [3, 13, 20, 17, 9].

Knowledge Graphs (KGs) contribute significantly to semantic information by representing entities and their relationships in structured formats, allowing AI systems to access rich semantic context essential for understanding complex queries and generating coherent responses. For instance, KGs enhance comprehension of polysemous words and context-dependent meanings in dialogue systems, improving engagement in meaningful interactions [29]. Incorporating semantic information through RAG techniques enhances AI systems' capabilities to dynamically retrieve and integrate relevant data, ensuring responses are semantically accurate and contextually enriched with up-to-date, domain-specific information [39].

LLMs like BERT and GPT leverage semantic information during training to generate human-like text, understanding language nuances, including idiomatic expressions and contextual variations, enabling semantically coherent and contextually aware responses [27]. Additionally, systems employing advanced NLP techniques to manage dialogue states and track user intents across multi-turn interactions benefit from a comprehensive understanding of semantic context, providing more relevant and personalized responses, thus enhancing user satisfaction and engagement [35].

8 Applications and Case Studies

8.1 Enhancements in Recommendation Systems

The integration of Knowledge Graphs (KGs) and advanced AI techniques has significantly enhanced recommendation systems, enabling them to provide more personalized and contextually relevant suggestions. The KGenSam model exemplifies this progress by using contextual information from user interactions to dynamically adapt to preferences, thus improving recommendation quality [73]. The Knowledge-Enhanced Conversational Recommender (KECR) system advances this field by offering explainable recommendations and bridging semantic gaps through explicit reasoning, thereby increasing user engagement and satisfaction [74]. Furthermore, the Knowledge-Enhanced Entity Representation Learning (KERL) framework achieves state-of-the-art performance by integrating rich intrinsic information from entity descriptions, enhancing the understanding of user preferences and recommendation accuracy [34].

The application of Knowledge Augmented Generation (KAG) in government service QA applications underscores the role of structured knowledge in achieving high accuracy in user inquiries, highlighting

KGs' potential to enhance recommendation precision across domains [39]. In clinical settings, knowledge-grounded conversational systems improve the detection of implicit symptoms, enriching recommendation systems with domain-specific knowledge for more informed decisions [75]. These advancements illustrate the transformative potential of integrating KGs and AI techniques, enabling recommendation systems to deliver personalized, precise, and explainable suggestions. Leveraging large language models (LLMs) and item metadata encoders further optimizes user experiences by generating high-quality recommendations and responses [6, 12].

8.2 Advancements in Question Answering Systems

Recent advancements in question answering (QA) systems have been significantly bolstered by the integration of Knowledge Graphs (KGs) and advanced AI methodologies, enhancing their ability to deliver precise and contextually relevant answers. KGs enable QA systems to perform complex reasoning with interconnected data points, as demonstrated by the LAGRANGE framework, which surpasses existing benchmarks in cyclic evaluations with superior precision and recall rates in KG reconstruction [76]. In specialized domains like medicine, incorporating domain-specific knowledge into QA systems has led to notable performance improvements, as evidenced by the MEG framework, which embeds medical knowledge to enhance accuracy [27].

Cognitive QA platforms, such as the one described by Mani et al., showcase AI-driven methodologies' potential to automate processes and improve IT support efficiency by reducing incident resolution times [77]. The emphasis on explainability in QA systems has grown, with methods like Template-based Explanation Generation for QA Systems (TEG-QA) improving user trust and understanding through transparent explanations [25, 31, 20, 78, 79]. Knowledge Graph Embeddings (KGE) have significantly boosted QA system performance by transforming high-dimensional knowledge graphs into low-dimensional embedding spaces, preserving essential structural information and enhancing interconnectivity among diverse resources [80, 81].

