2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

# UNIFYING LARGE LANGUAGE MODELS AND KNOWLEDGE GRAPHS: A COMPREHENSIVE SURVEY

**Rajeev Kumar**

Machine Learning Leader

**Abstract**

Large Language Models (LLMs) and Knowledge Graphs (KGs) were combined to create a pioneering artificial intelligence tool that combines LLMs' generative capability with KGs' systematic representation of data and reasoning. This roadmap outlines a comprehensive strategy for integrating these two paradigms to address key challenges, including knowledge grounding, contextual accuracy, scalability, and interpretability. We explore methodologies for enhancing synergy between LLMs and KGs, such as knowledge-augmented pretraining, hybrid architectures, and real-time knowledge updates. Additionally, we highlight emerging applications in natural language understanding, conversational AI, decision-making, and personalized systems. This roadmap seeks to provide academics and practitioners a starting point for maximising an opportunity of combined LLM-KG techniques for developing AI abilities by highlighting open problems and possible future paths.

***Keywords:*** Large Language Models, Knowledge Graphs, Artificial Intelligence, Knowledge Integration, Natural Language Processing

## 1. Introduction

Large language models (LLMs)1 (for example, BERT [1], RoBERTA [2], and T5 [3]), which have been pre-trained on the large-scale corpus, have shown remarkable performance in a variety of natural language processing (NLP) tasks. These tasks include question answering [4], machine translation [5], and text synthesis [6]. Lately, the exponential growth in model size has made it possible for LLMs to exhibit emergent capabilities [7], opening the door for their use as Artificial General Intelligence (AGI).

With billions of parameters, sophisticated LLMs like ChatGPT2 and PaLM23 show promise in a variety of challenging real-world applications, including recommendation [10], code creation [9], and teaching [8].Although LLMs have been successful in many cases, their lack of factual understanding has drawn criticism. In particular, LLMs commit information from the training corpus to memory [11]. Nevertheless, other research shows that LLMs struggle to remember information and often suffer from hallucinations by making factually erroneous claims [12]. When questioned, for instance, "When did Einstein discover gravity?" LLMs would respond, "Einstein discovered gravity in 1687," which runs against to the idea that Isaac Newton developed the gravitational theory. This problem seriously damages LLMs' credibility.

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

The lack of interpretability of LLMs is another reason they are criticised as black-box models. The parameters of LLMs indirectly indicate knowledge. Interpreting and validating the information acquired by LLMs is challenging. Furthermore, LLMs use an unsatisfactory argumentation process based on the likelihood model. Humans are unable to reach or interpret LLM structures and processes for forecasting or choosing. Although some LLMs can use a chain-of-thought to justify their predictions, the hallucination problem also affects their reasoning. This significantly hinders the use of LLMs in situations with significant stakes, such making legal decisions and diagnosing illnesses. For example, in a situation involving medical diagnosis, LLMs can make a mistaken diagnosis and provide justifications that defy conventional logic. In the absence of domain-specific understanding or new training information, LLMs learnt on general corpus may not generalise to specific categories or incorporate fresh data. Adding knowledge graphs (KGs) to LLMs is one way to potentially solve the aforementioned problems. A systematic and definitive method of representing knowledge is via knowledge graphs (KGs), which store vast amounts of information as triples (head entity, relation, tail entity) (e.g., Wikidata YAGO, and NELL Because they provide precise explicit information, KGs are essential for a number of applications. Additionally, they are well known for producing interpretable conclusions by symbolic reasoning. With fresh information being provided on a regular basis, KGs may also actively change. In order to give accurate and trustworthy domain-specific knowledge, professionals may also create domain-specific KGs.


**Figure 1: An overview of the benefits and drawbacks for KGs and LLMs**

However, KGs are challenging to create, and existing KG techniques are insufficient to deal with the imperfect and constantly evolving character of real-world KGs. These methods fall short in representing new facts and modelling invisible things. Furthermore, they often overlook the wealth of textual information included in KGs. Furthermore, current approaches in KGs are often tailored for particular KGs or objectives, making them insufficiently generalisable. As a result, using LLMs to solve the issues encountered in KGs is equally essential. In Fig. 1, we list the benefits and drawbacks of KGs and LLMs, respectively. Researchers and practitioners have recently been more interested in the potential for combining LLMs and KGs. Because of their intrinsic connections, KGs and LLMs may benefit from one another. In KG-enhanced LLMs, KGs may supply external information during pre-training and inference and be used for LLM analysis and interpretability. Although the majority of research on knowledge-enhanced LLMs focusses on using Knowledge Graphs (KGs) as external information resources in enhancing LLMs, it often ignores other methods for combining KGs and LLMs as well as the possible functions that LLMs may have in KG-related activities. In this work, we provide a future-oriented plan for merging LLMs and KGs, with the goal of using both methods' advantages while resolving their drawbacks for a variety of downstream applications. We conduct comprehensive analyses, propose a detailed categorization, and identify emerging opportunities in these rapidly evolving fields. Below is a summary of our key contributions:

1) A roadmap. We provide a proactive integration plan for KGs and LLMs. Our roadmap offers guidance for the unification of these two different but complimentary technologies.

