2025 European Conference on Computing in Construction
CIB W78 Conference on IT in Construction
Porto, Portugal
July 14-17, 2025

INTELLIGENT KNOWLEDGE GRAPH QUESTION ANSWERING METHOD FOR HEALTH AND SAFETY HAZARD MANAGEMENT USING LARGE LANGUAGE MODELS

Chunmo Zheng[1,2], Wahib Saif[3], Yinqiu Tang[4], Xing Su[1], Mohamad Kassem[2]
[1]Zhejiang University, Hangzhou, China
[2]Newcastle University, Newcastle, United Kingdom
[3]Northumbria University, Newcastle, United Kingdom
[4]PowerChina Huadong Engineering Corporation Limited, Hangzhou, China

Abstract

Construction safety knowledge is often scattered across various unstructured or semi-structured sources, complicating its retrieval, reasoning, and application for effective safety management. This paper introduces a novel Knowledge Graph Question Answering (KGQA) method that leverages Large Language Models (LLMs) for intelligent QA over safety hazard knowledge. The method integrates an LLM-based assistant for natural language understanding (NLU), allowing users to query a domain-specific safety knowledge graph (KG) using natural language. To evaluate its effectiveness, a natural language query (NLQ) dataset is developed and used for benchmarking QA performance across prominent LLMs. Results demonstrate optimal accuracy in retrieving safety knowledge efficiently.

Introduction

Construction industry is inherently associated with safety risks that significantly impact the lives and well-being of construction workers. According to Health and Safety Executive (HSE), one-third of all occupational fatalities in the United Kingdom are attributed to the construction sector, highlighting the critical importance of safety management in the industry (Health and Safety Executive, 2024).

Devising comprehensive safety measures for risk scenarios is the cornerstone of safety management practices. Construction managers often face the challenge of formulating cohesive safety measures under time-sensitive and demanding conditions, necessitating access to reliable, context-specific safety knowledge for informed decision-making (Hong et al., 2024). However, the vast volume of safety hazard information, coupled with fragmented and inconsistent data storage, makes it difficult for managers to retrieve and apply relevant knowledge in a timely manner. Additionally, the formulation of safety measure heavily relies on safety manager's personal experience. Therefore, it is of critical importance to develop automated approach to retrieve safety hazard knowledge for safety measure generation decision support.

Natural language processing (NLP) techniques are widely used for various tasks in safety management including safety accident mining (Cheng et al., 2020; Ma & Chen, 2024), accident case retrieval (Kim & Chi, 2019; Zou et al., 2017) and safety measure generation (Zhou et al., 2023). Machine learning prediction models have emerged as powerful tools as they automatically learn data patterns from safety text to perform safety measure prediction. For example, decision trees, support vector machines, K-nearest neighbor and random forests are used to extract information and classify accidents from construction safety dataset. Deep learning approaches are reported to generate superior performance than ML-based method, owing to their ability to capture the complex non-linear relationships in safety dataset. For example, Baker et al. (2020) utilized deep learning models to predict injury precursors in safety incidents, while Mostofi and Togan (2023) developed a graph neural network-based model for risk assessment, highlighting the effectiveness of the self-attention mechanism in capturing the significance of connected accident records.

Despite these advances, two major challenges hinder the development of intelligent safety decision support systems. First, existing approaches lack a unified framework for safety management tasks. The prediction and recommendation of safety measures and accident cases require training multiple models on safety datasets, leading to inefficiencies and integration challenges. Second, while previous research has focused on improving safety measure prediction and accident case retrieval, current NLP approaches remain suboptimal in terms of retrieving and utilizing safety knowledge effectively. There is no flexible mechanism for safety managers to interact with and integrate construction safety knowledge bases seamlessly.

Safety hazard knowledge is predominantly stored in the form of unstructured or semi-structured text, scattered across various documents within construction companies. Therefore, efficiently extracting and organizing this knowledge to meet the needs of safety managers is of critical importance. Since safety scenarios and their corresponding measures or accident cases are often framed as question-answer pairs in safety management, formulating this task as a question-answering (QA) problem can significantly enhance knowledge retrieval efficiency.