The convergence of KGs and advanced AI techniques marks a transformative advancement in QA systems, evidenced by sophisticated architectures leveraging Natural Language Processing and Machine Learning for effective knowledge extraction from scholarly publications. Multi-document knowledge graphs have improved retrieval efficiency and context understanding in question answering. Furthermore, conversational interfaces in scholarly search engines enhance user experiences by facilitating semantic search, addressing the complexities of navigating intricate data. Collectively, these innovations significantly enhance the accuracy and relevance of responses while contributing to more explainable and user-friendly systems across various domains [24, 82, 66, 23, 26].

8.3 Innovations in Educational and Knowledge Management Systems

Innovations in educational and knowledge management systems have been significantly propelled by the integration of Knowledge Graphs (KGs) and Large Language Models (LLMs), enhancing the management and dissemination of information. The KSR model demonstrates considerable improvements over baseline models, underscoring its effectiveness in semantic representation and its potential applications in knowledge and language integration [83]. This capability is essential for educational systems requiring precise semantic understanding to tailor content to diverse learning needs.

The KGV framework exemplifies advancements in knowledge management by enhancing LLM performance in assessing cyber threat intelligence quality, improving accuracy while reducing extensive data annotation requirements, which is particularly valuable in educational contexts where timely access to accurate information is crucial [37]. In educational applications, the COMET model has shown significant improvements in generating commonsense knowledge, outperforming traditional extractive methods and facilitating the creation of tools that provide learners with contextualized knowledge, thereby enhancing their understanding of complex concepts [84].

A comprehensive survey by Garg et al. offers a comparative analysis of software libraries for Knowledge Graph Embedding (KGE) training, providing valuable insights for educational institutions and knowledge management systems seeking to implement KGE methodologies to improve data integration and retrieval processes [85]. Furthermore, the model proposed by Dey et al. indicates its applicability beyond entertainment, suggesting its utility in developing adaptive learning environments and personalized educational content delivery [86]. The Text2KG-Bench benchmark, which includes datasets like Wikidata-TekGen and DBpedia-WebNLG, offers diverse examples for evaluating ontology-driven knowledge integration, essential for educational systems aiming to enhance their ontology-based knowledge representation capabilities [87].

These innovations underscore the transformative potential of integrating KGs and LLMs into educational and knowledge management systems. By employing advanced semantic representation techniques, enhancing commonsense knowledge generation, and integrating ontology-driven frameworks, these systems significantly improve their capacity to provide tailored, precise, and contextually appropriate educational experiences. This is achieved through semantically enhanced models that resolve ambiguities in commonsense knowledge and comprehensive knowledge enhancement methods leveraging both internal and external data sources, ultimately leading to more effective and high-quality educational content delivery [88, 21].

9 Challenges and Future Directions

9.1 Challenges and Limitations

The integration of Knowledge Graphs (KGs), Retrieval-Augmented Generation (RAG), and Large Language Models (LLMs) into AI systems faces several critical challenges impacting their effectiveness and scalability. A major issue is the dependence on the quality of KGs, crucial for performance in systems like emergency management, where decision reliability hinges on KG integrity [18]. Emotional support systems relying on mental health KGs also encounter accessibility and quality issues that can degrade performance [19].

Constructing and maintaining KGs, particularly for multi-modal tasks, involves significant complexity. The integration of external knowledge with minimal retrieval errors and contextual accuracy remains problematic, as existing methods often fall short [2]. Additionally, the computational demands of multi-hop KGs and adversarial training can become prohibitive, affecting scalability [21].

Addressing complex queries that require reasoning beyond simple fact retrieval is another limitation. Biological systems, for example, struggle with intricate multi-layered queries, leading to timeouts and failures [33]. Current benchmarks inadequately evaluate LLMs' reasoning capabilities in KG completion tasks, overlooking prompt engineering intricacies and output variability [36].