2) Review and classification. Our plan includes a thorough classification and new taxonomy of research on combining LLMs and KGs for each integration framework. To get a deeper understanding of each framework, we examine the research in each category from the viewpoints of various integration tasks and methodologies.

3) Reporting on new developments. We go over the more complex methods in both KGs and LLMs.

4) A synopsis of the difficulties and potential paths. We point out the shortcomings of the current body of research and provide a number of exciting avenues for further investigation.

**2. Common Topics of Discussion in the Community**

The Knowledge Computing community is divided on the use of parametric and explicit knowledge together, with supporters and detractors presenting opposing viewpoints. Here are some synopses of often raised issues.

Reasoning and Knowledge Representation: Inference and reasoning are made possible by KGs, which provide an organised representation of information with clear linkages. Parametric knowledge in LLMs, according to critics, is based more on statistical patterns than on actual comprehension and logic [13]. Advocates of LLMs, such as ChatGPT, emphasise their superior language comprehension skills, capacity to generalise from extensive text corpora, and ability to capture a variety of information. On the one hand, since LLMs lack explicit knowledge representation, they may produce answers that are believable yet inaccurate or illogical, like

2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

hallucinations [14]. Additionally, there are questions about whether LLMs can infer subsumption between concepts [15] or acquire directional entailments [16]. However, KGs may be expensive to construct. Even while training LLMs might be costly, they can be easily used to serve a wide range of downstream applications, elevating AI from the background to the forefront. For LLMs, parametric knowledge is thus not the (only) goal.

Extremely Accurate Techniques: The capacity of KGs to provide accurate and truthful information on entities is partly responsible for their success. YAGO for example, claims an accuracy rate of above 95%. Similar to this, Google requires high accuracy in their KG for operational usage. For example, Knowledge Vault's semi-automated building approach was not used in production, in part because it failed to meet their benchmark's required 99% accuracy.

Values in numbers: It is often known that LLMs have difficulty grasping numerical quantities. According to a Big-bench research, LMs may find it difficult to complete even simple arithmetic problems. This flaw also applies to KG completion assignments. The potential of many LLMs to finish KGs has been assessed using Wikidata's numerical information, including the years of birth and death of persons. But not a single year was correctly predicted by any of the models that were examined. This calls into question whether contemporary LLMs can accurately memorise numbers during pre-training in a manner that will allow them to be used later on in KG completion. Although LLMs such as PaLM show some aptitude for handling numbers, smaller models that are more often used seem to be unsuited for this work. When taking into account the complexities of measurements and various numbered formats and kinds, the complexity increases. Since there is currently no solution for adapting LLMs to handle numerical values, using them for numerical KG completion seems to be far from feasible.

**3. Prospects and Visions**

What new chances do we have now that parametric information has emerged? This is one of the main topics this study must address. Here are some of our opinions about these new prospects brought about by parametric knowledge's introduction and possible combination with explicit knowledge.

1. Instant access to vast text corpora: As the Introduction states, humans have traditionally transmitted their knowledge via texts. As a result, a large portion of information nowadays is found in texts. Large text corpora may be accessed quickly using LLMs, and more recently, on consumer hardware. This enables AI developers to stay clear of previously significant obstacles pertaining to large-scale data collection, preparation, storage, and querying. Additionally, it lessens hitherto significant reliance on the information retrieval industry.

2. Richer knowledge for many subtasks: LLMs' most significant characteristics, like as dialogue and inquiry responding, are still being studied, but they have substantially enhanced and simplified many of the information engineering pipeline's more common tasks. Dependency and structured parsing, entity identification, and connection extraction are only a few of the tasks that

2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

LLMs have progressed by unconventional methods, fine-tuning on a small number of samples, or few-shot prompting. Additionally, improvements spread as faults do down a pipeline, allowing for KG development at a scale and quality never before possible. In addition, LLMs may be easily used to a wide range of downstream activities outside of knowledge engineering.

The idea that "Knowledge is power" may be further realised by incorporating explicit information—especially structured knowledge—into LLMs, for example, by using retrieval-augmented approaches. This would make explicit knowledge easier to use for a variety of downstream activities.

3. Even more sophisticated language comprehension: Tasks like textual entailment, summarisation, paraphrase generation, and others demonstrate how much LLMs alone have improved our "understanding" of real language. These attributes are crucial for information engineering to be robust against typographical errors, redundant employees, linguistic variety, and other features of human-written, web-scraped, and many noisy document forms. It is now feasible to achieve even more sophisticated language comprehension for textual entailments as well as other NLP tasks like summarisation and consistent generation thanks to potentially innovative methods for fusing parametric and explicit information.