Intelligent QA system has been proposed as a potential solution for querying safety database using natural language. Existing approaches employ deep learning-based frameworks for safety hazard measure QA. For example, Tian et al. (2023) developed a model integrating Bidirectional Encoder Representations from Transformers (BERT), Bidirectional Gated Recurrent Units (BiGRU), and self-attention mechanisms to extract semantic features and match safety-related questions with relevant safety measures. However, these QA systems primarily operate on unstructured and dispersed text, lacking a unified and standardized representation of safety knowledge. While ontology- and knowledge graph-based approaches have been widely adopted for organizing and structuring safety knowledge, their application in intelligent QA remains largely unexplored.

To address these gaps, this study introduces a KGQA framework that integrates state-of-the-art LLMs to facilitate safety hazard knowledge retrieval and reasoning. Specifically, the proposed framework constructs a safety hazard knowledge graph by integrating multi-source safety data, including risk registers and accident reports, into a structured ontology-based representation. The KG serves as a structured knowledge base that connects safety risks, construction activities, risk categories, risk factors, accident cases, regulations, and corresponding risk responses. By leveraging this structured knowledge, the KGQA framework employs a natural language understanding (NLU) module to process user queries, mapping them to relevant entities and relationships within the knowledge graph. The system supports multi-hop reasoning and constraint-based retrieval, allowing users to query safety risks, mitigation measures and relevant accident cases through natural language inputs. Furthermore, the integration of LLMs enhances the adaptability of the system, enabling dynamic query interpretation and response generation beyond predefined templates. By combining structured knowledge representation with advanced language modeling techniques, the KGQA framework provides an intelligent, context-aware decision support system for construction safety management, facilitating efficient knowledge retrieval, accident case recommendation, and safety measure generation.

Literature Review

Intelligent question answering system for safety hazard knowledge

Question answering system is basically information retrieval tools that allow users to search and retrieve documents or database with natural language queries. Traditional QA system based on templates maps the user query to predefined template, which can only handle very limited questions. The fixed pattern of the predefined template hampers the scalability and flexibility of the developed QA system. In comparison, deep learning techniques are widely used in developing intelligent QA systems with their ability to automatically learn, extract and understand text patterns in the user query. Zhong et al. (2020) developed a deep learning-based building regulation QA system by finetuning Bidirectional Encoder Representation from Transformer (BERT) on regulation QA pairs. The deep learning model facilitates an end-to-end methodology to improve the efficient information retrieval.

While deep learning models yield superior performance than traditional approaches, they require large amount of annotated training data which is notably limited in the construction industry. Transfer learning is introduced to adapt scarce datasets as it reduces training time and improves the prediction performance on small datasets. Wang et al. (2022) developed a Robustly Optimized BERT Pretraining Approach (RoBERTa)-based text classification method for building information-related query classification, focusing on the intention identification in the QA system.

Recent advances in generative models especially transformer-based large language models (e.g. Generative Pre-trained Transformer models) presents unprecedented opportunities for intelligent QA systems to operates without collection of large training or finetuning dataset. Combined with prompt engineering, LLMs demonstrate advanced in-context learning ability in NLP tasks such as information extraction and text classification. It is demonstrated that LLM-based assistant alone presents accurate and efficient performance in the BIM information search with natural language query (J. Zheng & Fischer, 2023).

Despite the developments in construction field QA models, there are less attention in the safety knowledge base question answering. Recently, researchers have identified the need to extract safety hazard information accurately and efficiently from unstructured text. Tian et al. (2023) builds a deep learning network fusing BERT, Bidirectional Gated Recurrent Unit (BiGRU), and Self-attention mechanism to extract text semantic features for supporting intelligent safety hazard knowledge QA.

As safety hazard knowledge in enterprises is often presented in a semi-unstructured format across safety databases, risk registers, and other sources, there is a pressing need to develop intelligent QA systems for safety knowledge bases—an area that has received limited attention to date.

Ontology and graph-based approaches for safety management

Unstructured data in construction poses challenges for the systematic safety knowledge retrieval and management. Recent research aims to represent safety knowledge in consistent and machine-readable format. A knowledge graph represents data in a graph structure, where nodes correspond to entities and edges define their relationships (Hogan et al., 2021). By transforming textual information into a graph-based format, it becomes possible to perform deeper semantic reasoning, capturing the complex associations between different aspects of safety knowledge. Wu et al. (2022) leveraged a knowledge graph modeling approach to extract textual entities from accident reports, establishing a systematic organization and visualization method for safety knowledge.