Dependence on specific document sources for KG generation limits model generalizability, restricting applicability. In conversational AI, models often struggle with delivering accurate responses and managing diverse question types, especially when incorporating recent KG information [89]. Current QA methods, including factoid-based, KG-based, and retrieval-based approaches, fail to address the complex nature of answers required in enterprise IT support [77]. Additionally, navigating graphical interfaces can cause cognitive overload, hindering effective information discovery.

The labor-intensive nature of writing and validating templates in methods like SyntheT2C also limits scalability. Open-domain QA frameworks may falter due to the absence of structured graph properties [32]. KERL faces challenges in effectively integrating multiple data sources, impacting performance if not managed properly [34]. Moreover, JAKET may not fully resolve all integration challenges associated with KGs, particularly in specialized domains [35].

Addressing these challenges requires ongoing research and development to enhance AI systems' ability to leverage structured knowledge, ensuring accuracy and coherence across applications. Recent advancements, such as the Knowledge Augmented Generation (KAG) framework, illustrate how bidirectional enhancements between LLMs and KGs can improve reasoning and generation performance in professional contexts. KAG advances existing RAG methods by focusing on LLM-compatible knowledge representation, logical reasoning, and mutual indexing between KGs and original data. Furthermore, integrating LLMs with structured repositories like the ANU Scholarly Knowledge Graph (ASKG) has significantly improved query accuracy and efficiency, highlighting these technologies' transformative potential in knowledge management and retrieval [39, 3].

9.2 Future Directions and Research Opportunities

Future research should focus on developing mechanisms to ensure consistency across different KGs and enhance model robustness against knowledge noise [1]. This involves refining retrieval mechanisms for more precise fact selection and exploring multi-hop knowledge retrieval to significantly improve AI-generated content accuracy [2].

The MKGformer model offers promising opportunities for enhancing visual representation capabilities and applying this model to additional natural language processing tasks [41]. Further research should aim to develop more effective answer selection models that operate on smaller knowledge subgraphs (KSGs) identified by ranking models, thereby improving information retrieval efficiency [33].

In Knowledge Graph Completion (KGC), future studies will focus on integrating additional LLMs and refining benchmarks to enhance their applicability across diverse KGC-related tasks [36]. Improving the language understanding capabilities of models like CR-Walker and exploring complex reasoning rules are essential for enhancing performance in tasks requiring intricate reasoning [89].

The expansion of cognitive systems in IT support, including enhancing system learning capabilities, broadening knowledge bases, and improving the handling of diverse user queries, remains a significant area for future exploration [77]. In personalized recommendations, research should focus on enhancing the interpretability and explainability of recommendations while exploring pre-trained user profiles to further personalize interactions [34].

These research directions are poised to significantly enhance AI systems' capabilities by integrating advanced technologies such as neural retrieval and natural language processing, thereby improving their sophistication, reliability, and contextual awareness across a wide range of applications. Innovative platforms like the AI Research Navigator facilitate efficient literature analysis and discovery, while integrating large language models with knowledge graphs enhances semantic query processing in scholarly contexts. Furthermore, developing hybrid answering models in enterprise IT support illustrates AI's adaptability to complex, multi-source information environments. Collectively, these advancements contribute to the evolution of more effective and adaptable AI technologies capable of better managing, analyzing, and disseminating knowledge [3, 77, 9, 26].

10 Conclusion

The convergence of Knowledge Graphs (KGs), Retrieval-Augmented Generation (RAG), and Large Language Models (LLMs) marks a pivotal advancement in Conversational AI, significantly enhancing the precision and contextual relevance of AI-generated outputs. By embedding structured knowledge, KGs have been instrumental in refining the representation of scholarly information, thus improving the accuracy of conversational question answering systems and setting new standards in the field. This integration underscores the critical role of comprehensive knowledge amalgamation, as evidenced by systems like TEGKE, which highlight the necessity of integrating both internal and external knowledge sources to produce superior content quality.