4. Consolidation involves compression: In conventional knowledge engineering, combining contradicting and confirming bits of information is a crucial stage that calls for often complex techniques for combining phrase observations, patterns, and restrictions. An aggregation happens automatically during LLM training. This stage presents a significant knowledge engineering issue in terms of outsourcing, despite the fact that it is not well understood.

**3.1 Classification**

We also provide a detailed classification for every framework in the roadmap to help readers better comprehend the study on combining LLMs and KGs. We specifically concentrate on three methods of combining KGs and LLMs: synergised LLMs + KGs, KG-augmented LLMs, and KGenhanced LLMs. Figure 2 depicts the research's fine-grained classification.

LLMs with improvements in KG. LLM efficiency and accessibility in downstream processes may be enhanced by incorporating KGs. The study on KG-enhanced LLMs is divided into 3 categories:

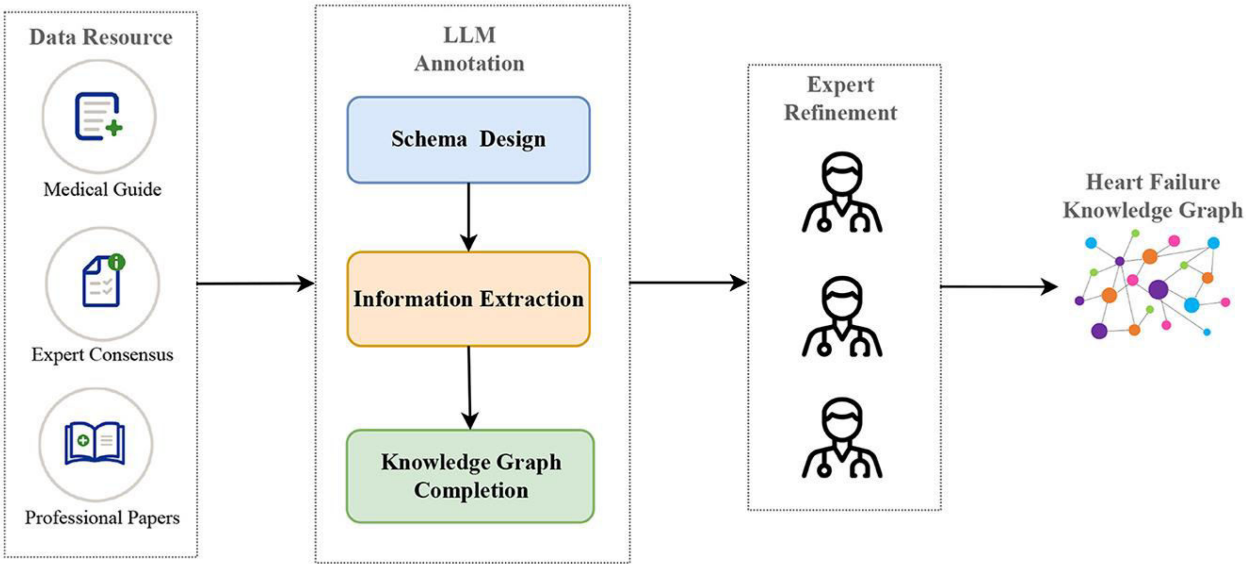


**Figure 2: Detailed classification of studies on integrating knowledge graphs (KGs) with large language models (LLMs).**

KG-Enhanced LLM Pre-Training: This group comprises methods that use Knowledge Graphs (KGs) to improve the modelling and display for understanding in Large Language Models (LLMs) during the pre-training stage.

KG-Enhanced LLM Interpretation: By including KGs into the LLM interpretation phase, these techniques provide the algorithms access to current data without necessitating rehabilitation. This method guarantees that LLMs may respond using the most up-to-date and pertinent information.

KG-Enhanced LLM Interpretation: This technique makes use of KGs to better comprehend the deductive methods of LLMs and to interpret the information they contain, which increases explanation and visibility.

LLM-Augmented KGs include five areas of study on LLMs' assistance for KG responsibilities:

LLM-Augmented KG Construction: This technique makes it easier to create knowledge charts through the utilisation of LLMs for activities like entity exploration, coreference conclusion, and connection analysis.

LLM-Augmented KG-to-Text Generation: These works use LLMs to construct descriptions in natural language of KG facts.

Combined LLMs and KGs: This study examines how LLMs and KGs might be integrated from the standpoints of understanding and logic in order to capitalise on their complementary advantages for applications with greater complexity.

## 4. LLMs for KGs: Building Knowledge Graphs

2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

We highlight how important LLMs are to improving KG construction, focussing on the latest advancements, issues, and unanswered concerns in this field. We start by discussing link estimation, a technique for generating new facts from a previous KG. Next, we look into inductive link forecasting, which generates triple forecasts for invisible links. Next, we turn our attention to a more recent technique that directly extracts triples from an LLM's parametric knowledge.