Integrating graph-based knowledge representation with deep learning methods has emerged as a promising solution for safety management. Deep learning models excel at identifying complex text patterns, while graph-based representations effectively preserve the interconnected relationships within safety data. Moreover, the explicit structure of knowledge graphs enhances the interpretability of deep learning models, addressing the black-box nature of traditional deep learning approaches. For instance, Pan et al. (2022) developed a graph-based deep learning framework to classify accident reports based on accident type, injury type, and affected body parts. This hybrid approach demonstrates the potential of combining graph structures with deep learning to improve accident analysis and safety management in the construction industry.

Pretrained large language models

Large language models (LLMs) have revolutionized NLP-based methods by enabling human-like language generation and instruction following. These models are pretrained on vast amounts of text data using sophisticated transformer architectures and attention mechanisms, equipping them with advanced in-context learning capabilities.

One of the most prominent LLMs is the GPT series, such as GPT-3 (Radford et al., 2018) with 175 billion parameters.), which features 175 billion parameters. ChatGPT, an extension of the GPT foundation model, is further refined using reinforcement learning with human feedback (RLHF) to enhance its ability to follow human instructions (Ouyang et al., 2022).

The substantial improvements in natural language understanding observed in these models have introduced a new paradigm—interacting with LLMs through natural language, a technique known as prompt engineering. Traditional NLP approaches rely heavily on task-specific training data, whereas the prompt-based paradigm enables zero-shot or few-shot learning by reframing tasks as prompt-completion problems (Brown et al., 2020). A user-provided instruction, known as a prompt, guides the model's responses to align with human preferences. Prompt engineering has demonstrated superior performance across various NLP tasks, including information retrieval (Maoro et al., 2024) and question answering (Maharjan et al., 2024).

In construction domain, the potential of LLMs for automating construction tasks has recently gained attention in both research and industry. Zheng et al. (2022) highlight the necessity of adopting pretrained LLMs for NLP tasks in the architecture, engineering, and construction (AEC) sector, such as automated compliance checking, and they developed a domain-specific dataset for this purpose.

Despite their potential, LLMs present several challenges when applied to domain-specific tasks, including hallucinations and the inability to retrieve dynamic data. Hallucinations refer to the generation of incorrect or misleading responses. To address this issue, retrieval-augmented generation (RAG) has been introduced, integrating LLMs with external knowledge retrieval mechanisms. Wong et al. (2024) developed an RAG-based approach using vector database to enhance the LLM-based automated contract risk identification. Lee et al. (2024) compare RAG pipeline with fine-tuned LLMs for safety knowledge retrieval and found that while the RAG approach improved performance by 21.5%, fine-tuning enhanced it by 26%.

However, the integration of LLMs with graph-based approaches in the construction domain remains largely unexplored. Given the vast amount of scattered unstructured and semi-structured knowledge in this field, representing domain knowledge in structured formats such as knowledge graphs (KGs) and leveraging LLM-based assistants for KG interaction present a promising direction. For instance, KG-based LLM agents have been introduced for complex domain-specific question answering and have demonstrated significant improvements over traditional models (Su et al., 2024).

Methodology

This paper presents a novel KGQA framework for addressing safety knowledge-related queries, as shown in Fig. 1. The framework consists of two main steps/components: a) knowledge graph construction, b) natural language understanding (NLU) for safety knowledge QA.

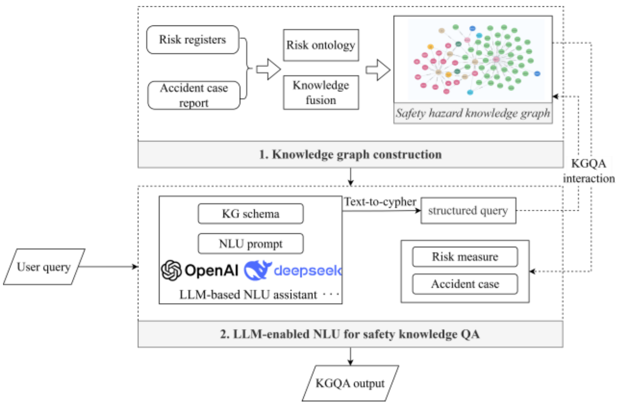

Figure 1: Framework overview

The knowledge graph construction module integrates multi-source safety knowledge into a standardized ontology-based graph, forming the backbone for structured retrieval. The NLU module transforms natural language queries into structured formats, enabling efficient mapping to ontology-defined concepts and seamless interaction with the knowledge graph. These components collectively operate as an information retrieval pipeline, augmented by LLM-based prompting techniques. The following sections detail the knowledge graph construction and the NLU-driven QA process.

