Revista de Informática Teórica e Aplicada - RITA - ISSN 2175-2745
Vol. 33, Num. 3 (2026) 24-44

RESEARCH ARTICLE

A Bibliometric Analysis on the Integration of Large Language Models and Knowledge Graphs: Exploring Trends, Techniques, and Emerging Directions

Uma Análise Bibliométrica sobre a Integração de Grandes Modelos de Linguagem e Grafos de Conhecimento: Exploração de Tendências, Técnicas e Direções Emergentes

Wafa Ghemmaz [1]*, Maroua Bouzid [2], Souad Bouaicha [1], Naila Marir [3]

Abstract: In recent years, the integration of large language models and knowledge graphs has increasingly attracted the attention of both researchers and practitioners. While existing reviews provide valuable and helpful insights for better understanding the research field, they lack a quantitative perspective. To address this limitation, a bibliometric analysis was performed in this study to fill the gap. In particular, sourcing data from the Web of Science and Scopus databases, our analysis focuses on: (1) publication and citation trends, (2) the most productive countries and authors, (3) the most influential sources and scientific papers, (4) the main research tasks particularly in KG-enhanced LLMs and LLM-augmented KGs, (5) the most frequently employed techniques and their evolution over time. In addition to providing a comprehensive overview of the most active and recent areas of research, this study identifies several emerging themes, namely: "Knowledge-aware Prompt Engineering", "Retrieval-Augmented Generation for LLM–KG Integration", "Knowledge Graph-based Hallucination Detection" and the convergence of "Multimodal Knowledge Graphs and Multimodal Large Language Models". These insights contribute to a clearer understanding of the field's development and highlight emerging research trends grounded in bibliometric evidence.
Keywords: knowledge graphs — large language models — research trends — emerging topics — VOSviewer

Resumo: Nos últimos anos, a integração de grandes modelos de linguagem e grafos de conhecimento tem atraído cada vez mais a atenção de pesquisadores e profissionais. Embora as revisões existentes ofereçam percepções valiosas e úteis para uma melhor compreensão deste campo de pesquisa, elas carecem de uma perspectiva quantitativa. Para superar essa limitação, foi realizada, neste estudo, uma análise bibliométrica com o objetivo de preencher essa lacuna. Em particular, utilizando dados das bases Web of Science e Scopus, nossa análise concentra-se em: (1) tendências de publicação e citação, (2) países e autores mais produtivos, (3) fontes e artigos científicos mais influentes, (4) as principais tarefas de pesquisa, especialmente em LLMs aprimorados por KGs e KGs aumentados por LLMs, e (5) as técnicas mais frequentemente empregadas e sua evolução ao longo do tempo. Além de fornecer uma visão abrangente das áreas de pesquisa mais ativas e recentes, este estudo identifica vários temas emergentes, nomeadamente: "Engenharia de Prompts com Consciência de Conhecimento", "Geração Aumentada por Recuperação para Integração LLM–KG", "Deteção de Alucinações Baseada em Grafos de Conhecimento" e a convergência entre "Grafos de Conhecimento Multimodais e Grandes Modelos de Linguagem Multimodais". Estas percepções contribuem para uma compreensão mais clara da evolução do campo e destacam tendências de pesquisa emergentes fundamentadas em evidências bibliométricas.
Palavras-Chave: grafos de conhecimento — grandes modelos de linguagem — tendências de pesquisa — tópicos emergentes — VOSviewer

[1] *LIRE Laboratory, University of Constantine 2 - Abdelhamid Mehri (UC2), Constantine, Algeria*
[2] *GREYC Laboratory, University of Caen Lower Normandy (UNICAEN), Caen, France*
[3] *College of Engineering, Effat University (EU), Jeddah, Saudi Arabia*
***Wafa Ghemmaz**: wafa.ghemmaz@univ-constantine2.dz
DOI: http://dx.doi.org/10.22456/2175-2745.151034 • **Received:** 19/10/2025 • **Accepted:** 08/04/2026

1. Introduction

Large Language Models (LLMs), including encoder-based models like BERT [1] and RoBERTa [2], decoder-only models such as GPT [3], and encoder-decoder models like T5 [4], which are pre-trained on vast corpora; have achieved great potential in various Natural Language Processing (NLP) tasks including for example question answering (QA) [5] and text generation [6]. Due to the rapid expansion of model size, they have developed emerging abilities enabling them to address complex challenges [7]. Indeed, LLMs especially the advanced ones which are equipped with billions of parameters have demonstrated significant promise in handling complex tasks such as education assistance and recommendation.

Despite their advances and powerful capabilities, LLMs still struggle with important limitations. One of the major challenges is hallucination. LLMs are mainly based on facts and knowledge from data on which they are trained. However, they often generate responses that appear plausible but deviate from factual reality. In this context, several studies have been conducted to address this challenge [8, 9, 10]. Furthermore, the lack of transparency and interpretability is another major challenge, as LLMs are often criticized for their black-box nature [11]. In fact, their probabilistic reasoning is prone to hallucination issues. Consequently, their reliability becomes limited especially in high-stakes domains such as medicine and law.

To address these challenges, an effective strategy is the incorporation of Knowledge Graphs (KGs) into LLMs. KG is a graph-structured representation of human knowledge, where nodes denote entities and edges capture the relationships between them [12]. So, it stores facts as triples consisting of a head entity, a relation, and a tail entity. In fact, KGs play a pivotal role in many applications since they provide precise and explicit knowledge. Consequently; and by integrating KG in LLMs; LLMs can leverage a robust foundation of reliable and interpretable explicit knowledge [13]. Moreover, KGs are widely recognized for their reasoning capabilities [14].

On the other hand, the construction of KGs is a difficult task especially when using classical methods which are considered inappropriate and insufficient for managing the incomplete and dynamically evolving nature of real-world KGs. Consequently, LLMs have recently emerged as powerful technology to address the challenges faced by a variety of KG-related tasks such as KG construction, embedding and completion.

In recent years, unifying LLMs and KGs has increasingly attracted the attentions of researchers and practitioners where some works focus on KG-enhanced LLMs, others on LLM-augmented KGs, while other researches introduced a collaborative approach in which LLMs and KGs mutually reinforce each other.

Given the growing importance of unifying LLMs and KGs, several survey papers that analyze the methods, techniques and models employed in LLM-augmented KGs, KG-enhanced LLMs, and synergized LLM+KG, have been published. In [15]; one of the most cited papers in this field; the authors present a forward-looking roadmap for unifying both LLMs and KG. Moreover, they propose a detailed categorization, conduct comprehensive reviews, and pinpoint emerging directions in the field. Similarly, [16] presents a detailed categorization and novel taxonomies of research on unifying LLMs and KGs, and covers the advanced techniques. Additionally, it highlights the challenges in existing research and presents several promising future research directions. In [17], the common debate points within the community are summarized. Furthermore, the state-of-the-art for a comprehensive set of topics involving the integration of KGs and LLMs is introduced. The opportunities and challenges are also discussed in the paper.

While existing reviews provide valuable and helpful insights for better understanding the research field, none of them provide:

- A quantitative assessment of how research in this field has evolved over time.
- A comparative view of which tasks and techniques are most frequently adopted.
- Identification of the most influential authors, publications, and emerging research topics.

In order to address these limitations, this study provides the following main contributions:

- It presents the first bibliometric analysis of LLM–KG integration research, based on publications indexed in Web of Science Core Collection (WoS) and Scopus databases between 2020 and 2025.
- It shows how research has changed over time and highlights emerging trends including "Knowledge-aware Prompt Engineering", "Retrieval-Augmented Generation for LLM–KG Integration," "Knowledge Graph-based Hallucination Detection" and the convergence of "Multimodal Knowledge Graphs and Multimodal Large Language Models", that are likely to continue to attract attention in the incoming years.

This quantitative and synthesized study aims not only to capture the current research landscape but also to guide future efforts by identifying both saturated and under-explored directions.

The rest of the paper is organized as follows: Section 2 presents our research methodology including data collection and analyzing procedures. Section 3 reports the main results and discussion. Section 4 discusses the emerging themes identified from the bibliometric results. Finally, section 5 concludes this paper

2. Methodology

In this study, the analyzed scientific literature was retrieved from WoS and Scopus databases which are considered as the world's most leading citation databases and are extensively

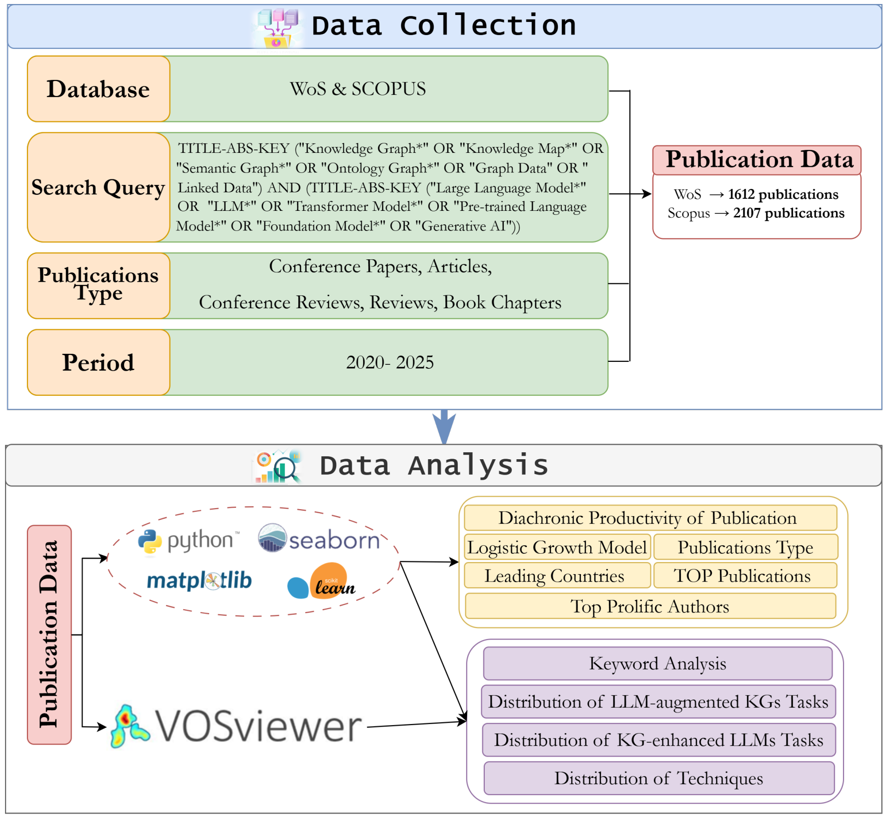

Figure 1. Bibliometric Analysis Process

employed in bibliometric analysis [18]. Furthermore, they provide comprehensive and high-quality bibliographic records of high-impact journals worldwide enabling the tracking of research progress and trends in combining KGs and LLMs.

The overall process of this review, encompassing two main steps, data collection with search terms, and bibliometric methodologies, is presented in Figure 1. The first step consists of collecting existing previous research works that combine LLMs and KGs. Then, the exported data are preprocessed where missing bibliometric information is eliminated. In the second step, we perform a bibliometric analysis by using several tools. The two main steps are detailly presented in the next sub-sections.

2.1 Data Collection

We searched the WoS and Scopus databases for documents on KGs and LLMs, focusing on titles, abstracts, and keywords. The search was conducted on December 1, 2025, using the following query:

TITLE-ABS-KEY (”Knowledge Graph” OR ”Knowledge Map*” OR ”Semantic Graph*” OR ”Ontology Graph*” OR ”Graph Data” OR ”Linked Data”) AND (TITLE-ABS-KEY (”Large Language Model*” OR ”LLM*” OR ”Transformer Model*” OR ”Pre-trained Language Model*” OR ”Foundation Model*” OR ”Generative AI”)).*

This search query was used in the WoS and Scopus database interfaces, focusing exclusively on relevant research published from 2020 to 2025, excluding documents as editorial, note, and letter. Note that the combined dataset includes a considerable number of publications dated 2026. However, we excluded these publications from the present analysis which focuses only on the 2020-2025 period. Particularly, we identified 196 publications dated 2026 that demonstrates the rapidly growing interest in research combining KGs and LLMs, even for articles that have not yet reached their official publication year. After merging, cleaning and de-duplication, the final corpus used for our analysis consists of 3719 publications (2107 from Scopus and 1612 from WoS).

2.2 Data Analysis

In this step, the retrieved data were further visualized and analyzed by using Python and a bibliometric mapping tool VOSviewer 1.6.19, as illustrated in Figure 1. Diverse python libraries are used such as Matplotlob, seaborn and Skitlearn.