We address the difficulties with LLM-based approaches for KG building as a conclusion to this section. These include problems with numerical values, long-tail entities, and the accuracy of these techniques.

**4.1 Predicting Links**

Link prediction is the process of forecasting a triple's missing element based on its other two members. It comprises the following predictions: relation (h,?, t), head entity (?, r, t), and tail (h, r,?). The majority of research on KG link prediction techniques has focused on static KG snapshots. Many of these models are thus limited to using things for which an embedding was discovered during the training process. Because of this, they are unable to forecast linkages for any previously unknown entities, such recently added individuals or goods. ILP, on the other hand, emphasises on techniques that may anticipate links to new organisations that hadn't previously been included in a KG. Additionally, textual information and other literal information are often not used by current KG embedding-based KG completion techniques .

**4.2 KG Connectivity to LLM Inputs**

This kind of study places a strong focus on include pertinent information subsections in LLM inputs, as seen in Fig. 3. Information graph triples and words are tokenised and spliced using ERNIE 3.0. Additionally, the relation tokens in the triple or the corresponding tokens in every sentence are randomly hidden to further improve the pairing of understanding with cultural representations. Since the sentence's variables might communicate heavily with each knowledge sub-graph's signs, immediate information triple combining may cause Knowledge Distortion. K-BERT first integrates the skill triple into the sentence using an animated matrix to fix this. Only knowledge entities may access learning triple data and sentence tokens can only see each other in the self-attention section. Colake proposes a unified word-knowledge network (shown in Fig. 3) as an additional measure to reduce Information Distortion. Knowledge entity tokens are related to their surroundings, whereas input word tokens form a comprehensive word graph.

It is true that the aforementioned techniques may impart a great deal of information to LLMs. They ignore the longtail and infrequent entities, however, and mostly concentrate on the well-known ones. Enhancing the LLMs' representations of such things is the goal of DkLLM . Dict-BERT supplements input text with dictionary entries to enhance rare word representations. It also trains the neural model to locally align unusual spellings in the input phrases and terms from dictionaries and identify correctly translated input text from interpretations.

2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

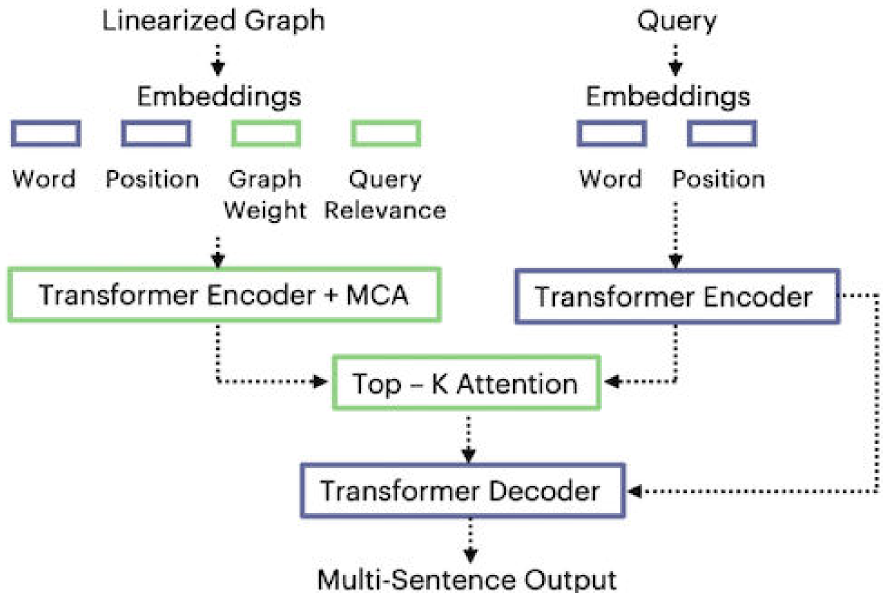


**Figure 3: Utilising a graph framework to inject KG data into LLM's inputs**

### 4.3 Improved LLM Inference with KG

The aforementioned techniques might successfully incorporate information into LLMs. The drawback of these methods is that they do not allow changes to the integrated information without retraining the model, and real-world knowledge is prone to change. They could thus not be able to generalise successfully to the information that is not visible during inference. As a result, a lot of work has gone into separating the text and knowledge spaces and incorporating the information during inference. These techniques mainly focus on QA difficulties since Question Answering (QA) requires the representation to capture either textual semantic significance and current actual facts.

Table annotation may be used to extract information for KG generation and populating by matching tables data to KG components, such as classes, entities, and properties. LLMs have been used for these objectives in a number of efforts. A table is serialised into a series of tokens using Doduo, which then trains BERT to anticipate the kinds of columns and the connections between them. ChatGPT is prompted for annotation of semantically column types.