In the realm of information retrieval, conversational exploratory search systems have demonstrated considerable promise in advancing scholarly publication discovery, indicating a fertile area for future research. Mixed-initiative dialogue systems, exemplified by the KEMI framework, have achieved substantial progress in facilitating emotionally supportive interactions, emphasizing the importance of sophisticated AI techniques for enhancing dialogue systems' contextual awareness and interaction depth. Furthermore, methods such as KG-ANN have proven effective in augmenting text classification and natural language inference, showcasing the benefits of leveraging world knowledge within deep learning frameworks.

The landscape of future research is rich with opportunities, focusing on refining these integrations to bolster model resilience and extend their applicability across diverse sectors. The potential of models like UniD2T, which adeptly unify structured datasets with AI-generated narratives, illustrates the transformative capacity of these technologies in generating coherent natural language text from varied data formats. Collectively, these innovations are driving the evolution of AI, fostering new developments, and expanding the horizons of Conversational AI applications.

References

[1] Yinquan Lu, Haonan Lu, Guirong Fu, and Qun Liu. Kelm: Knowledge enhanced pre-trained language representations with message passing on hierarchical relational graphs, 2022.

[2] Jinheon Baek, Alham Fikri Aji, and Amir Saffari. Knowledge-augmented language model prompting for zero-shot knowledge graph question answering, 2023.

[3] Runsong Jia, Bowen Zhang, Sergio J. Rodríguez Méndez, and Pouya G. Omran. Leveraging large language models for semantic query processing in a scholarly knowledge graph, 2024.

[4] Linmei Hu, Zeyi Liu, Ziwang Zhao, Lei Hou, Liqiang Nie, and Juanzi Li. A survey of knowledge enhanced pre-trained language models, 2023.

[5] Mojtaba Nayyeri, Zihao Wang, Mst. Mahfuja Akter, Mirza Mohtashim Alam, Md Rashad Al Hasan Rony, Jens Lehmann, and Steffen Staab. Integrating knowledge graph embedding and pretrained language models in hypercomplex spaces, 2023.

[6] Qingyao Li, Wei Xia, Kounianhua Du, Qiji Zhang, Weinan Zhang, Ruiming Tang, and Yong Yu. Learning structure and knowledge aware representation with large language models for concept recommendation, 2024.

[7] Xunzhu Tang, Tiezhu Sun, Rujie Zhu, and Shi Wang. Ckg: Dynamic representation based on context and knowledge graph, 2022.

[8] Priyanka Ranade and Anupam Joshi. Fabula: Intelligence report generation using retrieval-augmented narrative construction, 2024.

[9] Marzieh Fadaee, Olga Gureenkova, Fernando Rejon Barrera, Carsten Schnober, Wouter Weerkamp, and Jakub Zavrel. A new neural search and insights platform for navigating and organizing ai research, 2020.

[10] Pranav Sharma. Lpar – a distributed multi agent platform for building polyglot, omni channel and industrial grade natural language interfaces, 2020.

[11] Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, and Gerhard Weikum. Look before you hop: Conversational question answering over knowledge graphs using judicious context expansion, 2019.

[12] Bowen Yang, Cong Han, Yu Li, Lei Zuo, and Zhou Yu. Improving conversational recommendation systems' quality with context-aware item meta information, 2021.

[13] Chenxi Dong, Yimin Yuan, Kan Chen, Shupei Cheng, and Chujie Wen. How to build an adaptive ai tutor for any course using knowledge graph-enhanced retrieval-augmented generation (kg-rag), 2025.

[14] Gabriel R. Rosenbaum, Lavender Yao Jiang, Ivaxi Sheth, Jaden Stryker, Anton Alyakin, Daniel Alexander Alber, Nicolas K. Goff, Young Joon Fred Kwon, John Markert, Mustafa Nasir-Moin, Jan Moritz Niehues, Karl L. Sangwon, Eunice Yang, and Eric Karl Oermann. Medg-krp: Medical graph knowledge representation probing, 2024.

