

大语言模型赋能的地质找矿知识图谱与问答模型构建

张宝一¹, 唐嘉成¹, 张彤蕴², 王宾海², 史与正², 詹庆忠³, 方振西¹,

KABLAN Or Aimon Brou Koffi¹, 马凯^{3,*}

1.中南大学有色金属成矿预测与地质环境监测教育部重点实验室/地球科学与信息物理学院, 长沙 410083

2.湖南省地质地理信息所(湖南省地质大数据中心), 长沙 410007

3.三峡大学三峡数智研究院, 宜昌 443002

摘要: 当前地质找矿领域的大语言模型应用面临着专业知识不足、数据隐私安全和模型幻觉等问题, 同时大语言模型在地质找矿领域应用中仍缺乏高效快捷的知识推荐手段。本研究提出了知识图谱与检索增强生成相结合的 KG-RAG (Knowledge graph Retrieval-Augmented Generation) 框架, 以大语言模型为工具, 在地质本体约束下实现了地质找矿知识图谱的自动化抽取和结构化表达, 同时利用知识图谱的多跳检索算法实现检索内容的深度与广度优化, 实现了地质找矿智能知识问答模型。实验结果表明: KG-RAG 在准确率、召回率和可信度(F1-score)上分别取得的 0.807、0.833 和 0.819, 在知识图谱构建任务相比大语言基模型 GLM4-9B 的直接知识抽取, 分别取得了约 50%、8%和 29%的提升; 在问答任务中, KG-RAG 召回率和准确率上分别取得了 0.917 和 0.88, 相比文档向量检索增强生成方法分别取得了约 24%和 22%的提升。KG-RAG 在知识图谱构建与智能问答两方面均表现出了较好的性能, 能够有效从地质资料中进行地质找矿知识收集与表达, 支持地质工作者的地质调查与找矿预测工作, 本研究为大语言模型与

基金项目: 地球深部探测与矿产资源勘查国家科技重大专项(2024ZD1001201), 湖南省地质院重大科研项目(HNGSTP202301) **第一作者简介:** 张宝一(1979—), 男, 副教授, 博士, 博士生导师, 从事地理信息科学研究及相关教学工作, 主要从事地质大数据智能挖掘研究。Email: zhangbaoyi@csu.edu.cn, ORCID: 0000-0001-6075-9359

通讯作者简介: 马凯(1980—), 男, 教授, 博士, 博士生导师, 从事计算机科学与技术研究及相关教学工作, 主要从事知识图谱研究。Email: makai@ctgu.edu.cn, ORCID: 0000-0001-5432-1166, 通信地址: 湖北省宜昌市大学路8号 三峡数智研究院; 邮政编码: 443002

知识图谱的联合应用提供了借鉴。

关键词：地质找矿；知识图谱；大语言模型；检索增强生成；地质本体；思维链
中图分类号： P628+.3 **收稿日期：** 2025-07-22

Knowledge Graph and Question-Answering Model for Geological Prospecting Empowered by Large Language Models

ZHANG Baoyi¹, TANG Jiacheng¹, ZHANG Tongyun², WANG Binhai², SHI Yuzheng², ZHAN Qingzhong³, FANG Zhenxi¹, KABLAN Or Aimon Brou Koffi¹,
MA Kai^{3,*}

1. Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment Monitoring (Ministry of Education) / School of Geosciences and Info-Physics, Central South University, Changsha 410083
2. Geological and Geographic Information Institute of Hunan Province / Hunan Provincial Center of Geological Big Data, Changsha 410007
3. Three Gorges Digital Intelligence Institute, China Three Gorges University, Yichang 443002

Abstract: Current applications of large language models (LLMs) in geological prospecting face challenges including insufficient domain expertise, data privacy concerns, and model hallucinations. Furthermore, there remains a lack of efficient and rapid knowledge recommendation methods for LLMs in this field. This study proposes a KG-RAG (knowledge graph-embedded retrieval-augmented generation) framework that automates the extraction and structured representation of geological prospecting knowledge under the constraints of a geological ontology, leveraging large LLMs as tools. It further employs multi-hop retrieval algorithms within the knowledge graph to enhance the depth and breadth of retrieved content, thereby