The python libraries were used to fit and analyze the diachronic productivity of publication and citation, the Logistic Growth Model, Publications type, Leading Countries, Top publications and Top Prolific Authors. These libraries were also used with VOSviewer to construct and visualize the keywords co-occurrences and their connections which can effectively reflect the topic's prominence within the field, the distribution of LLM-augmented KGs tasks, KG-enhanced LLMs tasks and techniques.

3. Results and Discussions

3.1 Temporal Evolution of Publications

In order to prove the growing scientific interest and global activities in combining KGs and LLMs, we present the annual publication trend which is considered as a clear yet insightful indicator of this evolving research focus. As shown in Figure 2, the analysis of research shows a rapid and exponential increase over the past years. From 2020 to 2025, the publication increased from 37 to 1825 with a growth of 4832% over only five years. In 2020, only 37 publications were recorded. Then, this number rose to 51 in 2021 (+37%), and reached 103 in 2022 (+101%). The year 2023 recorded 374 publications (+263%). Between 2023 and 2024, the number of publications increased from 374 to 1329 (+255% in one year) representing the largest growth. Finally, for 2025, the number of publications has already reached 1825 (+37%), which clearly confirming the continuous acceleration of research dynamics and increasing relevance of KG and LLMs across several research topics.

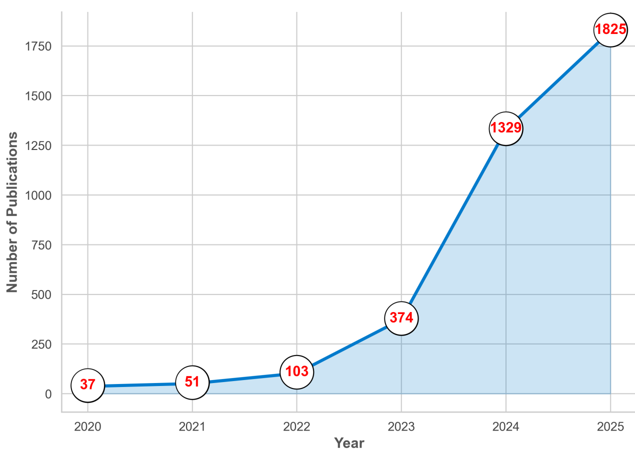

Figure 2. Annual Trend of Publications during 2020-2025

For observing trend in combining KGs and LLMs, and similar to [19, 20], we adopt a logistic growth model of the form of equation (1). The model is employed here as an indicative tool rather than a definitive forecasting method.

$$y = \frac{K}{1+e^{-r(t-t_0)}} \quad (1)$$

Where:

- K: maximum number of publications (carrying capacity)
- r = growth rate
- t_0: inflection point
- t: Year

The plot in Figure 3 illustrates the evolution of the number of publications between 2020 and 2030. Based on the Logistic Growth model, the evolution of combining KGs and LLMs may be divided into three main stages: (i) infant stage (2020–2022) which characterized by a moderate growth where the annual publication number increased slowly, (ii) growth stage (2022–2026) in which the number of publications could rise, and (iii) a mature stage (beyond 2026) where the growth tends to stabilize as it approaches the carrying capacity.

It should be emphasized that this domain remains highly dynamic, and the predicted year of the growth and maturity phases may evolve in response to new technological developments.

3.2 Type of Publication

As shown in Figure 4, the results of publication type distribution indicate that Conference Papers dominate the dataset, accounting for 62.1% of the 3719 extracted papers, followed by Articles (28.8%) and Conference Reviews (6.5%). Reviews and Book Chapters represent a minor fraction of the total publications (1.9% and 0.7% respectively). This distribution analysis highlights the predominant role of conference proceedings in disseminating research findings compared to other publication categories.

3.3 Leading Countries

In our analysis of scientific productivity in the field of KG and LLMs, we notice that a strong research effort is revealed, with notable contributions from a few leading countries. According to the first author of the retrieved articles and as shown in Figure 5, the most prolific contributor is China, with 1531 publications and 7154 citations, demonstrating its dominant role in advancing research related to KG and LLMs. The United States follows with 685 publications and 5621 citations. Germany ranks third with 224 publications and 636 citations. India, Italy and the United Kingdom produced between 107 and 118 publications where the citations impact is varied (383, 236 and 412 respectively). Additionally, Japan, France, South Korea, and Canada each published between 48 and 61 publications.

The results of our analysis indicate that the leading countries are Asia, North America, and Europe. We notice that China is leading in both publication volume and citation impact. In fact, it accounts with the United States more than 74% of global output.

3.4 Top Sources and Most cited Publications

As discussed above, conference papers and journal articles constitute the dominant publication types in our dataset. Therefore, the Top 10 conferences and Top 10 journals published studies related to the integration of LLMs and KGs are presented in Figure 6 and 7 respectively. For conferences, in terms of publications and citations, the most active source was The Annual Meeting of the Association for Computational Linguistics (ACL) with 122 publications and 1284 citations, reflecting (i) its pivotal role in driving scientific production and knowledge exchange in the field, and (ii) its strong influence in the NLP community. The International Semantic Web Conference (ISWC) conference ranks second with 63

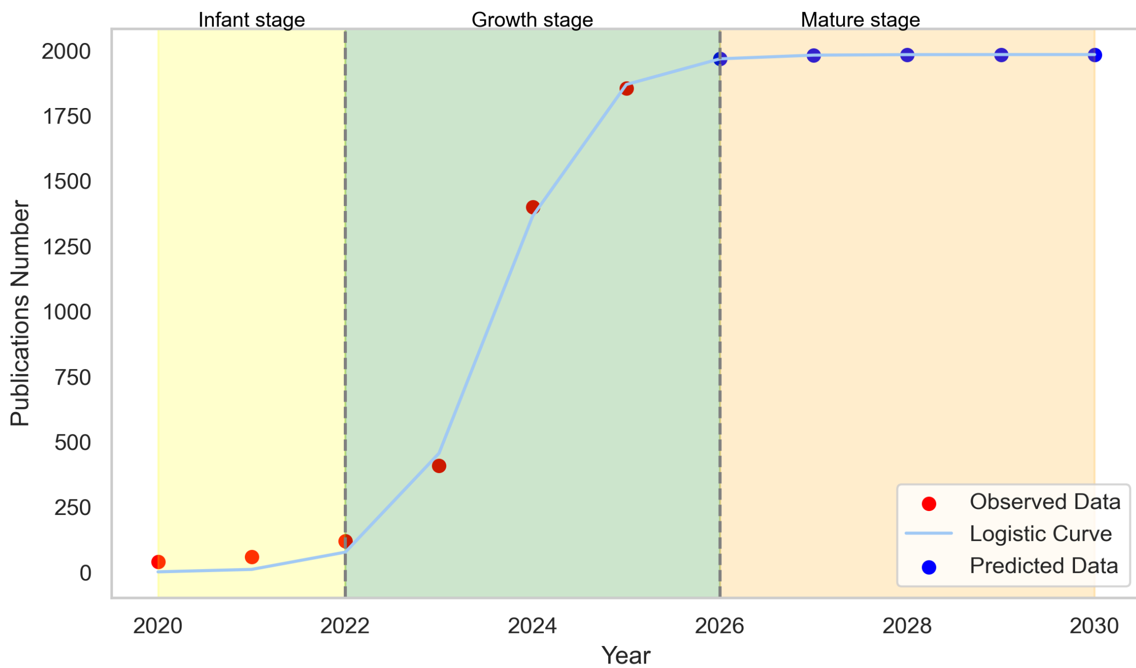

Figure 3. Publications Growth Curve using Logistic Model (2020-2030)

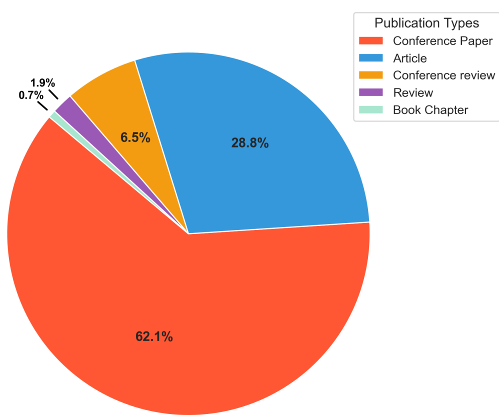

Figure 4. Annual Trend of Publications during 2020-2025

publications and 75 citations. In contrast, the Annual Meeting of the Association for Computational Linguistics (AAAI) received a remarkably citation count (481 citations for 62 publications) witch indicates that its contributions are highly influential in the field. Notably, The Annual Conference on Neural Information Processing Systems (NeurIPS) exhibits a distinct profile. it achieves a relatively high citation count (395) despite a lower number of publications (34). This suggests a higher average citation impact per paper, reflecting consequently the strong visibility and the influence of research presented at this venue. Other conferences such as The ACM Web Conference (ACM WebConf), The ACM International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD), The ACM International Conference on Information and Knowledge Management (ACM CIKM), The International Conference on Computational Linguistics (COLING) and International Joint Conference on Artificial Intelligence (IJCAI) show moderate publication volumes. Note that the full descriptions of each conference acronym used in Figure 6 are provided in Appendix 1.

In addition to conference venues, journals play a crucial role in the dissemination of scientific knowledge. As illustrated in Figure 7, the ”IEEE Access” has the highest number of publications (52) with 123 citations, followed by ”Knowledge-based Systems” (37 publications, 178 citations). ”Applied Sciences-Basel” and ”Electronics” also contribute significantly with 35 and 31 publications respectively (106 and 66 citations respectively). The ”Information Processing and Management” journal and ”Expert Systems with Applications” received 101 and 161 citations for 25 and 21 publications respectively.

”IEEE Transactions on Knowledge and Data Engineering” stands out for its exceptionally high citation count (820) despite a moderate publication volume (19). This indicates that articles published in this venue have a high scientific impact. Similarly, ”Advanced Engineering Informatics” showed a high citation impact relative to its number of publications. Other journals, including ”NeuroComputing” and ”Scientific Reports” , exhibit lower publication and citation numbers, reflecting either reduced research output in the field .

As shown in table 1, the top 10 most cited papers having

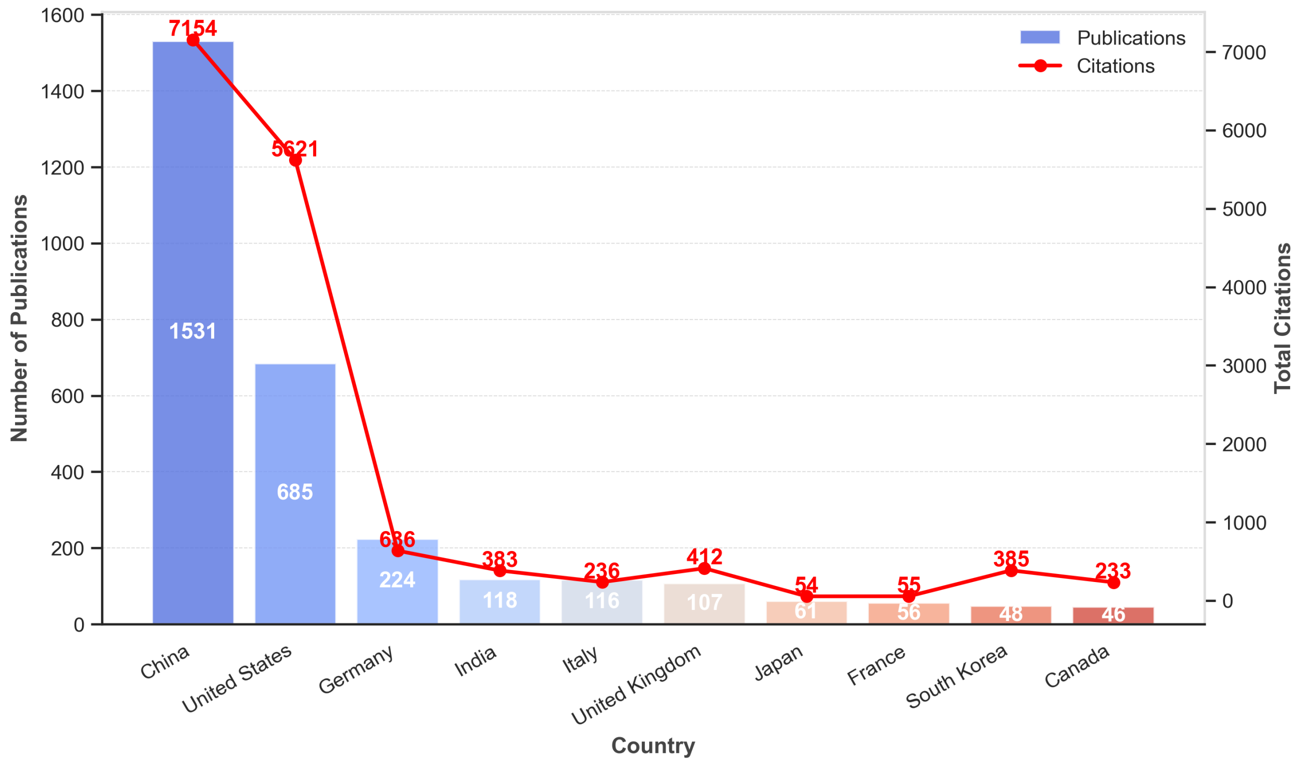

Figure 5. Number of Documents for the Top 10 Productive Countries during 2020-2025

more than 100 citations among the collected publications show strong academic influence, with citations ranging from 129 to 831 (3359 citations in total) which confirm the growing interest in combining LLMs and KGs.