ChatGPT performs similarly to RoBERTa model without task-specific examples. Even while using LLMs for KG generation and tabular data processing has received significant attention, there is still much to learn, particularly with the following issues:

Table contents must be turned into sequencing before being supplied into LLMs. distinct LLM utilisation situations, such as instructional tweaking of LLMs, LLM inferences with prompts, and fine-tuning LLMs, call for distinct transformation techniques.

Using and displaying information in tables that is not textual: In addition to lengthy and brief text, a table frequently incorporates other kinds of data, such as dates and numbers. Few works still take this facts into account.

Extracting tabular knowing: LLMs are seldom used for the last stage of knowledge removal, but they are mostly used to analyse and comprehend tables. People know about OntoGPT, which uses ChatGPT to pull examples from texts and fill in an an ontological framework but there aren't any similar tools for tables. It is more difficult to extract information with relationships over instances.

There are a number of difficulties, such as those listed in Section 2, in addition to the obviously high computing resource needs for training and using these LLM. More precisely, the following avenues for future development remain open:

Effective retrieval from lengthy texts. Long documents, such as novels, cannot be processed in one sitting by current LLMs. This makes it possible to further enhance corpus-level information extraction and long-range dependency modelling.

retrieval of high-coverage data. High accuracy is the main emphasis of almost all extracting pipes. High recall is disregarded or not sufficiently investigated. The development of data extractors with high recall and accuracy will be a significant step towards creating data extraction methods that last a lifetime.

### 4.3.1 Knowledge Fusion Enhanced by Retrieval

Retrieval-Enhanced Understanding One common technique for adding information to LLMs during inference is fusion. The main concept is to extract pertinent information from a big corpus, then combine that information to create LLMs. RAG suggests combining parametric and nonparametric modules to manage the external information, as seen in Fig 4. RAG first uses MIPS to look for pertinent KG in the non-parametric module given the input text, yielding a number of documents. According to the study, it is more effective to use many recovered documents as conditions at various stages of the creation process rather of relying just on one document to direct the whole process. According to the experimental findings, RAG performs better in open-domain QA than other baseline models that are parametric-only and non-parametric-only. Compared to other parameter-only baselines, RAG may also produce content that is more truthful, varied, and particular. By adding an extra module to identify important knowledge entities and include them into the generator, Story-fragments enhances the architecture even more and produces longer tales of higher quality. By using the quick maximum inner product search for memory querying and encoding external information into a keyvalue memory, EMAT significantly increases the effectiveness of such a system. In order to produce factual sentences, KGLM uses the present context to choose facts from a knowledge graph. KGLM might use out-of-domain terms or phrases to explain facts with the aid of an external knowledge graph.

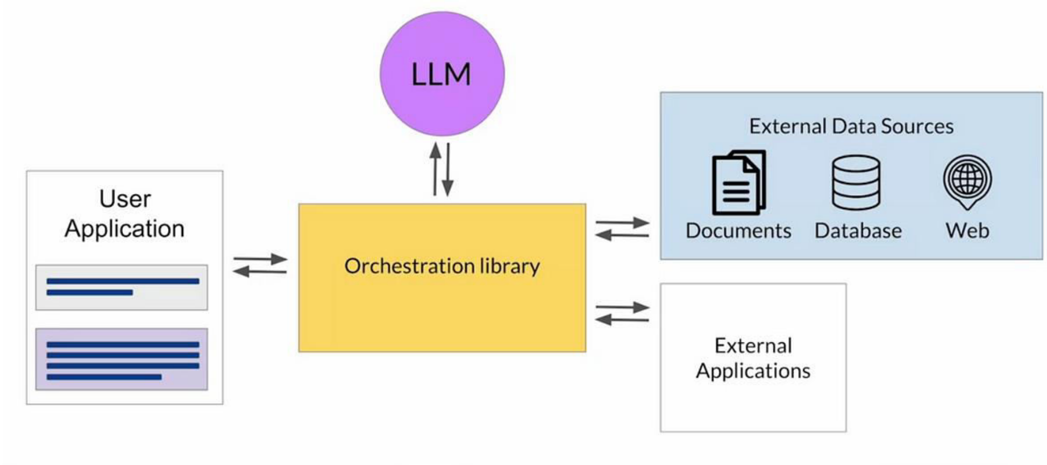


**Figure 4: Obtaining outside information to improve the creation of LLMs**

### 4.3.2 Electronic Medical Records

The automation of clinical recording, the synthesis of patient histories, and the identification of possible candidates for clinical trials are only a few of the many opportunities for the implementation of LLMs that the digital healthcare industry offers. Even if these developments are impressive, it's important to be aware of the possible hazards involved with using LLMs in the medical field. In fact, one of the most important application fields for LLM adoption is digital healthcare. The paradigm that led to the development of LLMs is in opposition to the demands of the main players, which include doctors, healthcare providers, and legislators. Specifically, the accuracy of the model and privacy issues arising from its use are the two main and important hazards.