Knowledge graph construction

The KG is designed to represent safety hazard knowledge in the context of safety management. This knowledge encompasses safety hazards, risks, related risk categories, contributing factors, and corresponding responses. To ensure consistency and interoperability, a safety ontology has been developed to serve as a unified schema for the safety knowledge graph.

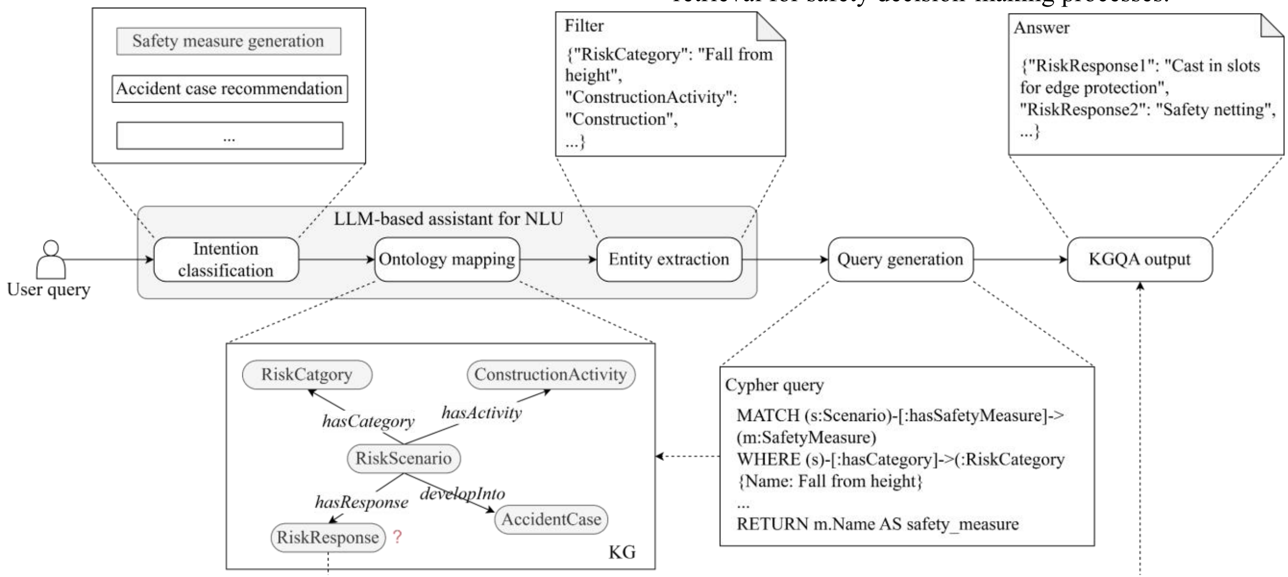

Figure 3: KGQA pipeline through natural language understanding

As shown in Figure 2, the ontology comprises several core classes, including RiskScenario, RiskCategory, ConstructionActivity, ConstructionObject, Agent, RegulationAndStandard, RiskFactor, RiskResponse, RiskConsequence and AccidentCase. The RiskScenario class describes the scenario that might cause safety accident or incident. Each risk scenario is categorized under a specific RiskCategory and is caused by one or more RiskFactors.

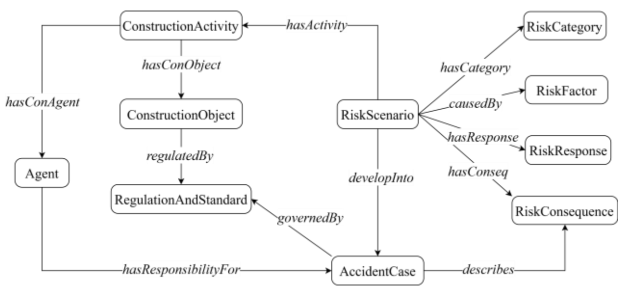

Figure 2: Ontology for safety risk hazard knowledge

Safety risks often emerge from construction activity safety hazards, which are further linked to specific ConstructionObjects. In cases where safety risks lead to accidents, they are recorded as AccidentCases. These cases are governed by relevant RegulationAndStandard entities, ensuring compliance with safety guidelines. Moreover, Agents (responsible parties) are associated with accident cases, and the consequences of risks are documented under the RiskConsequence class. To mitigate or control risks, corresponding RiskResponses are implemented.

This interconnected structure enables the knowledge graph to comprehensively capture and represent safety hazard information, facilitating effective management and retrieval for safety decision-making processes.