[15] Zhuo Chen, Yichi Zhang, Yin Fang, Yuxia Geng, Lingbing Guo, Xiang Chen, Qian Li, Wen Zhang, Jiaoyan Chen, Yushan Zhu, Jiaqi Li, Xiaoze Liu, Jeff Z. Pan, Ningyu Zhang, and Huajun Chen. Knowledge graphs meet multi-modal learning: A comprehensive survey, 2024.

[16] Debanjan Chaudhuri, Md Rashad Al Hasan Rony, Simon Jordan, and Jens Lehmann. Using a kg-copy network for non-goal oriented dialogues, 2019.

[17] Vivek Khetan, Annervaz K M, Erin Wetherley, Elena Eneva, Shubhashis Sengupta, and Andrew E. Fano. Knowledge graph anchored information-extraction for domain-specific insights, 2021.

[18] Minze Chen, Zhenxiang Tao, Weitong Tang, Tingxin Qin, Rui Yang, and Chunli Zhu. Enhancing emergency decision-making with knowledge graphs and large language models, 2023.

[19] Yang Deng, Wenxuan Zhang, Yifei Yuan, and Wai Lam. Knowledge-enhanced mixed-initiative dialogue system for emotional support conversations, 2023.

[20] Jinyuan Fang, Zaiqiao Meng, and Craig Macdonald. Trace the evidence: Constructing knowledge-grounded reasoning chains for retrieval-augmented generation, 2024.

[21] Zhiyue Liu, Jiahai Wang, and Zhenghong Li. Topic-to-essay generation with comprehensive knowledge enhancement, 2021.

[22] Garima Agrawal, Tharindu Kumarage, Zeyad Alghamdi, and Huan Liu. Mindful-rag: A study of points of failure in retrieval augmented generation, 2024.

[23] Phillip Schneider and Florian Matthes. Conversational exploratory search of scholarly publications using knowledge graphs, 2024.

[24] Aman Kumar and Swathi Dinakaran. Textbook to triples: Creating knowledge graph in the form of triples from ai textbook, 2021.

[25] Sumit Bhatia, Purusharth Dwivedi, and Avneet Kaur. Tell me why is it so? explaining knowledge graph relationships by finding descriptive support passages, 2018.

[26] Danilo Dessì, Francesco Osborne, Diego Reforgiato Recupero, Davide Buscaldi, and Enrico Motta. Generating knowledge graphs by employing natural language processing and machine learning techniques within the scholarly domain, 2020.

[27] Laura Cabello, Carmen Martin-Turrero, Uchenna Akujuobi, Anders Søgaard, and Carlos Bobed. Meg: Medical knowledge-augmented large language models for question answering, 2024.

[28] Yong He, Cheng Wang, Shun Zhang, Nan Li, Zhaorong Li, and Zhenyu Zeng. Kg-mtt-bert: Knowledge graph enhanced bert for multi-type medical text classification, 2022.

[29] Ashwini Jaya Kumar, Sören Auer, Christoph Schmidt, and Joachim köhler. Towards a knowledge graph based speech interface, 2017.

[30] Md. Rezaul Karim, Hussain Ali, Prinon Das, Mohamed Abdelwaheb, and Stefan Decker. Question answering over biological knowledge graph via amazon alexa, 2022.

[31] Fengshi Jing and Qingpeng Zhang. Knowledge-enhanced attentive learning for answer selection in community question answering systems, 2019.

[32] Siyu Li, Lu Chen, Chenwei Song, and Xinyi Liu. Text classification based on knowledge graphs and improved attention mechanism, 2024.

[33] Hanning Gao, Lingfei Wu, Po Hu, Zhihua Wei, Fangli Xu, and Bo Long. Graph-augmented learning to rank for querying large-scale knowledge graph, 2022.

[34] Zhangchi Qiu, Ye Tao, Shirui Pan, and Alan Wee-Chung Liew. Knowledge graphs and pre-trained language models enhanced representation learning for conversational recommender systems, 2024.