Precision. Impressive powers have been shown in several LLM demonstrations. Nevertheless, there have also been recorded cases of LLMs acting erratically or making errors. Understanding the possible hazards linked with LLM use is essential for healthcare organisations operating in the digital healthcare sector, where patient safety is of the highest significance. LLMs have shown accuracy when used to diagnose fictitious patient cases that is on par with third- or fourth-year medical students, but falls short of a professional's level of competence. LLMs have been known to provide inaccurate material, fabricate sources, make logical mistakes, and give improper or immoral answers despite their high performance level. Given the potential to provide domain-specific information capable of mitigating the aforementioned difficulties, the incorporation of KGs would undoubtedly improve the capabilities of LLMs. The two main areas where LLMs+KG may significantly contribute to are avoiding hallucinations and maintaining ethics.

privacy. The fact that using any third-party application requires sending data to that party is a significant worry with LLMs. When a covered entity, such as a hospital, manages data, including protected health information (PHI), the data is subject to the laws of the country in which the

2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

institution is based (e.g., GDPR). Additionally, organisations lose control over the handling of PHI as they distribute information to more third parties. Healthcare companies that utilise LLMs need to be aware that their data may be more vulnerable to abuse or breaches. By modelling axioms that specify which data may be shared, with whom, and how personal knowledge may be anonymised before being communicated to potential external systems, KGs may play the function of protecting private information.

### 4.3.3 Knowledge of Commonsense

Most KGs record information similar to what one may find in a relational database or an encyclopaedia. Nonetheless, another crucial kind of world information for AI systems is commonsense knowledge. For example, a KG could want to include information about the Congo rainforest's location in Central Africa as well as the fact that tropical rainforests get a lot of rainfall and feature verdant flora. The most popular commonsense knowledge graph, ConceptNet, was created by combining automatic refining methods with human crowdsourcing. Alternative methods of gathering such information have long been sought after, however, since crowdsourcing is highly expensive and labour-intensive.

LLMs' Commonsense Knowledge: To the best of our knowledge, the first work to look at knowledge extraction from a language model focused on commonsense knowledge . The Google Web 1T n-gram data and Microsoft's Web-scale smoothed language models were used by the authors to harvest commonsense triples like hasProperty (apples, green). Later on, this was expanded into a comprehensive commonsense knowledge graph that included a variety of relations and was included into the WebChild KG. Table 1 summarizes the key concepts.

**Table 1: Encapsulating the key points about Knowledge Graphs (KGs) and Large Language Models (LLMs)**

| Aspect | Knowledge Graphs (KGs) | Large Language Models (LLMs) | Integration Opportunities |
|---|---|---|---|
| Benefits | - Structured, explicit knowledge representation<br>- Reasoning and inference support<br>- High accuracy (e.g., YAGO >95%) | - Generalizes across diverse tasks<br>- Learns from vast text corpora<br>- Enhances NLP tasks (e.g., summarization, parsing)<br>- Easy adaptability to downstream applications | - KGs provide external knowledge for pre-training and inference<br>- LLMs improve KG construction (e.g., entity recognition, link prediction)<br>- Better reasoning and explainability with combined methods |
| Drawbacks | - Costly and labor-intensive to construct | - Prone to hallucinations and inaccuracies | - LLMs can help automate KG |

| | | | |
|---|---|---|---|
| | - Struggles with dynamic and evolving data<br>- Limited generalizability to new tasks or KGs | - Limited handling of numerical data<br>- Lacks explicit knowledge representation | construction and handle noisy data<br>- KGs can mitigate LLM hallucinations and provide accurate factual data<br>- Combined frameworks allow dynamic updating of knowledge |
| Applications | - Explicit knowledge for reasoning (e.g., search engines)<br>- Link prediction<br>- Entity relationships | - Textual tasks (e.g., summarization, paraphrasing)<br>- Few-shot learning<br>- Commonsense reasoning and natural language generation | - KG-augmented LLMs for factual correctness<br>- LLM-enhanced KG creation<br>- Applications in healthcare, commonsense knowledge, and dynamic inference |
| Challenges | - Scaling and adapting to new entities<br>- Numerical data representation<br>- Incorporating textual knowledge | - Addressing hallucinations<br>- Handling long-tail entities and rare terms<br>- Ensuring factual correctness | - Developing frameworks for synergized KG + LLMs<br>- Efficient retrieval and fusion of knowledge<br>- Privacy and data security (e.g., healthcare applications) |
| Key Integration Frameworks | - KG-Enhanced LLMs: Improves LLM pre-training, inference, and interpretation<br>- LLM-Augmented KGs: Improves KG construction (e.g., link prediction, KG-to-text) | - Synergized LLMs + KGs: Combines the reasoning power of KGs with the language understanding of LLMs | - Unified frameworks for enhanced reasoning, explainability, and accuracy<br>- Retrieval-augmented systems for dynamic knowledge updates |
| Future | - Better long-tail entity | - Addressing numerical | - Retrieval-augmented |

| | | | |
|---|---|---|---|
| Directions | representation<br>- Improved reasoning with evolving knowledge<br>- Automating KG construction and maintenance | reasoning gaps<br>- Advanced commonsense integration<br>- Handling large-scale, long-range dependencies | LLMs (e.g., RAG)<br>- LLM-enhanced commonsense and tabular data understanding<br>- Privacy-preserving LLMs using KGs |