The process of constructing knowledge graph is to organize the semi-structured construction safety risk database into structured schema defined in the safety hazard ontology. The data source for constructing safety hazard KG is collected from risk database by Health and Safety Executive (HSE). The risk registers detail defined risk categories, associated construction activities, locations, risk assessments, and corresponding risk responses for each category. Accident reports provide descriptions of accidents or incidents and link them to relevant risk categories and construction activities. These descriptions also include judgments on the causes of the accidents, associated regulations, and the resulting consequences (e.g., fines or identification of responsible parties).

Natural language understanding for safety knowledge QA

The goal of this module is to answering user's natural language query related to safety hazard knowledge through an NLU pipeline. Traditional approaches, which primarily rely on NLP-based named entity recognition (NER) and predefined templates to map queries into structured formats, often suffer from rigidity and an inability to accommodate diverse query patterns. To overcome these limitations, this study integrates a dynamic LLM-based assistant, enabling flexible and context-aware information retrieval from the safety knowledge base. The proposed framework facilitates querying from any node within the knowledge graph and dynamically retrieving related nodes. Furthermore, it supports multi-hop reasoning and multi-constraint QA, both of which are essential for safety hazard analysis given the intricate interconnections among risk attributes.

Table 1: Example NLQ dataset for safety hazard knowledge QA

Intention classification	NLQ	Risk Category	Construction Activity	Risk Location	Example of matched answer
Safety measure generation	What protective actions can be taken to mitigate the risk of falling from heights due to open or exposed edges of slabs during high-level construction of in-situ concrete?	Falls from height	Construction	High level	Cast in slots for edge protection; Install walkway routes to avoid edge; …
	What protective measures should be in place to ensure safety from gas, fumes, vapors, or oxygen deficiency risks in confined spaces while performing maintenance on below-ground drainage and pipework?	Confined space	Maintenance	Below ground	Select products without silica content. Refer to product specification, for walls, floors and ceiling finishes; Design in easily accessible termination and connection points.
Accident case recommendation	Which accident cases are related to the risk scenario involving Hazardous Substances Exposure during Maintenance in the site work area?	Hazardous substances exposure	Maintenance	Site wide	Two companies have been fined after asbestos was disturbed during refurbishment work; …
	Which accident cases involve fall through fragile material during install construction at roof level?	Falls from height	Construction	Roof level	A roofing company has been fined after an employee fell through a skylight; …

As illustrated in Figure 3, the overall LLM-based NLU pipeline comprises three key stages. The first stage, intention classification, involves categorizing the user's query into distinct high-level tasks, such as safety measure generation or accident case recommendation, where different intents correspond to different search paths within the knowledge graph. Once the intent is determined, the ontology mapping process aligns the query with relevant ontology concepts, ensuring consistency with structured entities in the knowledge base. For instance, a query related to safety measure generation typically necessitates retrieving RiskResponse entities based on constraints such as RiskCategory, ConstructionActivity, and RiskLocation. In the entity extraction phase, the system identifies key entities within the query, such as a specific RiskCategory (e.g., "Fall from height") or ConstructionActivity (e.g., "Construction"), which are then used to generate structured filters for querying the safety database.

The prompt designed for the NLU module is shown in Figure 4. User query first goes through the intention classification prompt, which identify the intention, namely the target entity in the KG. For example, in the user query “What protective actions can be taken to mitigate the risk of falling from heights due to open or exposed edges of slabs during high-level construction of in-situ concrete?” The intention is classified into “safety measure generation”, and the target entity in the KG belongs to “SafetyMeasure” class. The result determines the ontology mapping schema provided in the ontology mapping stage. In the entity extraction stage, the ontology schema, together with user query are embedded in the prompt template. And the LLMs are prompted to identify instances from the user query under each class in the ontology schema.

Following the NLU module, the query generation phase embeds the extracted entities into structured queries, such as Cypher queries for Neo4j. These structured queries facilitate entity matching within the knowledge graph and retrieve the corresponding answers. For example, the Cypher query depicted in Fig. 3 retrieves the appropriate safety measures for the risk of "Fall from height" during construction activities from the safety knowledge base.