The article "ERNIE: Enhanced Language Representation with Informative Entities" [21] was the most cited article (931 citations since 2020) followed by [22] that has reached 626 citations since 2020. [15] ranks third with 475 citations only since 2024. [23] published in 2021 have also received important citation (396). The high citation counts of these publications underscore the increasing importance of combining LLMs and KGs. In this case, this combination can be considered a prominent and emerging trend in current AI research.

After analyzing, these highly cited articles could be classified into the following three categories [15]: (1) KG-enhanced LLMs, (2) LLM-augmented KGs, and (3) Synergized LLM + KG.

KG-enhanced LLMs: By embedding structured knowledge into vector representations, KGs are integrated into LLMs to improve language comprehension and performance. Furthermore, this integration is used to address diverse problems like hallucinations and lack of interpretability. In this context, several models like ERNIE [21], CoLAKE [25], DRAGON [26], and KGLM [29] have been proposed.

In [21], an enhanced language representation model named ERNIE (Enhanced Representation through kNowledge IntEgration) is introduced in order to improve language understanding by integrating KGs with large-scale textual corpora. In [25], a Contextualized Language and Knowledge Embedding (CoLAKE), which jointly learns contextualized representation for both language and knowledge by pre-training on structured, unlabeled word-knowledge graphs (WK graphs). DRAGON (Deep bidiRectional LAnGuage-KnOwledge graph PretraiNing) [26] is an approach that performs deeply bidirectional, self-supervised pretraining of a language-knowledge model from text and KG. It has two main components: a cross-modal model that bidirectionally fuses text and KG, and a bidirectional self-supervised objective that learns joint reasoning over text and KG.

LLM-augmented KGs: By using LLMs generalization capabilities, this combination enhances various KG-related tasks including for example KG construction, completion, embedding, and QA. It involves using LLMs to process text, enrich graph representation, as well as generate new facts. Models like COMET [22] and SimKGC [24] have been introduced and demonstrated how LLMs can be exploited to enhance KG construction and completion.

COMET [22] is a COMmonsEnse Transformer that allows the construction of commonsense KBs by using existing tuples as a seed set of knowledge on which to train. By using this initial seed set, the PLM learns to adapt its learned representations to the knowledge generation task and produces novel and semantically accurate knowledge triples. SimKGC [24] ad-

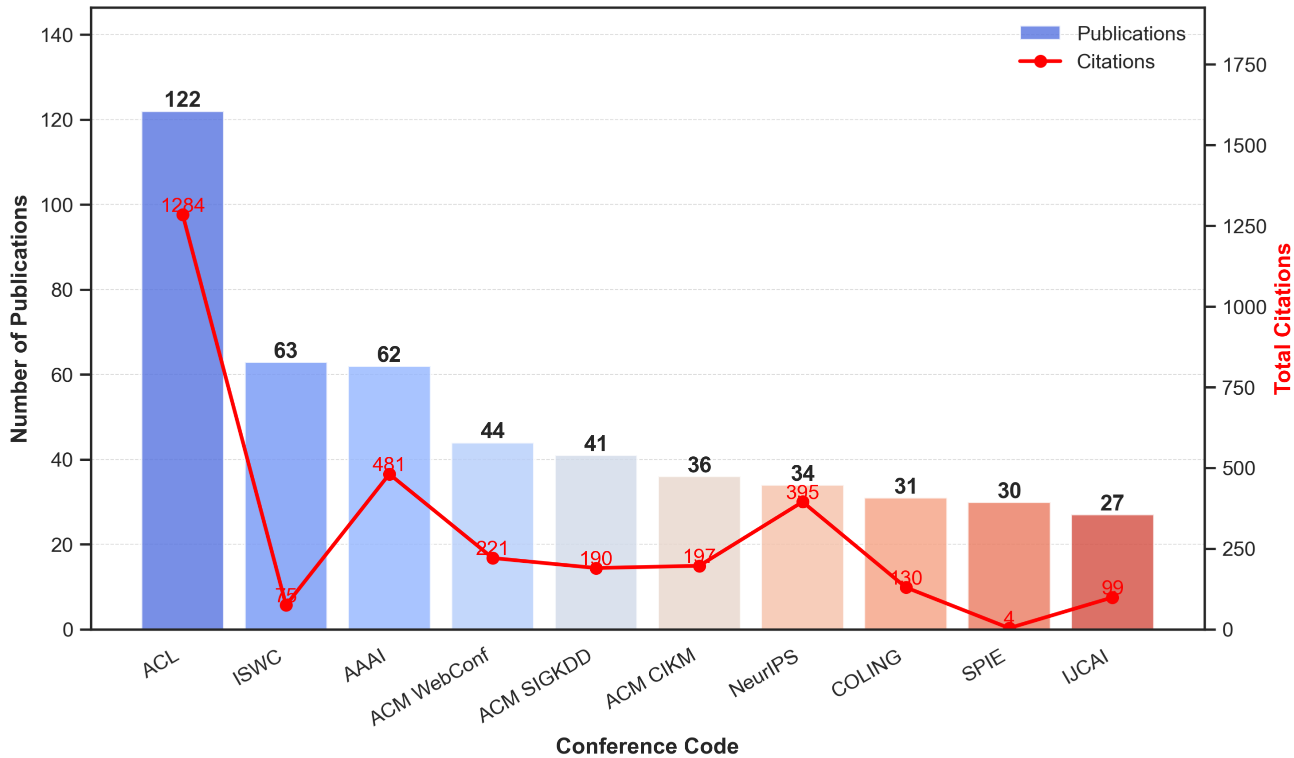

Figure 6. Top 10 Conferences during 2020-2025

vances contrastive learning for Knowledge Graph Completion (KGC) using PLMs. It introduces novel negative sampling strategies (in-batch negatives, pre-batch negatives, and self-negatives) in order to improve inductive KGC performance.

Synergized LLM + KG: Because LLMs and KG are two inherently complementary techniques, they are unified into a general framework to mutually enhance each other's capabilities. In recent years, achieving complex QA in open-domain and long-context settings has increasingly attracted researchers' interest. In [23], a novel model named QA-GNN which exploits LLMs and KGs is proposed. In this work, the authors use LLMs to estimate the importance of KG nodes relative to the given QA context. Then, they connect the QA context and KG to form a joint graph, and mutually update their representations through Graph Neural Networks (GNNs). In [27], a Knowledge Reasoning with Implicit and Symbolic rePresentations (KRISP) model allowing the combination of the implicit and symbolic knowledge is proposed. Firstly, it incorporates explicit reasoning in a KG with implicit reasoning in a multi-modal transformer. Then, it uses a graph network to make use of symbolic knowledge bases.

3.5 Top Prolific Authors

Based on the retrieved WoS and Scopus datasets, Table 2 presents the most prolific authors who appear in the first or second author positions only, in order to focus on primary contributors and to avoid inflated publication counts resulting from repeated co-authorship. From these listed authors, four are from Chinese institutions while three are affiliated with German institutions which prove the strong research presence of these two countries in this domain.

As researcher in semantic web technologies, Linked Data, KGs, and NLP, Mihindukulasooriya from IBM Research in New York, leads the list with 8 contributions [30, 31, 32, 33, 34, 35]. In the second position, four researchers share 7 publications: Buehler Markus j, Chen Zhuo, Li Xin and Frey Johannes. Buehler Markus j, affiliated with MIT School of Engineering in United States, focuses on artificial intelligence and machine learning, atomistic, molecular and multiscale modeling, Materials science and biomaterials [36, 37, 38, 37, 39, 40, 41]. Chen Zhuo from Zhejiang University, China, specializes in KG, Multi-Modal Learning and LLMs [42, 43, 44, 45, 46, 47, 48]. Li Xin affiliated with Beijing Institute of Technology, China, conducts research covering Machine learning, deep (reinforcement) learning and representation Learning [49, 50, 51, 52, 53]. Frey Johannes, from the University of Leipzig, Germany, is particularly interested in research topics related to the scope of the semantic web including Data integration and knowledge fusion; (Meta) data management in KGs [54, 55, 56, 57, 58, 59, 60].

In the third position, Perevalov Aleksandr [61, 62, 63, 64, 65, 66] and Hertling Sven [67, 68, 69, 32, 70] from Germany. Zhang Yu [71, 72, 73, 74, 75, 76] and Liu Ben [77, 78, 79, 80, 81, 82] from China, Luo Linhao [15, 83, 84, 85, 86, 87] from Australia, have actively participated with 6 contributions.

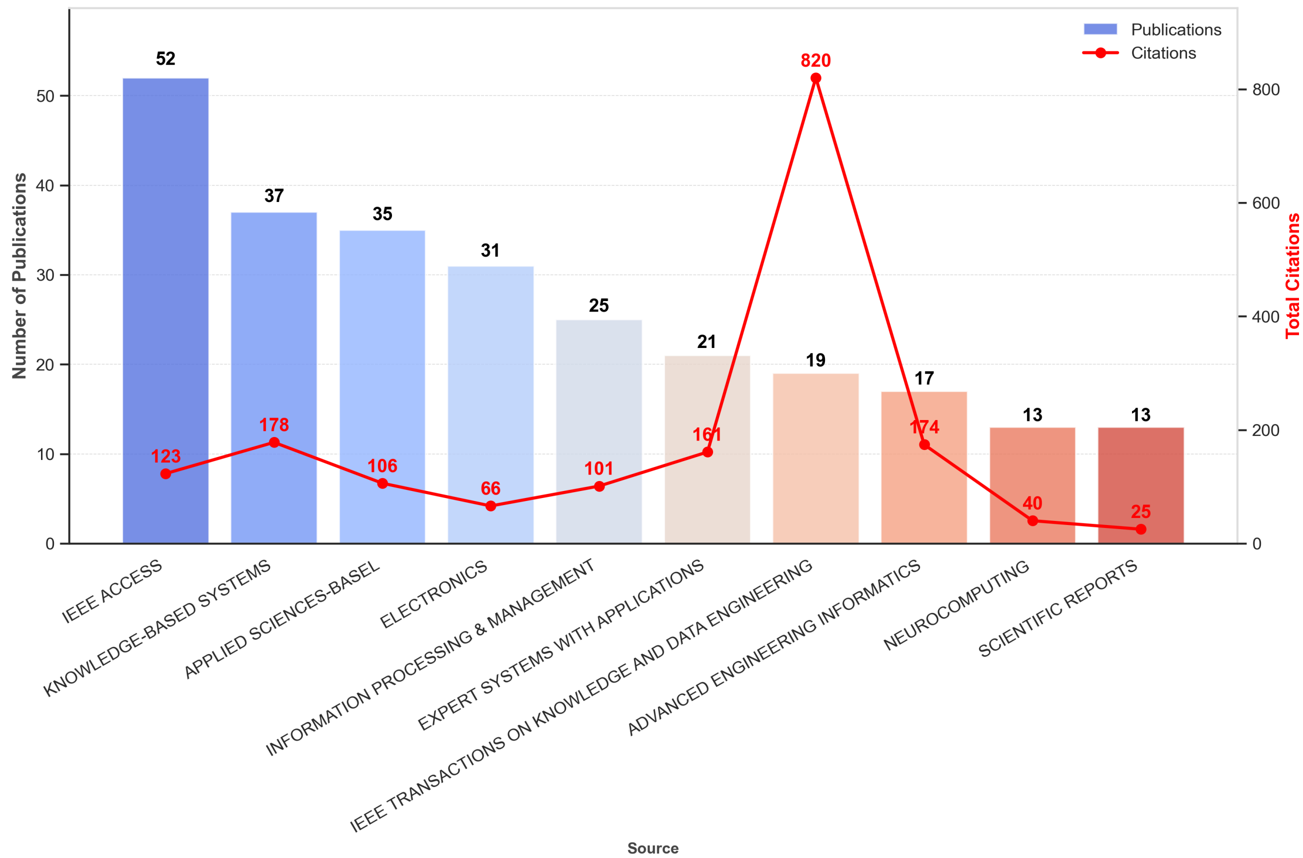

Figure 7. Top 10 Journals during 2020-2025

3.6 Thematic Analysis of Keyword Clusters

In this study, the co-occurrence analysis of keywords was performed using author keywords, leveraging the frequency and relationships among them; utilizing VOSviewer software for visualization and analysis.