## 5. Discussion and interpretations

The integration of Knowledge Graphs (KGs) and Large Language Models (LLMs) represents a significant step forward in advancing knowledge representation, reasoning, and language understanding. While both paradigms have their distinct strengths and limitations, their combination creates exciting opportunities and challenges that demand further exploration.

### 5.1 Strengths of KGs and LLMs in Complementary Roles

KGs provide structured and explicit representations of knowledge with well-defined relationships, which are critical to logical reasoning, inference, and maintaining high accuracy. Their ability to encode factual and hierarchical relationships makes them invaluable for tasks that require precision and structured querying. For instance, the accuracy of YAGO and Google's Knowledge Vault shows how reliable they are in operational use cases. However, their rigidity in construction, high cost of creation, and inability to adapt to changing information presents a significant challenge.

However, LLMs are amazing at capturing parametric knowledge from unstructured text corpora, with remarkable generalisation capabilities and natural language understanding capabilities. They can be used by downstream applications such as summarization, paraphrasing, or answering complex queries with ease. LLMs like GPT-4 and PaLM have demonstrated their ability to tackle diverse language-related tasks, including commonsense reasoning and semantic understanding, which are areas where KGs typically struggle. However, LLMs also face issues such as hallucinations, lack of interpretability, and difficulties with numerical reasoning or handling long-tail entities. The complementary nature of KGs and LLMs thus creates a compelling case for their integration. By leveraging the structured precision of KGs alongside the adaptive and contextual understanding of LLMs, hybrid systems can address the shortcomings of each technology individually.

### 5.2 Challenges in KG and LLM Integration

Despite their potential, integrating KGs and LLMs introduces complexities. Current methods for KG-enhanced LLMs often fall short in handling the dynamic nature of knowledge. For example, knowledge learned in pre-training of LLMs becomes frozen and stale, requiring retraining to be updated. Similarly, integrating KGs into inference pipelines can introduce knowledge distortion if not properly aligned, as observed with some input tokenization methods. Additionally, although LLMs have been used to augment KG construction and interpretation tasks, they suffer from numerical precision issues, underrepresentation of rare entities, and retrieving relationships between long-tail or infrequently appearing entities. Another critical issue is the computational resources needed to train and fine-tune LLMs for KG-related tasks. As KGs grow in complexity and scale, their integration with LLMs demands innovative approaches to enhance computational efficiency while maintaining accuracy.

### 5.3 Emerging Opportunities

The synergistic combination of LLMs and KGs opens new doors for applications across various domains. Retreival-Augmented methods, including RAG, have the potential to endow LLMs with knowledge from non-parametric sources during inference. Factual accuracy increases and hallucination decreases as a result. This happens especially for open-domain question-answering. Another avenue that leverages LLMs is for augmenting KG construction through the use of entity recognition, relationship extraction, and KG-to-text generation to facilitate the automation of constructing large-scale, high-quality knowledge graphs. This integration can transform the industries like healthcare, where patient data can be synthesized into actionable insights, or education, where knowledge can be structured and disseminated more effectively. In addition, LLMs can help address the high costs and scalability issues of KG construction. Their ability to process unstructured textual data and infer relationships makes them valuable tools for creating and updating KGs dynamically. This can enhance the representation of commonsense knowledge, which is traditionally challenging to capture in explicit KGs.

### 5.4 Ethical and Practical Considerations

The use of large language models (LLMs) in sensitive areas like healthcare brings up significant ethical issues. In these contexts, concerns about privacy—especially regarding protected health information (PHI)—and the risk of LLMs generating incorrect or potentially harmful content necessitate stricter safeguards. Incorporating knowledge graphs (KGs) as tools for interpretability can help provide a clear and traceable source of information to mitigate these risks. Additionally, LLMs must prioritize fairness and reduce bias, as they serve as assistants to KGs that may influence decision-making processes. Future research on the integration of KGs and LLMs should focus on addressing their limitations and enhancing their collaborative capabilities. Promising avenues include developing modular architectures that can update knowledge dynamically without retraining, improving the interpretability of hybrid systems, and optimizing computational efficiency for large-scale applications. Another intriguing direction involves innovative fusion techniques, such as integrating commonsense knowledge into LLMs

2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

or allowing KGs to dynamically query LLMs for reasoning tasks. In conclusion, the combination of KGs and LLMs presents a transformative opportunity for advancing AI systems. By tackling current challenges and seizing new opportunities, researchers can create systems that are not only accurate and scalable but also robust, interpretable, and adaptable to the changing needs of real-world applications.