[35] Donghan Yu, Chenguang Zhu, Yiming Yang, and Michael Zeng. Jaket: Joint pre-training of knowledge graph and language understanding, 2020.

[36] Vasile Ionut Remus Iga and Gheorghe Cosmin Silaghi. Assessing llms suitability for knowledge graph completion, 2024.

[37] Zongzong Wu, Fengxiao Tang, Ming Zhao, and Yufeng Li. Kgv: Integrating large language models with knowledge graphs for cyber threat intelligence credibility assessment, 2024.

[38] Bo Ni, Yu Wang, Lu Cheng, Erik Blasch, and Tyler Derr. Towards trustworthy knowledge graph reasoning: An uncertainty aware perspective, 2024.

[39] Lei Liang, Mengshu Sun, Zhengke Gui, Zhongshu Zhu, Zhouyu Jiang, Ling Zhong, Yuan Qu, Peilong Zhao, Zhongpu Bo, Jin Yang, Huaidong Xiong, Lin Yuan, Jun Xu, Zaoyang Wang, Zhiqiang Zhang, Wen Zhang, Huajun Chen, Wenguang Chen, and Jun Zhou. Kag: Boosting llms in professional domains via knowledge augmented generation, 2024.

[40] Diego Sanmartin. Kg-rag: Bridging the gap between knowledge and creativity, 2024.

[41] Xiang Chen, Ningyu Zhang, Lei Li, Shumin Deng, Chuanqi Tan, Changliang Xu, Fei Huang, Luo Si, and Huajun Chen. Hybrid transformer with multi-level fusion for multimodal knowledge graph completion, 2023.

[42] Bin He, Di Zhou, Jinghui Xiao, Xin jiang, Qun Liu, Nicholas Jing Yuan, and Tong Xu. Integrating graph contextualized knowledge into pre-trained language models, 2021.

[43] K M Annervaz, Somnath Basu Roy Chowdhury, and Ambedkar Dukkipati. Learning beyond datasets: Knowledge graph augmented neural networks for natural language processing, 2018.

[44] Nimesh Bhana and Terence L. van Zyl. Knowledge graph fusion for language model fine-tuning, 2022.

[45] Mufei Li, Siqi Miao, and Pan Li. Simple is effective: The roles of graphs and large language models in knowledge-graph-based retrieval-augmented generation, 2025.

[46] Yanbin Wei, Qiushi Huang, James T. Kwok, and Yu Zhang. Kicgpt: Large language model with knowledge in context for knowledge graph completion, 2024.

[47] Qi Liu, Dani Yogatama, and Phil Blunsom. Relational memory augmented language models, 2022.

[48] Changlong Yu, Hongming Zhang, Yangqiu Song, and Wilfred Ng. Cocolm: Complex commonsense enhanced language model with discourse relations, 2022.

[49] Sai Sathiesh Rajan, Ezekiel Soremekun, and Sudipta Chattopadhyay. Knowledge-based consistency testing of large language models, 2024.

[50] Hannah Sansford, Nicholas Richardson, Hermina Petric Maretic, and Juba Nait Saada. Grapheval: A knowledge-graph based llm hallucination evaluation framework, 2024.

[51] Yuan Sui, Yufei He, Nian Liu, Xiaoxin He, Kun Wang, and Bryan Hooi. Fidelis: Faithful reasoning in large language model for knowledge graph question answering, 2025.

[52] Jiapu Wang, Kai Sun, Linhao Luo, Wei Wei, Yongli Hu, Alan Wee-Chung Liew, Shirui Pan, and Baocai Yin. Large language models-guided dynamic adaptation for temporal knowledge graph reasoning, 2024.