Before starting this analysis, the collected bibliometric datasets were preprocessed in order to achieve better accuracy. A keyword normalization process is applied to reduce redundancy, for example, different forms of large language model (such as ”large language models (LLMs)”, ”LLMs”, and ”LLM” were replaced by a unified term: ”LLM”. The minimum threshold for keyword occurrence is 15, resulting 66 most frequent author keywords were used to generate our visualization. The size of each node reflects the occurrences of the keyword. In other words, higher frequency of occurrence yields large nodes while lower frequency results in smaller ones [88]. As depicted in Figure 8, among the most frequently occurring keywords were: KG (1504), LLM (1676), RAG (267), NLP (229), PLM (128), QA (125), GNN (122), DL (102), Completion (95), Ontology (93), Transformer (90).

Furthermore, links between nodes reflect co-occurrence relationships, with thicker lines denoting more frequent associations. Take the node KG as an example, it has a stronger relation with the node LLM, within a link strength (LS) of 972. It has also notable relation with NLP node (LS = 135). This suggests that these sets of keywords might be inclined to appear in the same articles.

In this analysis, the co-occurrence of keywords is supposed to represent the searching themes. As illustrated in Figure 7, the most frequent occurrence of authors keywords mainly formed 4 clusters (listed in Table 3) where the keywords with similarity in research topics are grouped together. The 4 clusters were analyzed as follows:

- Cluster 1 (red colored) includes key terms that are frequently used in studies focusing on knowledge access and reasoning powered by LLMs which are integrated with external resources through Retrieval-Augmented Generation (RAG), Graph RAG, KG-QA in order to improve factual accuracy and contextual understanding. It contains also terms that are usually used in studies of QA systems powered by LLMs like GPT and conversational tools such as ChatGPT. In fact, the strong occurrence of LLMs, GPT and ChatGPT and GenAI (Generative AI) highlights the central role of generative models in QA tasks. Advanced adaptation strategies such as Prompt Engineering are also presented in the cluster. Furthermore, the growing interest in making LLM-based reasoning over KG more transparent and understandable is affirmed by the inclusion of Explain-

Table 1. TOP 10 Cited publications

Rank	Publication Title	Citation	Year	Category
1	ERNIE: Enhanced language representation with informative entities [21]	931	2020	KG-enhanced LLMs
2	CoMET: Commonsense transformers for automatic knowledge graph construction [22]	626	2020	LLM-augmented KGs
3	Unifying Large Language Models and Knowledge Graphs: A Roadmap [15]	475	2024	Review Article
4	QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering [23]	396	2021	Synergized LLMs + KGs
5	SimKGC: Simple Contrastive Knowledge Graph Completion with Pre-trained Language Models [24]	191	2022	LLM-augmented KGs
6	CoLAKE: Contextualized Language and Knowledge Embedding [25]	158	2020	KG-enhanced LLMs
7	Deep Bidirectional Language-Knowledge Graph Pretraining [26]	155	2022	KG-enhanced LLMs
8	Krisp: Integrating implicit and symbolic knowledge for open-domain knowledge-based VQA [27]	151	2021	Synergized LLMs + KGs
9	Scalable multi-hop relational reasoning for knowledge-aware question answering [28]	147	2020	LLM-augmented KGs
10	Barack's wife Hillary: Using knowledge graphs for fact-aware language modeling [29]	129	2020	KG-enhanced LLMs

able AI (XAI). Moreover, the cluster includes concepts such as Ontology, RDF (Resource Description Framework), SPARQL (SPARQL Protocol and RDF Query Language), Linked Data, the Semantic Web, Knowledge Base, and Wikidata which can be considered as structured and symbolic semantic resources for enhancing LLMs reasoning capabilities.

- Cluster 2 (green colored) groups key terms related to the topic of knowledge extraction and enhancement from textual data using KGs dans deep learning techniques. It focuses on the integration of NLP methods such as ”Named Entity Recognition” (NER) and ”Relation Extraction”, deep learning architectures such as Transformers along with pre-trained models especially BERT; in order to extract, structure and enrich knowledge. For improving the accuracy and scalability of knowledge extraction tasks, advanced AI paradigms such as ”Attention Mechanism”, ”Prompt Learning” and ”Fine-Tuning” are also presented in the cluster. Furthermore, the cluster highlights several applications including ”Knowledge Engineering”, ”Knowledge Enhancement”, ”Natural Language Generation” (NLG), and ”Cybersecurity”. This shows how structured knowledge representations support reasoning and practical AI-driven solutions.
- Cluster 3 (blue colored) groups key terms related to the topic of representation learning and reasoning over KGs using Pre-trained Language Models (PLMs) and deep learning architectures such as GNNs; in order to encode semantic and structural information into semantic embeddings where techniques such as KG embedding is employed. These embeddings are then leveraged for KG reasoning tasks such as link prediction, KG completion, temporal kg modeling and entity alignment. The cluster also reflects the integration of advanced training paradigms; including contrastive learning to enhance the quality of semantic representations and reasoning capabilities. Furthermore, the presence of the term “Explainability” indicates an important research interest in making knowledge representation and reasoning models more interpretable and transparent which consequently reflecting the ongoing efforts to enhance the explainability of KG-based AI systems.
- Cluster 4 (yellow colored) includes key terms that focus on information extraction and retrieval from text in order to construct KGs. It also includes important techniques namely ”In-Context Learning” and ”zero-shot learning”, which help adapt language models to different tasks without retraining. Furthermore, the presence of ”Recommendation Systems” and ”Chatbots” shows how KGs can be used in real-world applications to deliver more relevant and personalized recommendations as well as to support intelligent conversational systems grounded in structured knowledge. Furthermore, the presence of the term ”Adaptation Models” highlights the efforts to improve system flexibility and responsiveness.

Table 2. TOP 10 Prolific Authors

Rank	Author	Institution	Publications
1	Mihindukulasooriya, Nandana	IBM Research, New York, United States	8
2	buehler, markus j	MIT School of Engineering, Cambridge, MA, United States	7
3	chen, zhuo	College of Computer Science and Technology, Zhejiang University, Hangzhou, Zhejiang, China	7
4	li, xin	School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China	7
5	frey, johannes	University of Leipzig, InfAI, Leipzig, Germany	7
6	perevalov, aleksandr	Leipzig University of Applied Sciences, Leipzig, Germany	6
7	hertling, sven	Data and Web Science Group, Universität Mannheim, Mannheim, Baden-Wurttemberg, Germany	6
8	zhang, yu	College of Intelligence Science and Technology, National University of Defense Technology China, Changsha, Hunan, China	6
9	liu, ben	School of Computer Science, Wuhan University, Wuhan, Hubei, China	6
10	luo, linhao	Department of Data Science and AI, Monash University, Melbourne, VIC, Australia	6

3.7 Analysis of LLM-Augmented KG Tasks

According to [15], we illustrated in Figure 9 the main KG-based tasks enhanced by using LLMs:

1. KG embedding in which LLMs are applied to enrich representations of KGs by encoding the textual descriptions of entities and relations.

2. KG completion by utilizing LLMs to encode text or generate facts

3. KG construction by applying LLMs to address the entity discovery, coreference resolution, and relation extraction tasks for KG construction.

4. LLM-augmented KG-to-text Generation by utilizing LLMs to generate natural language that describes the facts from KGs.

5. LLM-augmented KG QA by applying LLMs to bridge the gap between natural language questions and retrieve answers from KGs.

Our analysis reveals that KG-based tasks have attracted increasing attention from the research community, showing a clear upward trend in publication frequency from 2020 to 2025.

We can see from 9 that ”KG-to-text generation” has emerged as the most widely discussed research topic with an impressive increase from 4 occurrences in 2020 to a peak of 215 in 2025. This trend confirms the researcher’s interest in generating natural language from structured knowledge bases. Furthermore, ”KG-QA” task shows consistent growth, rising from 1 in 2020 to 153 in 2025. This highlights the efforts of the research community in leveraging KGs in open-domain and long-context QA systems. ”KG completion” task demonstrates also a gradual but consistent increase; reaching 62 in 2024 and 2025; which indicates an ongoing relevance. the ”KG construction” task shows a clear acceleration especially after 2023 where the frequency raises to 30 in 2024 and further increases to 55 in 2025. On other hand, the ”KG embedding” task have seen slow growth, increasing from 3 occurrences in 2022 to 22 in 2024 before slightly stabilizing at 19 in 2025.

After analyzing these results, we note that the focus has shifted toward more application-oriented research areas that synergize with the LLMs, rather than the traditional foundational KG tasks which represent; in our opinion, fertile ground for future research contributions. They present significant opportunities for further exploration and scientific contribution.

3.8 Analysis of KG-Enhanced LLM Tasks

According to [15], the performance and interpretability of LLMs can be enhanced by integrating KGs in different tasks. These tasks are categorized in three main groups:

1. KG-enhanced LLM pre-training in which KGs are applied during the pre-training stage and improve the knowledge expression of LLMs.

2. KG-enhanced LLM inference in which KGs are utilized during the inference stage of LLMs (LLMs will be enabled to access the latest knowledge without retraining).

3. KG-enhanced LLM interpretability in which the KGs are used to understand the knowledge learned by LLMs and interpret the reasoning process of LLMs.

In our analysis (see Figure 10), and as expected, KG-enhanced LLMs have also witnessed a remarkable and rapid development in the last years, especially from 2022 to 2025. What is clear and obvious from the plot is that the dominant category is KG-enhanced LLM inference with a dramatic