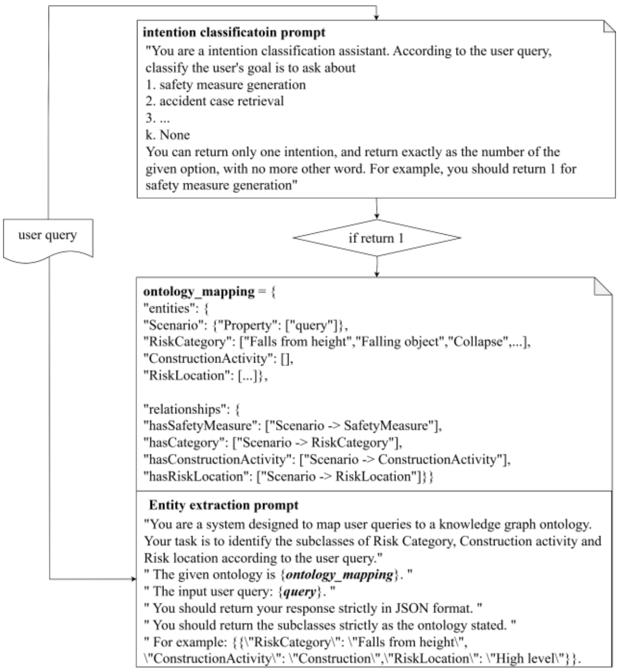

Figure 4: Prompt for NLU module in QA system

Case study

Data preparation

Natural language query (NLQ) dataset for NLU in the safety domain is limited which hampers the development of intelligent QA systems. To bridge this gap, we developed a safety knowledge QA dataset based on the HSE database, consisting of 192 natural language queries labeled according to the safety domain KG schema. These queries typically describe a risk scenario and request either risk control measures or related accident cases from the database.

For evaluation in the case study, three key constraints are considered: RiskCategory, ConstructionActivity, and RiskLocation. Each query is annotated with these labels from the KG ontology. The Risk category include 11 labels: "Falls from height", "Falling object", "Collapse", "Fire and explosion", "Load related", "Plant, vehicles and mobile equipment", "Noise", "Confined space", "Hazardous substance ill health", "Electrical Injury" and “Exposure to ionising radiation”. Additionally, the dataset incorporates 4 construction activities and 8 risk locations.

Table 1 presents a sample of this dataset, illustrating various query types, their associated risk scenario constraints, and the matched responses.

Performance results

To evaluate the effectiveness of our method in a multi-constraint scenario, we tested the QA system under two different conditions. In the QA with 3 constraints setting, entities from RiskCategory, ConstructionActivity, and RiskLocation are extracted and used as structured query filters, while in the QA with 2 constraints setting, only RiskCategory and ConstructionActivity are extracted.

As a baseline, we utilize a LangChain-based framework, which follows a similar text-to-query pipeline to interact with the knowledge graph using natural language queries. The QA system's accuracy is assessed based on the hit rate, where a successful query retrieval from the KG is counted as a hit. The accuracy is defined as follows.

$$\text{Acc} = \frac{Number\ of\ hits\ among\ all\ queris}{Totoal\ number\ of\ queries} \quad (1)$$

The performance results of the QA system across three different language models using both the baseline method and our approach are summarized in Table 2. The LangChain-based framework fails to retrieve any results from the KG in all cases, yielding a 0% hit rate across all models. This is due to the inaccurate mapping from natural language query to ontology. In contrast, our approach demonstrates significant improvements, achieving high accuracy in both 2-constraint and 3-constraint scenarios. The NLU module in our appoach is able to map the natural language query to KG nodes, instead of returning unstable entities. Therefore, langchain is more suitable for KGQA with simple and fixed schema, while ours can be adopted in more complex domain QA tasks.

For the 2-constraint setting, where only RiskCategory and ConstructionActivity are extracted, our method achieves near-perfect accuracy across all models: 100% with GPT-4o, 98.44% with DeepSeek-V3, and 92.71% with Claude-3.5. While in the 3-constraint setting, where an additional constraint is included, the accuracy slightly decreases but remains consistently high. DeepSeek-V3 achieves the highest accuracy at 91.67%, followed by Claude-3.5 (92.19%) and GPT-4o (89.58%).

Overall, these results demonstrate that our method significantly outperforms the LangChain-based framework and effectively retrieves structured answers from the KG.

Table 2: Testing results for safety knowledge QA

Model	Langchain	2 - constraint (ours)	3 - constraint (ours)
gpt-4o	0	100%	89.58%
deepseek-v3	0	98.44%	91.67%
Claude-3.5	0	92.71%	92.19%

Discussion

This paper demonstrates an intelligent KGQA system for safety hazard knowledge, enabled by LLMs. The core component of our system is the NLU module, which is responsible for classifying user intent and mapping it to a structured safety ontology. The generated filters from ontology mapping are then used to query the safety KG for relevant entities. Our approach introduces an efficient, flexible, and dynamic QA framework that facilitates multiple safety management tasks, including safety measure generation and accident case retrieval.