[53] Ziije Zhong, Linqing Zhong, Zhaoze Sun, Qingyun Jin, Zengchang Qin, and Xiaofan Zhang. Synthet2c: Generating synthetic data for fine-tuning large language models on the text2cypher task, 2025.

[54] Avinash Anand, Mohit Gupta, Kritarth Prasad, Ujjwal Goel, Naman Lal, Astha Verma, and Rajiv Ratn Shah. Kg-ctg: Citation generation through knowledge graph-guided large language models, 2024.

[55] Hakan T. Otal, Stephen V. Faraone, and M. Abdullah Canbaz. A new perspective on adhd research: Knowledge graph construction with llms and network based insights, 2024.

[56] Tuan Bui, Oanh Tran, Phuong Nguyen, Bao Ho, Long Nguyen, Thang Bui, and Tho Quan. Cross-data knowledge graph construction for llm-enabled educational question-answering system: A case study at hcmut, 2024.

[57] Hasan Abu-Rasheed, Mohamad Hussam Abdulsalam, Christian Weber, and Madjid Fathi. Supporting student decisions on learning recommendations: An llm-based chatbot with knowledge graph contextualization for conversational explainability and mentoring, 2024.

[58] Dean Allemang and Juan Sequeda. Increasing the llm accuracy for question answering: Ontologies to the rescue!, 2024.

[59] Mehran Kazemi, Sid Mittal, and Deepak Ramachandran. Understanding finetuning for factual knowledge extraction from language models, 2023.

[60] Chu Fei Luo, Ahmad Ghawanmeh, Xiaodan Zhu, and Faiza Khan Khattak. Biaskg: Adversarial knowledge graphs to induce bias in large language models, 2024.

[61] Lars-Peter Meyer, Johannes Frey, Kurt Junghanns, Felix Brei, Kirill Bulert, Sabine Gründer-Fahrer, and Michael Martin. Developing a scalable benchmark for assessing large language models in knowledge graph engineering, 2023.

[62] Rosario Uceda-Sosa, Karthikeyan Natesan Ramamurthy, Maria Chang, and Moninder Singh. Reasoning about concepts with llms: Inconsistencies abound, 2024.

[63] Wenquan Wu, Zhen Guo, Xiangyang Zhou, Hua Wu, Xiyuan Zhang, Rongzhong Lian, and Haifeng Wang. Proactive human-machine conversation with explicit conversation goals, 2019.

[64] Debanjan Chaudhuri, Md Rashad Al Hasan Rony, and Jens Lehmann. Grounding dialogue systems via knowledge graph aware decoding with pre-trained transformers, 2021.

[65] Weijie Liu, Peng Zhou, Zhe Zhao, Zhiruo Wang, Qi Ju, Haotang Deng, and Ping Wang. K-bert: Enabling language representation with knowledge graph, 2019.

[66] Reham Omar, Omij Mangukiya, Panos Kalnis, and Essam Mansour. Chatgpt versus traditional question answering for knowledge graphs: Current status and future directions towards knowledge graph chatbots, 2023.

[67] Rishi Kalra, Zekun Wu, Ayesha Gulley, Airlie Hilliard, Xin Guan, Adriano Koshiyama, and Philip Treleaven. Hypa-rag: A hybrid parameter adaptive retrieval-augmented generation system for ai legal and policy applications, 2025.

[68] Dong Shu, Tianle Chen, Mingyu Jin, Chong Zhang, Mengnan Du, and Yongfeng Zhang. Knowledge graph large language model (kg-llm) for link prediction, 2024.

[69] Linyao Yang, Hongyang Chen, Zhao Li, Xiao Ding, and Xindong Wu. Give us the facts: Enhancing large language models with knowledge graphs for fact-aware language modeling, 2024.

[70] Yi-Lin Tuan, Sajjad Beygi, Maryam Fazel-Zarandi, Qiaozi Gao, Alessandra Cervone, and William Yang Wang. Towards large-scale interpretable knowledge graph reasoning for dialogue systems, 2022.