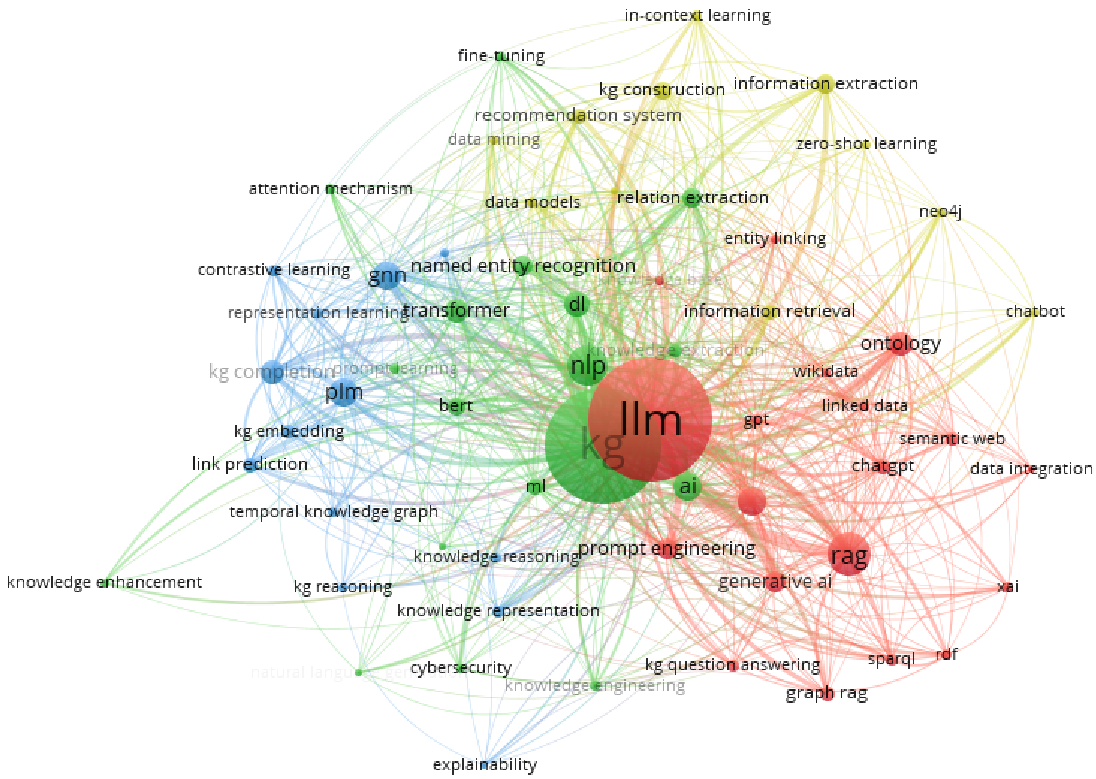

Figure 8. Network Visualization of the Author Keywords co-occurrences

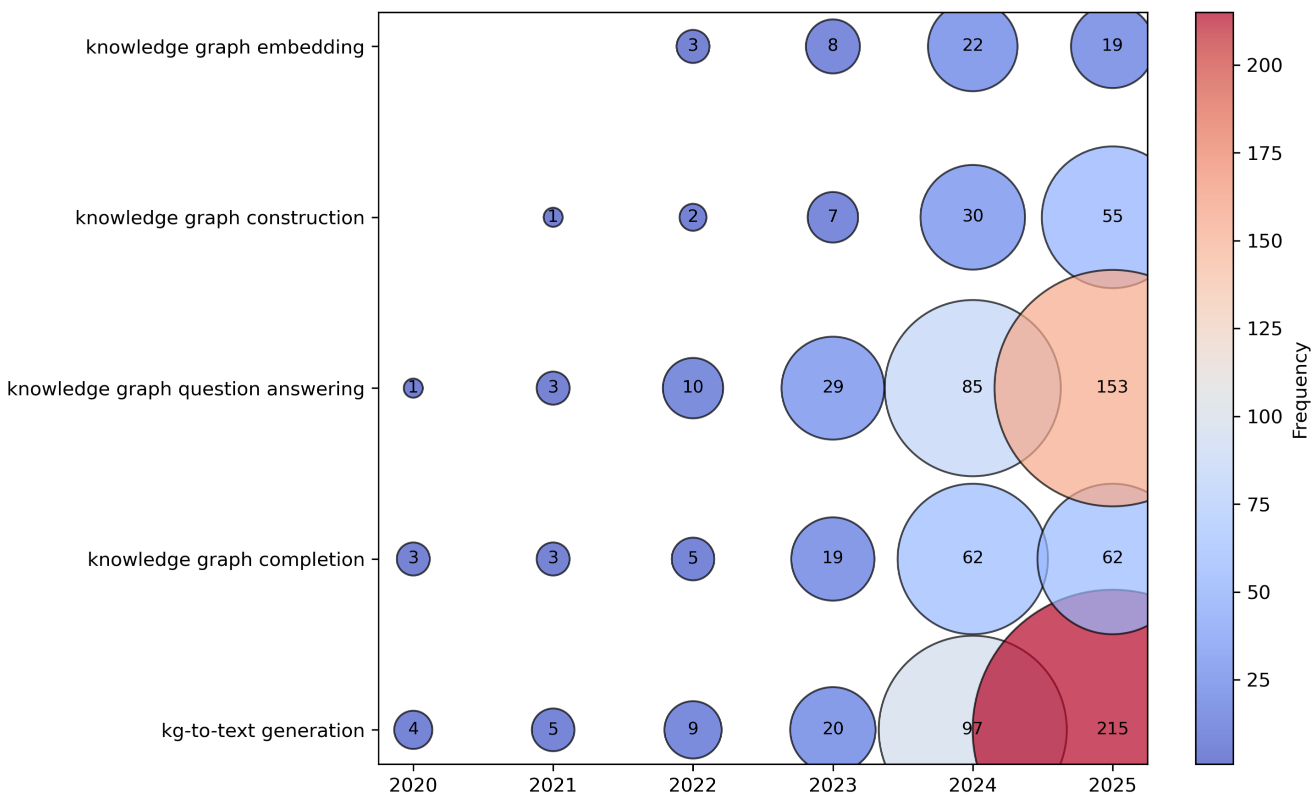

Figure 9. Distribution of KG-based tasks by year

Table 3. Clusters of keywords

Cluster	Keywords (TLS, Occurrences)	Interpretation
1: Red	LLM (2481, 1676), RAG (520, 261), QA (293,125), Ontology (227, 93), GenAI (132, 77), Prompt Engineering (188, 76), Chatgpt (87, 38), Graph RAG (83, 38), Semantic Web (74, 32), SPARQL (87, 30), KG_QA (55, 29), RDF (65, 23), GPT (53, 22), XAI (46, 22), Entity Linking (48, 20), Linked Data (38, 19), Knowledge Base (36, 17), Wikidata (43, 17), Data Integration (32, 15).	LLM-driven Knowledge Access and Retrieval-Augmented Question Answering.
2: Green	KG (2423, 1504), NLP (519, 229), AI (271, 128), DL (220, 102), Transformer (157, 90), Named Entity Recognition (134, 67), Relation Extraction (128, 62), Bert (125, 53), ML (116, 49), Knowledge Extraction (90, 37), Attention Mechanism (29, 24), Fine-tuning (50, 24), Knowledge Engineering (46, 22), Prompt Learning (29, 21), Cybersecurity (40, 17), Knowledge Enhancement (27, 17), Natural Language Generation (21, 15), RL (29, 15).	Integrating KGs, NLP, and Transformer-based Deep Learning for Knowledge Extraction and Enhancement.
3: Blue	PLM (156, 128), GNN (188, 122), KG Completion (156, 95), Link Prediction (86, 43), KG Embedding (62, 33), Contrastive Learning (55, 31), Knowledge Representation (50, 27), Representation Learning (44, 21), Temporal Knowledge Graph (38, 21), KG Reasoning (28, 20), Entity Alignment (43, 19), Knowledge Reasoning (29, 17), Explainability (22, 15).	Deep Representation Learning and KG Reasoning using PLMs.
4: Yellow	Information Extraction (121, 63), KG Construction (92, 54), Recommendation System (87, 41), Information Retrieval (85, 39), In-Context Learning (51, 24), Data Models (73, 23), Neo4j (55, 20), Data Mining (51, 19), Chatbot (50, 18), Zero-shot Learning (22, 16), Adaptation Models (44, 15).	Graph-based Recommendation and Adaptive Information Retrieval

surge: from only 4 mentions in 2020 to an impressive 463 occurrences in 2025. This remarkable growth proves the increasing importance of methods such as RAG, especially for QA tasks, and KGs prompting, which help LLMs to perform better.

On the other hand, KG-enhanced LLM pre-training has also experienced a robust upward trajectory (from 2 in 2020 to 90 in 2025) which confirms that the researchers are actively focused on how incorporate structured knowledge during the model training phase where the major aim is improving LLM robustness and generalization. Several techniques are used in this phase such as contrastive learning, graph embeddings and KG instruction tuning.

In contrast, KG-enhanced LLM interpretability has started to gain attention (from 2 occurrences in 2022 to 29 in 2025), even though it shows modest numbers compared to the two other categories. In fact, the emergence of KG-enhanced LLM interpretability reflects an excessing awareness of the need to explain and interpret LLM decisions, particularly in knowledge-intensive tasks where transparency and trustworthiness are critical.

3.9 Diachronic Evolution of Techniques for LLM-KG integration

The most techniques, methods or architectures that have been used in integrating LLMs and KGs to further enhance the performance are illustrated in Figure 11. Among these techniques, RAG [89] stands out as the most dominant, starting from only 1 occurrence in 2022 and reaching 192 occurrences in 2025. In this case, we can say that RAG is the most popular strategy for injecting structured knowledge into LLMs. This evolution is also tightly connected to the rapid adoption of Transformer architectures, which increased from 8 occurrences in 2020 to 125 in 2025. In this regard, we can say that Transformers not only serve as the backbone of modern LLMs but also enable advanced integration strategies with KGs.

Another technique has demonstrated strong momentum is Prompt Learning or knowledge-aware prompting (from only 2 occurrences in 2021 to 88 in 2025). On the other hand, the use of GNNs with KGs exhibit a notable progression (from 2 occurrences in 2021 to 81 in 2025). In this context, we can say that the researchers are interested in graph-based architectures for improving LLM reasoning. Similarly, the interest growth of Graph representation learning from 1 in 2020 to 77 in 2025 confirms that graph embeddings became essential for knowledge injection.

4. Emerging Research Themes

In this study, we analyzed the co-occurrence and the connections between author keywords of the selected publications. In addition, we examined the overlay visualization to identify the most recent and trending research themes (see Figure 12). This analysis not only provided us an overview about newest