Compared to traditional QA systems designed for safety hazard knowledge, our approach offers three key advantages. First, instead of treating safety knowledge as isolated pieces of information, our safety KG models interconnected relationships between safety scenarios, risk categories, and associated construction activities, effectively integrating scattered safety knowledge across the industry. It also captures the causes (RiskFactor), responses (RiskResponse), and consequences (AccidentCase) of safety risks, allowing for a more comprehensive representation of the dynamic flow of safety risks. This enriched knowledge base enables a more intelligent and holistic safety management system. Second, our approach provides greater flexibility in user interaction with the safety knowledge base through the NLU module. By leveraging LLMs, our system can interpret and understand user queries based on the general reasoning capabilities of LLMs, eliminating the need for extensive training on domain-specific data. Third, our approach allows for dynamic updates to the knowledge graph (KG) without the need for extensive retraining, unlike deep learning models that require additional training on new data to maintain performance. This scalability and adaptability make our method particularly well-suited for real-world safety management applications, where safety regulations, risk factors, and industry practices are constantly evolving.

However, this research has several limitations. Firstly, the query dataset used to evaluate the QA system is still limited. To ensure the robustness and generalizability of the proposed approach, future research should incorporate a larger and more diverse dataset. Expanding the dataset with additional data sources from industry practice could enhance the dataset's representativeness and practical relevance. Moreover, developing synthetic data generation methods presents a promising direction for advancing KGQA system development.

Secondly, this paper relies on a predefined ontology for KG construction and query transformation. Although ontologies provide rich semantics, their practical implementation faces challenges within the construction safety risk domain. The dynamic nature of risk scenarios and evolving industry practices limits ontology reusability and scalability. Future research should explore methods to test, refine, and enhance the adaptability of ontology to accommodate new safety regulations and changing construction context. Consequently, the development of automated approaches for extracting structured information from unstructured texts is critical for efficient and scalable KG population.

Thirdly, the evaluation of the QA system in this study primarily focuses on the NLU stage, specifically entity extraction accuracy based on ontology mapping. While this is an important aspect, it does not provide a full assessment of the system's practical performance in real-world construction settings. Future research should conduct a comprehensive evaluation of the final QA system, including quality of the outputs and integration with existing safety management platforms. Such evaluation is essential to validate the system's effectiveness in delivering accurate and actionable safety risk management knowledge.

Conclusion

This study proposes a KGQA framework that integrates LLMs to enhance safety knowledge retrieval and safety management. By structuring safety data into a safety KG, the system enables efficient querying, multi-hop reasoning, and context-aware responses. The NLU module maps user queries in natural language to structured knowledge, facilitating safety measure generation and accident case retrieval.

The proposed approach addresses key limitations of traditional safety QA systems by providing a holistic representation of safety risks, improving query flexibility, and ensuring scalability through LLM-enabled assistant. These features make the system well-suited for real-world safety management applications.

Reference

Baker, H., Hallowell, M. R., & Tixier, A. J.-P. (2020). Automatically learning construction injury precursors from text. Automation in Construction, 118, 103145. https://doi.org/10.1016/j.autcon.2020.103145

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.

Cheng, M.-Y., Kusoemo, D., & Gosno, R. A. (2020). Text mining-based construction site accident classification using hybrid supervised machine learning. Automation in Construction, 118, 103265. https://doi.org/10.1016/j.autcon.2020.103265

Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutierrez, C., Gayo, J. E. L., Kirrane, S., Neumaier, S., Polleres, A., Navigli, R., Ngomo, A.-C. N., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge Graphs. ACM Computing Surveys, 54(4), 1–37. https://doi.org/10.1145/3447772

Hong, E., Lee, S., Kim, H., Park, J., Seo, M. B., & Yi, J.-S. (2024). Graph-based intelligent accident hazard ontology using natural language processing for tracking, prediction, and learning. Automation in Construction, 168, 105800. https://doi.org/10.1016/j.autcon.2024.105800

Kim, T., & Chi, S. (2019). Accident Case Retrieval and Analyses: Using Natural Language Processing in the Construction Industry. Journal of Construction Engineering and Management, 145(3), 04019004.