[71] Feifei Xu, Shanlin Zhou, Xinpeng Wang, Yunpu Ma, Wenkai Zhang, and Zhisong Li. Open-domain dialogue generation grounded with dynamic multi-form knowledge fusion, 2022.

[72] Tianxiang Sun, Yunfan Shao, Xipeng Qiu, Qipeng Guo, Yaru Hu, Xuanjing Huang, and Zheng Zhang. Colake: Contextualized language and knowledge embedding, 2020.

[73] Mengyuan Zhao, Xiaowen Huang, Lixi Zhu, Jitao Sang, and Jian Yu. Knowledge graph-enhanced sampling for conversational recommender system, 2021.

[74] Xuhui Ren, Tong Chen, Quoc Viet Hung Nguyen, Lizhen Cui, Zi Huang, and Hongzhi Yin. Explicit knowledge graph reasoning for conversational recommendation, 2023.

[75] Hongyin Luo, Shang-Wen Li, and James Glass. Knowledge grounded conversational symptom detection with graph memory networks, 2021.

[76] Ali Mousavi, Xin Zhan, He Bai, Peng Shi, Theo Rekatsinas, Benjamin Han, Yunyao Li, Jeff Pound, Josh Susskind, Natalie Schluter, Ihab Ilyas, and Navdeep Jaitly. Construction of paired knowledge graph-text datasets informed by cyclic evaluation, 2023.

[77] Senthil Mani, Neelamadhav Gantayat, Rahul Aralikatte, Monika Gupta, Sampath Dechu, Anush Sankaran, Shreya Khare, Barry Mitchell, Hemamalini Subramanian, and Hema Venkatarangan. Hi, how can i help you?: Automating enterprise it support help desks, 2017.

[78] George Arthur Baker, Ankush Raut, Sagi Shaier, Lawrence E Hunter, and Katharina von der Wense. Lost in the middle, and in-between: Enhancing language models' ability to reason over long contexts in multi-hop qa, 2024.

AI-generated, for reference only.

[79] Saeedeh Shekarpour, Abhishek Nadgeri, and Kuldeep Singh. Qa2explanation: Generating and evaluating explanations for question answering systems over knowledge graph, 2020.

[80] Dat Quoc Nguyen. A survey of embedding models of entities and relationships for knowledge graph completion, 2020.

[81] Genet Asefa Gesese, Russa Biswas, Mehwish Alam, and Harald Sack. A survey on knowledge graph embeddings with literals: Which model links better literal-ly?, 2020.

[82] Yu Wang, Nedim Lipka, Ryan A. Rossi, Alexa Siu, Ruiyi Zhang, and Tyler Derr. Knowledge graph prompting for multi-document question answering, 2023.

[83] Han Xiao, Minlie Huang, and Xiaoyan Zhu. Ksr: A semantic representation of knowledge graph within a novel unsupervised paradigm, 2020.

[84] Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. Comet: Commonsense transformers for automatic knowledge graph construction, 2019.

[85] Satvik Garg and Dwaipayan Roy. A birds eye view on knowledge graph embeddings, software libraries, applications and challenges, 2022.

[86] Alvin Dey, Harsh Kumar Jain, Vikash Kumar Pandey, and Tanmoy Chakraborty. All it takes is 20 questions!: A knowledge graph based approach, 2019.

[87] Nandana Mihindukulasooriya, Sanju Tiwari, Carlos F. Enguix, and Kusum Lata. Text2kgbench: A benchmark for ontology-driven knowledge graph generation from text, 2023.

[88] Ikhlas Alhussien, Erik Cambria, and Zhang NengSheng. Semantically enhanced models for commonsense knowledge acquisition, 2018.

[89] Wenchang Ma, Ryuichi Takanobu, and Minlie Huang. Cr-walker: Tree-structured graph reasoning and dialog acts for conversational recommendation, 2021.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.