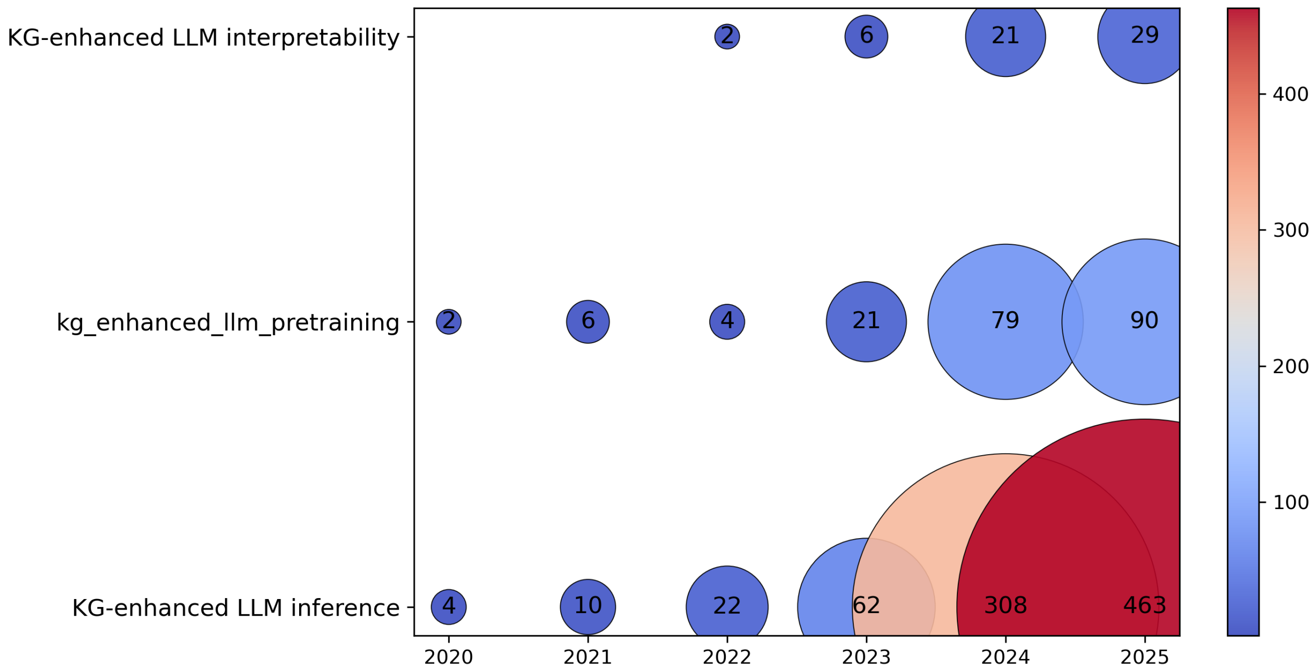

Figure 10. Distribution of KG-enhanced LLM tasks by year

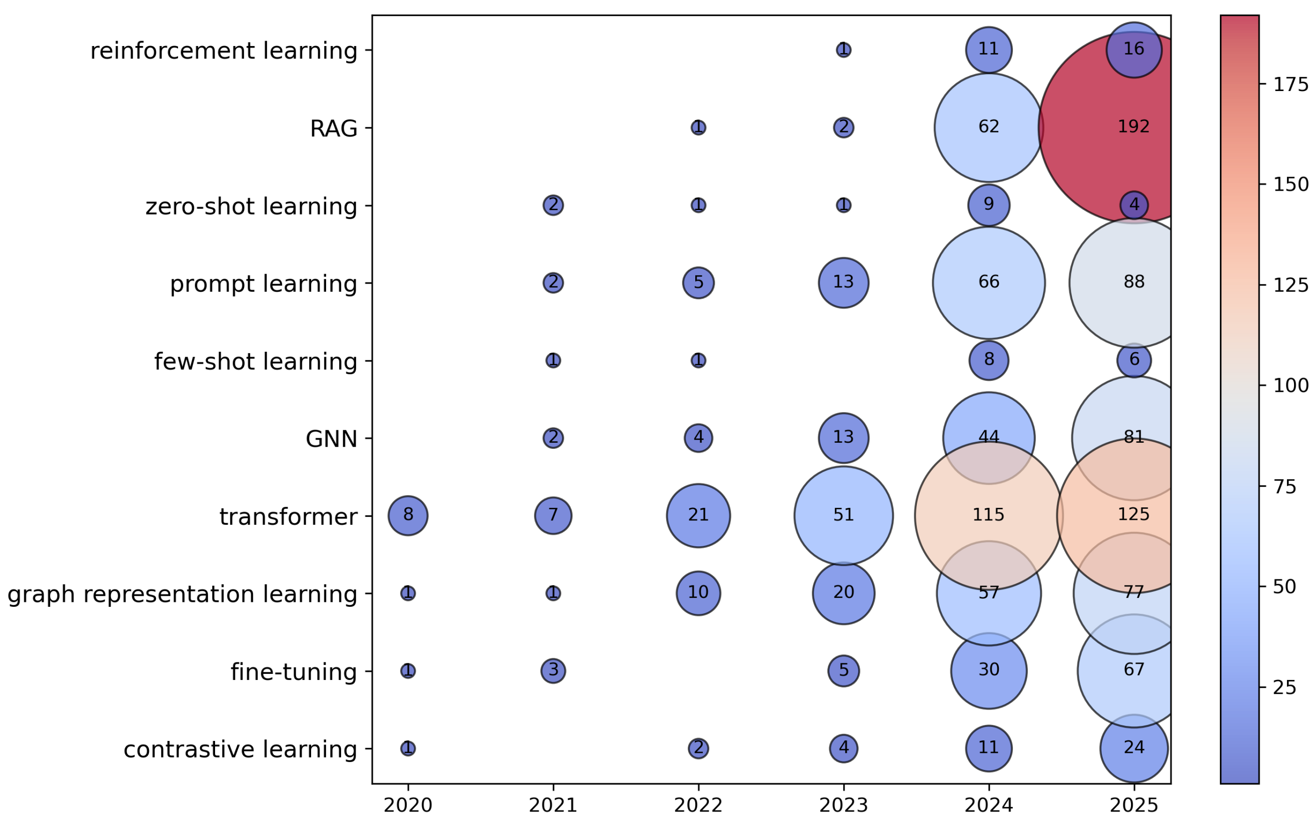

Figure 11. Distribution of techniques by year

or the most frequently addressed areas but also revealed several emerging themes or directions that have recently attracted the attention of researchers. It is important to note that the emerging research themes reported in this section are derived from bibliometric data analysis, using keyword co-occurrence and temporal trend visualization. Unlike existing surveys that primarily suggest future research directions from a conceptual point of view, our study identifies emerging themes based on large-scale bibliometric data. In fact, these themes which are summarized below, reflect actual trends in the literature, rather than speculative future research directions.

4.1 Knowledge-aware Prompt Engineering

Prompt engineering is the technique of guiding model outputs through carefully designed instructions (prompts) without modifying the underlying model parameters. Recently, it has increasingly gained attention as a pivotal mechanism to enhance the synergy between LLM and KG. Prompt engineering leverages LLMs and KGs to enhance several NLP and KG tasks such as knowledge extraction, QA, KG completion, link prediction and graph embedding etc.

In [90], the authors assessed the potential of advanced prompt engineering methods for improving outcomes in knowledge extraction. In [91], a novel prompting framework named KnowGPT is introduced. KnowGPT contains two main modules: (i) a knowledge extraction module to extract the most informative knowledge from KGs, and (ii) a context-aware prompt construction module to automatically convert extracted knowledge into effective prompts. In [92], the effects of various prompt engineering approaches have been explored on the task of Knowledge Graph QA for a complex data set.

[93] introduced a KG Prompting method to formulate the right context in prompting LLMs for Multi Document QA, which consists of a graph construction module and a graph traversal module. [94] proposes KG-LLM, a framework that applies prompt engineering to convert structured KG data into natural language prompts, enabling large language models (LLMs) to capture entity representations and improve multi-hop link prediction in KGs.

In [95], a pipeline based on LLMs and prompt engineering is proposed where the aim is to construct a heart failure knowledge graph to support diagnosis and treatment. The construction of the KG is divided into three phases: schema design, information extraction including named entity recognition and relation extraction, and knowledge graph completion, including triple classification, relation prediction and link prediction.

In [96], a framework named KICGPT that integrates LLM and a triple-based KG completion retriever is proposed. KICGPT uses Knowledge Prompt strategy allowing encoding structural knowledge into demonstrations to guide the LLM. [97] proposes a novel framework, called SIKGC, which builds the structural information prompt to assist the knowledge graph completion tasks. SIKGC reformulates KG triples into text sequences and integrates entity and relation descriptions with structural information as task-aware prompts. These prompts are provided to LLMS, which interpret the outputs as prediction tasks.

4.2 Retrieval-Augmented Generation for KG and LLM Integration

In recent years, RAG has become one of the most emerging technologies within generative AI, leveraging the synergy between retrieval systems and language models to enhance factual accuracy and produce contextually rich outputs. To enhance generations, several recent studies focus on retrieving structured knowledge from graphs. In fact, Knowledge graph retrieval-augmented generation not only improves retrieval precision but also mitigates hallucinations and enhances reasoning capabilities. Today, RAG is widely applied in areas such as information retrieval, QA, and intelligent assistants in sensitive domains like healthcare, law, and education.

KG2RAG [98] is a Knowledge Graph-Guided Retrieval Augmented Generation framework that utilizes KGs to provide fact-level relationships between chunks. After performing a semantic-based retrieval, it employs a KG-guided chunk expansion process to provide seed chunks. Then, a KG-based chunk organization process is used to deliver relevant and important knowledge in well-organized paragraphs.

[99] introduces a customer service question-answering method that integrates RAG with a KG. This later is constructed from historical issues, retaining both the intra-issue structure and inter-issue relations. During the question-answering phase, consumer queries are parsed and relevant sub-graphs are retrieved from the KG to generate answers.

G-Retriever [100] is a RAG framework designed for general textual graphs. To enhance graph understanding, this method leverages soft prompting for fine-tuning. Furthermore, G-Retriever formulates the retrieval process as a Prize-Collecting Steiner Tree optimization problem in order to mitigate hallucination and handle textual graphs that extend far beyond the context window of LLMs.

In [101] a personalization approach for LLM that integrates RAG and KGs is introduced. This approach focuses on using smaller models to achieve efficient results, which can be further deployed on personal devices such as smartphones.

[102] also combines KG and RAG techniques in a framework called d KG-RAG to enhance LLMs performance in the telecom domain. In this framework, a KG is built and encapsulates key entities, such as network protocols, hardware components, and signaling standards, as well as their relationships. Then, the retrieval mechanism dynamically retrieves relevant subgraphs or nodes from the KG.

K-RagRec [93] is a knowledge retrieval-augmented recommendation framework that provides up-to-date and reliable knowledge. Firstly, it performs knowledge sub-graph indexing on the items' KG at a coarse and fine granularity to construct the knowledge vector database. Then, it applied an adaptive retrieval policy to determine which items should be retrieved, followed by the retrieval of specific sub-graphs. Finally, the

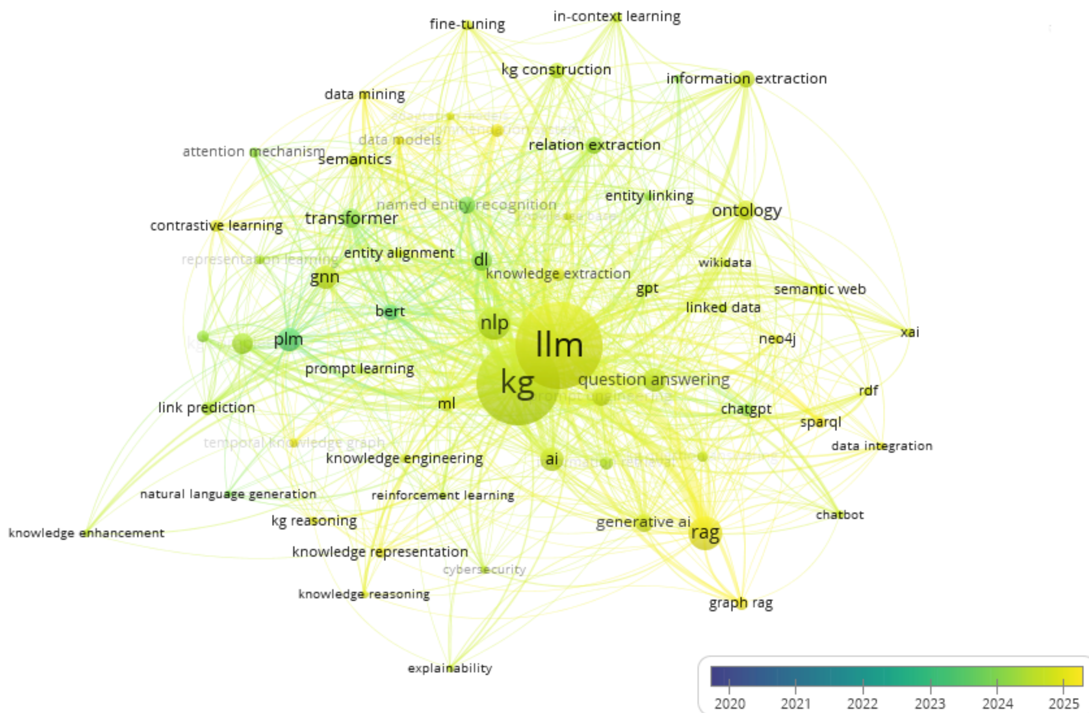

Figure 12. Distribution of Authors Keywords

retrieved knowledge sub-graphs are aligned into the semantic space of the LLM for knowledge-augmented recommendation.

4.3 Knowledge Graph–Based Hallucination Detection

One of the most critical challenges to deploy LLMs in the real-world applications is hallucination where models generate outputs that appear coherent and convincing, yet are factually incorrect. Despite efforts to address hallucination detection, it is expected to persist. Recently, leveraging KGs for reliable hallucination detection is emerging as a promising and trending research direction. According to [103], Hallucination Detection is still a very active area of research with various unresolved open problems.

[104] proposes a hallucination evaluation framework called GraphEval that is based on representing information in KG structures. It allows the identification of the specific triples in the KG that are prone to hallucinations and hence provides more insight into where in the response a hallucination has occurred.

[105] proposes a black-box approach named FactAlign which automatically detects and classifies hallucinations at a fact level by transforming the problem into a KG alignment task KGR.

[106] addresses the hallucination detection through a specialized framework that incorporates LLMs with KGs. By retrofitting the initial draft responses of LLMs based on the factual knowledge stored in KGs, it mitigates factual hallucination during the reasoning process.

CoKGLM [107] also performs hallucination detection by enhancing the use of both textual and structural information from KGs and thoroughly examines desirable KG information for hallucination detection.

In [108], hallucinations in LLMs are reduced by incorporating KGs as an additional modality. The input text is transformed into a set of KG embeddings. Then, and by using an adapter, these embeddings are integrated into language model space, without relying on external retrieval processes.

4.4 Multi-modal KGs (MM-KGs) and Multi-modal LLMs (MM-LLMs)

Most existing KGs are built from textual and graph structure sources, but this singular modality of knowledge constrains their ability to represent and handle the multimodal nature of real-world knowledge. Moreover, they remain limited in integrating information transmitted by other modalities (such as visual, auditory or spatio-temporal data) which can restrict their capacity to support complex reasoning tasks. Recent advances in LLMs, and more recently Multi-modal LLMs (MM-LLMs), offer promising avenues to address these limitations. So, the ability of MM-LLMs to reason over heterogeneous modalities can provide new opportunities for the construction of MM-KGs. On the other hand, MM-KGs can also be leveraged to empower LLMs with advanced multimodal reasoning skills by using the rich knowledge contained in MMKGs.

In [109], a Multimodal Reasoning with Multimodal Knowledge Graph (MR-MKG) method is introduced. This method exploits multimodal KG to learn rich and semantic knowledge across modalities. This integration allows enhancing the multimodal reasoning capabilities of LLMs. For encoding MMKGs, a relation graph attention network is utilized while a cross-modal alignment module is designed for optimizing image-text alignment. Furthermore, a MMKG-grounded dataset is developed for enhancing multimodal reasoning.

[110] proposes a multimodal KG reasoning method based on a cross-attention mechanism. In this method, the BERT model is used to encode textual information while the VGG-16 model, with a multi-scale attention module, is utilized to encode visual information. The graph attention mechanism is employed to extract the structural information of the KG. In order to reduce the semantic inconsistency between images and text, contrastive learning is applied.

In [111], a novel approach which exploits LLMs to automate the construction of synthetic multi-modal KGs specifically designed for a Smart City Cognitive Digital Twin (SC CDT). Based on the power of LLMs for natural language understanding, entity recognition, and relationship extraction, this approach facilitates the integration of data originating from sensor networks, social media feeds, official reports, and other relevant sources.

[112] introduces a GPT-based multimodal knowledge graph completion (MLKGC) framework which combines multimodal data and LLMs for enhancing the completeness and accuracy of KG completion tasks. Audio and image modalities are introduced to enhance the expressive power of KGs, enriching knowledge from diverse domains and strengthening multimodal reasoning.