https://doi.org/10.1061/(ASCE)CO.1943-7862.0001625

Lee, J., Ahn, S., Kim, D., & Kim, D. (2024). Performance comparison of retrieval-augmented generation and fine-tuned large language models for construction safety management knowledge retrieval. Automation in Construction, 168, 105846. https://doi.org/10.1016/j.autcon.2024.105846

Ma, Z., & Chen, Z.-S. (2024). Mining construction accident reports via unsupervised NLP and Accimap for systemic risk analysis. Automation in Construction, 161, 105343. https://doi.org/10.1016/j.autcon.2024.105343

Maharjan, J., Garikipati, A., Singh, N. P., Cyrus, L., Sharma, M., Ciobanu, M., Barnes, G., Thapa, R., Mao, Q., & Das, R. (2024). OpenMedLM: Prompt engineering can out-perform fine-tuning in medical question-answering with open-source large language models. Scientific Reports, 14(1), Article 1. https://doi.org/10.1038/s41598-024-64827-6

Maoro, F., Vehmeyer, B., & Geierhos, M. (2024). Leveraging Semantic Search and LLMs for Domain-Adaptive Information Retrieval. In A. Lopata, D. Gudonienė, & R. Butkienė (Eds.), Information and Software Technologies (pp. 148–159). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-48981-5_12

Mostofi, F., & Toğan, V. (2023). Construction safety predictions with multi-head attention graph and sparse accident networks. Automation in Construction, 156, 105102. https://doi.org/10.1016/j.autcon.2023.105102

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P. F., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems, 35, 27730–27744. https://proceedings.neurips.cc/paper_files/paper/2022/hash/b1efde53be364a73914f58805a001731-Abstract-Conference.html

Pan, X., Zhong, B., Wang, Y., & Shen, L. (2022). Identification of accident-injury type and bodypart factors from construction accident reports: A graph-based deep learning framework. Advanced Engineering Informatics, 54, 101752.

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.

Su, X., Wang, Y., Gao, S., Liu, X., Giunchiglia, V., Clevert, D.-A., & Zitnik, M. (2024). Knowledge Graph Based Agent for Complex, Knowledge-Intensive QA in Medicine (arXiv:2410.04660). arXiv. http://arxiv.org/abs/2410.04660

Tian, D., Li, M., Ren, Q., Zhang, X., Han, S., & Shen, Y. (2023). Intelligent question answering method for construction safety hazard knowledge based on deep semantic mining. Automation in Construction, 145, 104670. https://doi.org/10.1016/j.autcon.2022.104670

Wang, N., Issa, R. R. A., & Anumba, C. J. (2022). Transfer learning-based query classification for intelligent building information spoken dialogue. Automation in Construction, 141, 104403. https://doi.org/10.1016/j.autcon.2022.104403

Wong, S., Zheng, C., Su, X., & Tang, Y. (2024). Construction contract risk identification based on knowledge-augmented language models. Computers in Industry, 157–158, 104082. https://doi.org/10.1016/j.compind.2024.104082

Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., & Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. Automation in Construction, 134, 104059. https://doi.org/10.1016/j.autcon.2021.104059

Zheng, J., & Fischer, M. (2023). Dynamic prompt-based virtual assistant framework for BIM information search. Automation in Construction, 155, 105067. https://doi.org/10.1016/j.autcon.2023.105067

Zheng, Z., Lu, X.-Z., Chen, K.-Y., Zhou, Y.-C., & Lin, J.-R. (2022). Pretrained domain-specific language model for natural language processing tasks in the AEC domain. Computers in Industry, 142, 103733. https://doi.org/10.1016/j.compind.2022.103733

Zhong, B., He, W., Huang, Z., Love, P. E. D., Tang, J., & Luo, H. (2020). A building regulation question answering system: A deep learning methodology. Advanced Engineering Informatics, 46, 101195. https://doi.org/10.1016/j.aei.2020.101195

Zhou, H., Tang, S., Huang, W., & Zhao, X. (2023). Generating risk response measures for subway construction by fusion of knowledge and deep learning. Automation in Construction, 152, 104951. https://doi.org/10.1016/j.autcon.2023.104951

Zou, Y., Kiviniemi, A., & Jones, S. W. (2017). Retrieving similar cases for construction project risk management using Natural Language Processing techniques. Automation in Construction, 80, 66–76. https://doi.org/10.1016/j.autcon.2017.04.003