constructing an intelligent question-answering model for geological prospecting. Experimental results demonstrate that KG-RAG achieved scores of 0.807 (Precision), 0.833 (Recall), and 0.819 (F1-score) in knowledge graph construction tasks. Compared to direct knowledge extraction using the baseline LLM (GLM4-9B), KG-RAG delivers improvements of approximately 50% (Precision), 8% (Recall), and 29% (F1-score), respectively. In question-answering tasks, KG-RAG achieved 0.917 (Recall) and 0.88 (Precision), outperforming document vector-embedded retrieval-augmented generation methods by approximately 24% (Recall) and 22% (Precision), respectively. KG-RAG exhibits superior performance in both knowledge graph construction and intelligent question-answering. It effectively collects and represents geological prospecting and mineral exploration knowledge, providing a valuable reference to geologists for the combined application of LLMs and knowledge graphs.

Key words: Geological Prospecting; Knowledge Graph (KG); Large Language Models (LLMs); Retrieval Augmented Generation (RAG); Geological Ontology; Chain of Thought (CoT).

1 引言

大数据与人工智能技术的迅猛发展为科学研究范式带来了深刻变革。在地学领域，地质勘探作为典型的数据密集型科学工作，其数据的收集、更新、维护、挖掘与分析已成为地学研究的重要内容(王登红等, 2015, 周永章等, 2021, Zhang et al., 2023)。当前人工智能逐步演进到大语言模型 (Large language models, LLMs) 阶段, LLMs 已经在各种自然语言处理任务中取得了巨大成功, 表现出了令人印象深刻的 Zero-shot 和 Few-shot 性能表现, 其代表模型有 GPT-3、GPT-4、LLaMa、ChatGLM 和 DeepSeek 等(Floridi and Chiriatti, 2020, Katz et al., 2024, 徐春等, 2024, Dreyer, 2025, Xie et al., 2025)。受益于大量的模型参数和广泛的预训练语料库, LLMs 可以通过提示词工程和上下文学习来解决复杂的推理任务(Zhang et al., 2024a, 佟彬等, 2025, 吴华意等, 2025)。在地学领域, LLMs 有助于

地学知识理解与问题解决, Hu et al. (2023)将位置描述的地理知识和 LLMs 进行结合,有效提取了自然灾害的位置信息; Memduhoglu et al. (2024)提出了使用 LLMs 来解读 OpenStreetMap 标签的方法,通过 LLMs 嵌入捕捉深度语境理解,结合物理和空间指标来对城市建筑功能进行分类; Zhang et al. (2024b)开发了一款融合了 LLMs 和地理空间数据的人工智能助手 GeoGPT,结合知识链和扩展工具实现了空间分析和文本报告生成; De Almeida et al. (2025)利用 GPT-4 和微软 Azure 创建了一个文本转视频生成工具,以增强采矿中的安全报告; Fu et al. (2025)开发了云南省地质和矿产勘探智能问答系统 GeoMinLM,通过知识图谱整合专家知识,减少幻觉现象并提高专业性; Li et al. (2025)开发了面向地理空间数据问答系统 GDQA,将地理空间数据知识图谱 GDKG 与 LLMs 相结合,使 LLMs 能够在 GDKG 上进行推理和探索,从而增强其理解和回答复杂地球科学查询的能力; 吴若玲和郭旦怀 (2025)开展了 LLMs 在空间认知问题上的测试标准研究,最终形成一套具备 LLMs 空间认知能力的测试标准框架 SRT4LLM 和测试流程。然而需要注意的是,LLMs 在地学领域应用过程中,也面临着专业知识不足、数据隐私安全和模型幻觉等挑战。

为解决 LLMs “幻觉”问题,检索增强生成 (Retrieval-Augmented Generation, RAG) 技术应运而生。RAG 以外挂知识库的方式,能有效避免隐私数据泄露,同时在内容生成过程中,会检索大规模知识集合,生成的内容更丰富,有效增强 LLMs 的泛化能力(Chen et al., 2025, 江双五等, 2025)。目前主流使用的 RAG 框架有 LangChain、Langchain-ChatChat、LlamaIndex 和 RAGFlow(Church et al., 2024, Vidivelli et al., 2024, Hosseini and Seilani, 2025)。随着 DBpedia(Bizer et al., 2009)、多模态地球科学知识图谱 GAKG(董少春等, 2020)、地球科学数据语料库 ESDC(Li et al., 2024)等国内外地学领域开放知识库的迅猛增长,知识图谱(Knowledge Graph, KG)已广泛用于支持语义网络的各类地学应用(Ma, 2022, 邱芹军等, 2023, 郭飞等, 2024, 王成彬等, 2024)。Jiang et al. (2020)提出了一种知识增强对话生成模型 KADG,通过预先优化每个知识候选对象并结合推断出的话题转换线索来进行一跳主题推理,并充分利用所选知识生成上下文连贯且富有知识性的回答; Liang et al. (2025)提出了基于图的检索增强生成