VaLiK [113] is a Vision-align-to-Language integrated Knowledge Graph framework allowing the construction of MMKGs that enhances LLMs' multimodal reasoning capabilities. For MMKG construction, a zero-shot method that captures deep semantic connections beyond traditional predetermined labels with an effective verification system.

5. Conclusion

The aim of our study is to provide; by using bibliometric techniques; a quantitative analysis of the current state and future trends of KGs and LLMs integration research. The results of our analysis show that there is a rapid and exponential increase over the past years, particularly since 2023. As predicted by the Logistic Growth Model, this growing is expected to continue. In term of publications and citations, China is the most contributor. Furthermore, from the most prolific authors, four are from Chinese institutions which prove the country's strong research presence in this domain. The most active source was ACL as conference and IEEE Access as journal.

Our analysis reveals the KG-based tasks have attracted increasing attention where "KG-to-text generation" has emerged as the most widely discussed research topic with an impressive increase. KG-enhanced LLMs have also witnessed a remarkable and rapid development where the dominant category is KG-enhanced LLM inference. Among the identified techniques, RAG is the most widely used technique for injecting structured knowledge into LLM. Four research directions are expected to dominate future work in the field: "Knowledge-aware Prompt Engineering", "RAG for KGs and LLMs Integration", "KG-Based Hallucination Detection" and "Multi-modal KGs and Multi-modal LLMs".

To conclude, our analysis provides researchers and practitioners with systematic information on KGs and LLMs integration. It can help them to better understanding the current state of the field and its research trends. Moreover, the findings of this review highlight the key research techniques and topics as well as the Emerging research themes.

Author Contributions

Conceptualization: Wafa Ghemmaz and Maroua Bouzid; Methodology: Wafa Ghemmaz and Maroua Bouzid; Software: Wafa Ghemmaz and Souad Bouaicha; Data curation: Wafa Ghemmaz and Naila Marir; Formal analysis: Wafa Ghemmaz; Visualization: Wafa Ghemmaz; Validation: Wafa Ghemmaz, Maroua Bouzid, Naila Marir and Souad Bouaicha; Writing and original draft preparation: Wafa Ghemmaz; Review and Editing: Wafa Ghemmaz, Maroua Bouzid, Souad Bouaicha and Naila Marir.

References

[1] DEVLIN, J. et al. Bert: pre-training of deep bidirectional transformers for language understanding. *arXiv*, arXiv, n. arXiv:1810.04805, maio 2019.

[2] LIU, Y. et al. Roberta: A robustly optimized bert pretraining approach. *arXiv*, arXiv, n. arXiv:1810.04805, 2019.

[3] RADFORD, A. et al. Improving language understanding by generative pre-training. San Francisco, CA, USA, 2018.

[4] RAFFEL, C. et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research*, v. 21, n. 140, p. 1–67, 2020.

[5] SU, D. et al. Generalizing question answering system with pre-trained language model fine-tuning. In: *Proceedings of the 2nd Workshop on Machine Reading for Question Answering*. [S.l.: s.n.], 2019. p. 203–211.

[6] LI, J. et al. Pre-trained language models for text generation: A survey. *ACM Computing Surveys*, v. 56, n. 9, p. 1–39, 2024. ISSN 0360-0300, 1557-7341.

[7] WEI, J. et al. Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*, 2022.

[8] ZHANG, Z. et al. Llm hallucinations in practical code generation: Phenomena, mechanism, and mitigation. *Proceedings of the ACM on Software Engineering*, v. 2, p. 481–503, 2025.

[9] JI, Z. et al. Survey of hallucination in natural language generation. *ACM Computing Surveys*, v. 55, n. 12, p. 1–38, 2023.

[10] BANG, Y. et al. A multitask, multilingual, multimodal evaluation of chatgpt on reasoning, hallucination, and interactivity. In: *Proceedings of the 13th International Joint Conference on Natural Language Processing*. [S.l.]: Association for Computational Linguistics, 2023. p. 675–718.

[11] LIAO, Q. V.; VAUGHAN, J. W. Ai transparency in the age of llms: A human-centered research roadmap. *Harvard Data Science Review*, n. Special Issue 5, maio 2024. ISSN 2644-2353, 2688-8513. Disponível em: ⟨https://hdsr.mitpress.mit.edu/pub/aelql9qy/release/2⟩.

[12] HOGAN, A. et al. *Knowledge graphs*. [S.l.]: Springer Verlag, 2021. v. 12. (Synthesis Lectures on Data, Semantics, and Knowledge, v. 12). ISBN 9783031019180.

[13] JI, S. et al. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, v. 33, n. 2, p. 494–514, fev. 2022. ISSN 2162-2388. Disponível em: ⟨https://ieeexplore.ieee.org/document/9416312⟩.

[14] ZHANG, J. et al. Neural, symbolic and neural-symbolic reasoning on knowledge graphs. *AI Open*, v. 2, p. 14–35, jan. 2021. ISSN 2666-6510.

[15] PAN, S. et al. Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering*, v. 36, n. 7, p. 3580–3599, jul. 2024.

[16] CAI, L. et al. Practices, opportunities and challenges in the fusion of knowledge graphs and large language models. *Frontiers in Computer Science*, v. 7, jul. 2025. ISSN 2624-9898.

[17] PAN, J. Z. et al. Large language models and knowledge graphs: Opportunities and challenges. *Transactions on Graph Data and Knowledge (TGDK)*, v. 1, n. 1, p. 2:1–2:38, 2023.

[18] SINGH, V. K. et al. The journal coverage of web of science, scopus and dimensions: A comparative analysis. *Scientometrics*, v. 126, n. 6, p. 5113–5142, jun. 2021.

[19] ZENG, L. et al. Landscapes and emerging trends of virtual reality in recent 30 years: A bibliometric analysis. In: *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*. Guangzhou: IEEE, 2018. p. 1852–1858.

[20] SCHRAVEN, D.; JOSS, S.; JONG, M. de. Past, present, future: Engagement with sustainable urban development through 35 city labels in the scientific literature 1990-2019. *Journal of Cleaner Production*, v. 292, p. 125924, abr. 2021.

[21] ZHANG, Z. et al. Ernie: Enhanced language representation with informative entities. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. [S.l.]: Association for Computational Linguistics, 2019. p. 1441–1451.

[22] BOSSELUT, A. et al. Comet: Commonsense transformers for automatic knowledge graph construction. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. [S.l.]: Association for Computational Linguistics, 2019. p. 4762–4779.

[23] YASUNAGA, M. et al. Qa-gnn: Reasoning with language models and knowledge graphs for question answering. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. [S.l.]: Association for Computational Linguistics, 2021. p. 535–546.

[24] WANG, L. et al. Simkgc: Simple contrastive knowledge graph completion with pre-trained language models. *arXiv*, 2022.

[25] SUN, T. et al. Colake: Contextualized language and knowledge embedding. In: *Proceedings of the 28th International Conference on Computational Linguistics*. [S.l.]: International Committee on Computational Linguistics, 2020. p. 3660–3670.

[26] YASUNAGA, M. et al. Deep bidirectional language-knowledge graph pretraining. *arXiv*, 2022.

[27] MARINO, K. et al. Krisp: Integrating implicit and symbolic knowledge for open-domain knowledge-based vqa. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.]: IEEE, 2021. p. 14106–14116.

[28] FENG, Y. et al. Scalable multi-hop relational reasoning for knowledge-aware question answering. *arXiv preprint arXiv:2005.00646*, 2020.

[29] LOGAN, R. et al. Barack's wife hillary: Using knowledge graphs for fact-aware language modeling. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics, 2019. p. 5962–5971.

[30] KHORASHADIZADEH, H. et al. Construction and canonicalization of economic knowledge graphs with llms. In: SPRINGER. *International Knowledge Graph and Semantic Web Conference*. [S.l.], 2024. p. 334–343.

[31] KHORASHADIZADEH, H. et al. Exploring in-context learning capabilities of foundation models for generating knowledge graphs from text. *arXiv preprint arXiv:2305.08804*, 2023.

[32] VAKAJ, E. et al. Knowledge graphs for responsible ai. In: *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. [S.l.: s.n.], 2024. p. 5596–5598.

[33] MIHINDUKULASOORIYA, N. et al. Scholarly wikidata: Population and exploration of conference data in wikidata using llms. In: SPRINGER. *International Conference on Knowledge Engineering and Knowledge Management*. [S.l.], 2024. p. 243–259.

[34] MIHINDUKULASOORIYA, N. et al. Text2kgbench: A benchmark for ontology-driven knowledge graph generation from text. In: SPRINGER. *International semantic web conference*. [S.l.], 2023. p. 247–265.

[35] MIHINDUKULASOORIYA, N. et al. Unleashing the potential of data lakes with semantic enrichment using foundation models. In: *ISWC (Posters/Demos/Industry)*. [S.l.: s.n.], 2023.

[36] BUEHLER, M. J. Accelerating scientific discovery with generative knowledge extraction, graph-based representation, and multimodal intelligent graph reasoning. *Machine Learning: Science and Technology*, IOP Publishing, v. 5, n. 3, p. 035083, 2024.

[37] BUEHLER, M. J. Generative retrieval-augmented ontologic graph and multiagent strategies for interpretive large language model-based materials design. *ACS Engineering Au*, ACS Publications, v. 4, n. 2, p. 241–277, 2024.

[38] HU, Y.; BUEHLER, M. J. End-to-end protein normal mode frequency predictions using language and graph models and application to sonification. *ACS nano*, ACS Publications, v. 16, n. 12, p. 20656–20670, 2022.

[39] BUEHLER, M. J. Mechgpt, a language-based strategy for mechanics and materials modeling that connects knowledge across scales, disciplines, and modalities. *Applied Mechanics Reviews*, American Society of Mechanical Engineers, v. 76, n. 2, p. 021001, 2024.

[40] GHAFAROLLAHI, A.; BUEHLER, M. J. Sciagents: automating scientific discovery through bioinspired multi-agent intelligent graph reasoning. *Advanced Materials*, Wiley Online Library, v. 37, n. 22, p. 2413523, 2025.

[41] BUEHLER, E. L.; BUEHLER, M. J. X-lora: Mixture of low-rank adapter experts, a flexible framework for large language models with applications in protein mechanics and molecular design. *APL Machine Learning*, AIP Publishing, v. 2, n. 2, 2024.

[42] CHEN, Z. et al. Duet: Cross-modal semantic grounding for contrastive zero-shot learning. In: *Proceedings of the AAAI conference on artificial intelligence*. [S.l.: s.n.], 2023. v. 37, n. 1, p. 405–413.

[43] CHEN, Z. et al. Knowledge graphs meet multi-modal learning: A comprehensive survey. *arXiv preprint arXiv:2402.05391*, 2024.

[44] CHEN, Z. et al. Lako: Knowledge-driven visual question answering via late knowledge-to-text injection. In: *Proceedings of the 11th International Joint Conference on Knowledge Graphs*. [S.l.: s.n.], 2022. p. 20–29.

[45] LI, Q. et al. Llm-based multi-level knowledge generation for few-shot knowledge graph completion. In: *Proceedings of the 33rd International Joint Conference on Artificial Intelligence*. [S.l.: s.n.], 2024. v. 3, p. 2135–2143.

[46] ZHANG, Y. et al. Making large language models perform better in knowledge graph completion. In: *Proceedings of the 32nd ACM international conference on multimedia*. [S.l.: s.n.], 2024. p. 233–242.

[47] CHEN, Z. et al. Noise-powered multi-modal knowledge graph representation framework. In: *Proceedings of the 31st International Conference on Computational Linguistics*. [S.l.: s.n.], 2025. p. 141–155.

[48] CHEN, Z. et al. Self-improvement programming for temporal knowledge graph question answering. *arXiv preprint arXiv:2404.01720*, 2024.

[49] NA, Q. et al. A pre-training method inspired by large language model for power named entity recognition. In: *Proceedings of the 2023 5th International Conference on Internet of Things, Automation and Artificial Intelligence*. [S.l.: s.n.], 2023. p. 308–312.

[50] LI, X. et al. An improved transformer model based graph node classification method. In: IEEE. *2022 41st Chinese Control Conference (CCC)*. [S.l.], 2022. p. 7144–7149.

[51] FANG, Y. et al. Boosting with fewer tokens: Multi-query optimization for llms using node text and neighbor cues. In: IEEE. *2025 IEEE 41st International Conference on Data Engineering (ICDE)*. [S.l.], 2025. p. 2684–2697.