## 6. Conclusion

Massive Language Models and Knowledge Graphs can change the scenario of artificial intelligence. The generative capabilities of LLMs combined with the knowledge-fact contextual relationship provided by KGs create more intelligent, context-aware, and effective systems. This union improves natural language understanding, reasoning, decision-making, and personalization. However, knowledge grounding, scalability, and interpretability issues remain challenges. The way forward includes innovative techniques such as hybrid models, attention mechanisms, and novel training strategies. The pathway laid down in this paper outlines future development LLM-KG systems are expected to usher in across various domains-including search engines, conversational AI, and knowledge-based systems. To realize their full potential for intelligent systems, deep knowledge, and contextual awareness, it will require continued collaboration between these two powerful technologies.

## References

[1] C. Rosset, C. Xiong, M. Phan, X. Song, P. Bennett, and S. Tiwary, “Knowledge-aware language model pretraining,” arXiv preprint arXiv:2007.00655, 2020.

[2] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Kuttler, M. Lewis, W.-t. Yih, T. Rockt ̈ aschel, S. Riedel, ̈ and D. Kiela, “Retrieval-augmented generation for knowledgeintensive nlp tasks,” in NeurIPS, vol. 33, 2020, pp. 9459–9474.

[3] Y. Zhu, X. Wang, J. Chen, S. Qiao, Y. Ou, Y. Yao, S. Deng, H. Chen, and N. Zhang, “Llms for knowledge graph construction and reasoning: Recent capabilities and future opportunities,” arXiv preprint arXiv:2305.13168, 2023.

[4] Z. Zhang, X. Liu, Y. Zhang, Q. Su, X. Sun, and B. He, “Pretrainkge: learning knowledge representation from pretrained language models,” in EMNLP Finding, 2020, pp. 259–266.

[5] A. Kumar, A. Pandey, R. Gadia, and M. Mishra, “Building knowledge graph using pre-trained language model for learning entity-aware relationships,” in 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON). IEEE, 2020, pp. 310–315.

[6] X. Xie, N. Zhang, Z. Li, S. Deng, H. Chen, F. Xiong, M. Chen, and H. Chen, “From discrimination to generation: Knowledge graph completion with generative transformer,” in WWW, 2022, pp. 162–165.

[7] Singh, S. K., Choudhary, S. K., Ranjan, P., & Dahiya, S. (2022). Comparative Analysis of Machine Learning Models and Data Analytics Techniques for Fraud Detection in Banking System. International Journal of Core Engineering & Management, 7(1), 64. ISSN 2348-9510.

[8] Rekha, P., Saranya, T., Preethi, P., Saraswathi, L., & Shobana, G. (2017). Smart Agro Using Arduino and GSM. International Journal of Emerging Technologies in Engineering Research (IJETER) Volume, 5.

[9] Suresh, K., Reddy, P. P., & Preethi, P. (2019). A novel key exchange algorithm for security in internet of things. Indones. J. Electr. Eng. Comput. Sci, 16(3), 1515-1520.

[10] Bharathy, S. S. P. D., Preethi, P., Karthick, K., & Sangeetha, S. (2017). Hand Gesture Recognition for Physical Impairment Peoples. SSRG International Journal of Computer Science and Engineering (SSRG-IJCSE), 6-10.

[11] Sujithra, M., Velvadivu, P., Rathika, J., Priyadharshini, R., & Preethi, P. (2022, October). A Study On Psychological Stress Of Working Women In Educational Institution Using Machine Learning. In 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-7). IEEE.

[12] Laxminarayana Korada, D. M. K., Ranjidha, P., Verma, T. L., & Mahalaksmi Arumugam, D. R. O. Artificial Intelligence On The Administration Of Financial Markets.

[13] Korada, L. (2024). Data Poisoning-What Is It and How It Is Being Addressed by the Leading Gen AI Providers. European Journal of Advances in Engineering and Technology, 11(5), 105-109.

[14] S. Li, X. Li, L. Shang, C. Sun, B. Liu, Z. Ji, X. Jiang, and Q. Liu, “Pre-training language models with deterministic factual knowledge,” in EMNLP, 2022, pp. 11 118–11 131.

[15] M. Kang, J. Baek, and S. J. Hwang, “Kala: Knowledge-augmented language model adaptation,” in NAACL, 2022, pp. 5144–5167.

[16] W. Xiong, J. Du, W. Y. Wang, and V. Stoyanov, “Pretrained encyclopedia: Weakly supervised knowledge-pretrained language model,” in ICLR, 2020.