[52] HE, Y. et al. Hop-aware linear transformers for scalable graph neural networks. In: IEEE. *2025 44th Chinese Control Conference (CCC)*. [S.l.], 2025. p. 8676–8681.

[53] LI, X. et al. Toward cognitive digital twin system of human–robot collaboration manipulation. *IEEE Transactions on Automation Science and Engineering*, IEEE, 2024.

[54] MEYER, L.-P. et al. Assessing sparql capabilities of large language models, 2024. *URL: https://arxiv.org/abs/2409.05925*, 2024.

[55] FREY, J. et al. Assessing the evolution of llm capabilities for knowledge graph engineering in 2023. In: SPRINGER. *European Semantic Web Conference*. [S.l.], 2024. p. 51–60.

[56] FREY, J. et al. Benchmarking the abilities of large language models for rdf knowledge graph creation and comprehension: how well do llms speak turtle? *arXiv preprint arXiv:2309.17122*, 2023.

[57] MEYER, L. et al. Developing a scalable benchmark for assessing large language models in knowledge graph engineering. *arXiv preprint arXiv:2308.16622*, 2023.

[58] MEYER, L.-P. et al. Llm-kg-bench 3.0: A compass for semantic technology capabilities in the ocean of llms. In: SPRINGER. *European Semantic Web Conference*. [S.l.], 2025. p. 280–296.

[59] FREY, J.; FERRAZ, L.; HOFER, M. Pots-a polyparadigmatic ontology term search with fine-grained context steering using hyper-level vector spaces. In: *Companion Proceedings of the ACM on Web Conference 2025*. [S.l.: s.n.], 2025. p. 2831–2834.

[60] HOFER, M.; FREY, J.; RAHM, E. Towards self-configuring knowledge graph construction pipelines using llms-a case study with rml. In: *Fifth International Workshop on Knowledge Graph Construction@ ESWC2024*. [S.l.: s.n.], 2024. v. 3718.

[61] PEREVALOV, A. et al. Language models as sparql query filtering for improving the quality of multilingual question answering over knowledge graphs. In: SPRINGER. *International Conference on Web Engineering*. [S.l.], 2024. p. 3–18.

[62] PEREVALOV, A.; BOTH, A.; NGOMO, A.-C. N. Multilingual question answering systems for knowledge graphs–a survey. *Semantic Web*, SAGE Publications Sage UK: London, England, v. 15, n. 5, p. 2089–2124, 2024.

[63] SCHIESE, D.; PEREVALOV, A.; BOTH, A. Post-hoc natural-language explanations of component-based knowledge graph question answering systems generated by llms. *IADIS International Journal on Computer Science & Information Systems*, v. 20, n. 1, 2025.

[64] PEREVALOV, A. et al. Sparql query candidate filtering for improving the quality of multilingual question answering over knowledge graphs using language models. *Journal of Web Engineering*, River Publishers, v. 24, n. 4, p. 563–592, 2025.

[65] PEREVALOV, A.; BOTH, A. Towards llm-driven natural language generation based on sparql queries and rdf knowledge graphs. In: *3rd international workshop on knowledge graph generation from text (Text2KG) at ESWC*. [S.l.: s.n.], 2024.

[66] SCHIESE, D.; PEREVALOV, A.; BOTH, A. Towards llm-generated explanations for component-based knowledge graph question answering systems. *arXiv preprint arXiv:2508.14553*, 2025.

[67] HERTLING, S.; PORTISCH, J.; PAULHEIM, H. Matching with transformers in melt. *arXiv preprint arXiv:2109.07401*, 2021.

[68] HERTLING, S.; NOROUZI, E.; SACK, H. Oaei machine learning dataset for online model generation. In: SPRINGER. *European Semantic Web Conference*. [S.l.], 2024. p. 239–243.

[69] HERTLING, S.; PAULHEIM, H. Olala: Ontology matching with large language models. In: *Proceedings of the 12th knowledge capture conference 2023*. [S.l.: s.n.], 2023. p. 131–139.

[70] HERTLING, S.; SACK, H. Towards large language models interacting with knowledge graphs via function calling. 2024.

[71] ZHANG, Y.; GAN, L.; PI, X. An assembly instruction question answering system based on knowledge graph. In: SPRINGER. *International Conference on Applied Intelligence*. [S.l.], 2023. p. 322–334.

[72] ZHANG, Y. et al. Entity set co-expansion in stackoverflow. In: IEEE. *2022 IEEE International Conference on Big Data (Big Data)*. [S.l.], 2022. p. 4792–4795.

[73] ZHANG, Y. et al. Question-guided knowledge graph re-scoring and injection for knowledge graph question answering. In: *Findings of the Association for Computational Linguistics: EMNLP 2024*. [S.l.: s.n.], 2024. p. 8972–8985.

[74] ZHANG, Y.; ZHANG, Y.; HAN, J. Mining structures from massive texts by exploring the power of pre-trained language models. In: *EDBT*. [S.l.: s.n.], 2023. p. 851–854.

[75] ZHANG, Y. et al. Image captioning with transformer and knowledge graph. *Pattern Recognition Letters*, Elsevier, v. 143, p. 43–49, 2021.

[76] WANG, K. et al. Visual semantic planning for service robot via natural language instructions. In: IEEE. *2021 China Automation Congress (CAC)*. [S.l.], 2021. p. 793–798.

[77] LIU, B. et al. Unilp: Unified topology-aware generative framework for link prediction in knowledge graph. In: *Proceedings of the ACM Web Conference 2024*. [S.l.: s.n.], 2024. p. 2170–2180.

[78] LIU, B. et al. Symagent: A neural-symbolic self-learning agent framework for complex reasoning over knowledge graphs. In: *Proceedings of the ACM on Web Conference 2025*. [S.l.: s.n.], 2025. p. 98–108.

[79] XU, W. et al. Pre-trained language model with prompts for temporal knowledge graph completion. *arXiv preprint arXiv:2305.07912*, 2023.

[80] XU, W. et al. Historical facts learning from long-short terms with language model for temporal knowledge graph reasoning. *Information Processing & Management*, Elsevier, v. 62, n. 3, p. 104047, 2025.

[81] LIU, B. et al. Filter-then-generate: Large language models with structure-text adapter for knowledge graph completion. In: *Proceedings of the 31st International Conference on Computational Linguistics*. [S.l.: s.n.], 2025. p. 11181–11195.

[82] PENG, M. et al. Deja vu: Contrastive historical modeling with prefix-tuning for temporal knowledge graph reasoning. *arXiv preprint arXiv:2404.00051*, 2024.

[83] LUO, L. et al. Systematic assessment of factual knowledge in large language models. In: *Findings of the association for computational linguistics: EMNLP 2023*. [S.l.: s.n.], 2023. p. 13272–13286.

[84] LUO, L. et al. Reasoning on graphs: Faithful and interpretable large language model reasoning. *arXiv preprint arXiv:2310.01061*, 2023.

[85] LUO, L. et al. Integrating large language models and knowledge graphs for next-level agi. In: *Companion Proceedings of the ACM on Web Conference 2025*. [S.l.: s.n.], 2025. p. 33–36.

[86] NGUYEN, M.-V. et al. Direct evaluation of chain-of-thought in multi-hop reasoning with knowledge graphs. In: *Findings of the Association for Computational Linguistics: ACL 2024*. [S.l.: s.n.], 2024. p. 2862–2883.

[87] LUO, L. et al. Chatrule: Mining logical rules with large language models for knowledge graph reasoning. arxiv 2024. *arXiv preprint arXiv:2309.01538*, 2024.

[88] CHEN, X. et al. Mapping the research trends by co-word analysis based on keywords from funded project. *Procedia computer science*, Elsevier, v. 91, p. 547–555, 2016.

[89] LEWIS, P. et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, v. 33, p. 9459–9474, 2020.

[90] POLAT, F.; TIDDI, I.; GROTH, P. Testing prompt engineering methods for knowledge extraction from text. *Semantic Web*, SAGE Publications Sage UK: London, England, v. 16, n. 2, p. SW–243719, 2025.

[91] ZHANG, Q. et al. Knowgpt: Knowledge graph based prompting for large language models. *Advances in Neural Information Processing Systems*, v. 37, p. 6052–6080, 2024.

[92] KOSTEN, C.; NOORALAHZADEH, F.; STOCKINGER, K. Evaluating the effectiveness of prompt engineering for knowledge graph question answering. *Frontiers in Artificial Intelligence*, Frontiers Media SA, v. 7, p. 1454258, 2025.

[93] WANG, S. et al. Knowledge graph retrieval-augmented generation for llm-based recommendation. *URL https://arxiv.org/abs/2501.02226*, 2025.

[94] SHU, D. et al. Knowledge graph large language model (kg-llm) for link prediction. arxiv. *arXiv preprint arXiv:2403.07311*, 2024.

[95] XU, T. et al. Knowledge graph construction for heart failure using large language models with prompt engineering. *Frontiers in Computational Neuroscience*, Frontiers Media SA, v. 18, p. 1389475, 2024.

[96] WEI, Y. et al. Kicgpt: Large language model with knowledge in context for knowledge graph completion. *arXiv preprint arXiv:2402.02389*, 2024.

[97] LI, W. et al. Sikgc: Structural information prompt based knowledge graph completion with large language models. In: IEEE. *2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. [S.l.], 2024. p. 1116–1123.

[98] ZHU, X. et al. Knowledge graph-guided retrieval augmented generation. *arXiv preprint arXiv:2502.06864*, 2025.

[99] XU, Z. et al. Retrieval-augmented generation with knowledge graphs for customer service question answering. In: *Proceedings of the 47th international ACM SIGIR conference on research and development in information retrieval*. [S.l.: s.n.], 2024. p. 2905–2909.

[100] HE, X. et al. G-retriever: Retrieval-augmented generation for textual graph understanding and question answering. *Advances in Neural Information Processing Systems*, v. 37, p. 132876–132907, 2024.

[101] PRAHLAD, D. et al. Personalizing large language models using retrieval augmented generation and knowledge graph. In: *Companion Proceedings of the ACM on Web Conference 2025*. [S.l.: s.n.], 2025. p. 1259–1263.

[102] YUAN, D. et al. Enhancing large language models (llms) for telecommunications using knowledge graphs and retrieval-augmented generation. In: IEEE. *2025 IEEE International Conference on Communications Workshops (ICC Workshops)*. [S.l.], 2025. p. 486–491.

[103] LAVRINOVICS, E. et al. Knowledge graphs, large language models, and hallucinations: An nlp perspective. *Journal of Web Semantics*, Elsevier, v. 85, p. 100844, 2025.

[104] SANSFORD, H. et al. Grapheval: A knowledge-graph based llm hallucination evaluation framework. *arXiv preprint arXiv:2407.10793*, 2024.

[105] RASHAD, M. et al. Factalign: Fact-level hallucination detection and classification through knowledge graph alignment. In: *Proceedings of the 4th Workshop on Trustworthy Natural Language Processing (TrustNLP 2024)*. [S.l.: s.n.], 2024. p. 79–84.

[106] GUAN, X. et al. Mitigating large language model hallucinations via autonomous knowledge graph-based retrofitting. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2024. v. 38, n. 16, p. 18126–18134.

[107] HASEGAWA, R.; ICHISE, R. Cokglm: Detecting hallucinations generated by large language models via knowledge graph verification. In: SPRINGER. *International Knowledge Graph and Semantic Web Conference*. [S.l.], 2024. p. 212–224.

[108] CHEKALINA, V. et al. Addressing hallucinations in language models with knowledge graph embeddings as an additional modality. *arXiv preprint arXiv:2411.11531*, 2024.

[109] LEE, J. et al. Multimodal reasoning with multimodal knowledge graph. *arXiv preprint arXiv:2406.02030*, 2024.

[110] HAN, Z.; HOU, C. Multimodal knowledge graph inference method based on cross-attention mechanism. In: *Proceedings of the 2025 2nd International Conference on Generative Artificial Intelligence and Information Security*. [S.l.: s.n.], 2025. p. 94–100.

[111] MANDAL, S.; O'CONNOR, N. E. Llmasmmkg: Llm assisted synthetic multi-modal knowledge graph creation for smart city cognitive digital twins. In: *Proceedings of the AAAI Symposium Series*. [S.l.: s.n.], 2024. v. 4, n. 1, p. 210–221.

[112]YUE, P. et al. Mlkgc: Large language models for knowledge graph completion under multimodal augmentation. *Mathematics*, MDPI, v. 13, n. 9, p. 1463, 2025.

[113]LIU, J. et al. Aligning vision to language: Text-free multimodal knowledge graph construction for enhanced llms reasoning. *arXiv preprint arXiv:2503.12972*, 2025.