框架 GeoGraphRAG，从外部知识库中检索相关子图，并将结构和语义信息注入到 LLMs 中用于地理空间建模。然而，地质找矿领域的知识推荐需要一套以成矿概念与数据混合驱动的背景信息组织方法，当前的文档向量 RAG 和知识库 RAG 并不能完全胜任这份工作。

面向地质找矿领域的大语言模型应用，本研究提出结合大语言模型和知识图谱的检索增强生成框架 KG-RAG (Knowledge Graph Retrieval-Augmented Generation)来实现地质领域的知识抽取与知识应用，解决地质找矿领域的知识图谱构建中依赖人工标注、扩展性差等问题，同时针对地质找矿应用的知识碎片化与智能化不足的问题，通过将结构化的知识图谱嵌入 RAG 增强生成机制中，实现地质找矿信息的深层次和多粒度检索。

2 KG-RAG 框架

针对地质知识图谱构建与找矿问答推荐任务，本研究融合大语言模型、知识图谱和检索增强生成构建了 KG-RAG 框架，如图 1 所示，包括：利用语料数据、大语言模型和提示词工程实现知识图谱的构建，以及通过大语言模型和检索增强生成和知识图谱信息实现问答推荐。

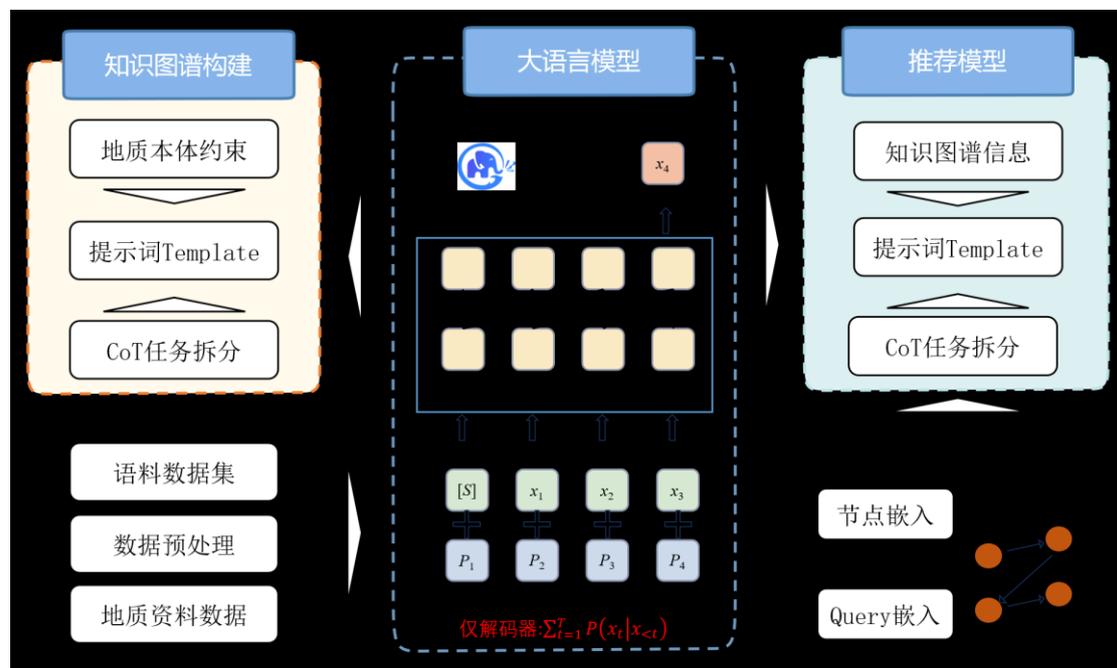


图 1 基于知识图谱的检索增强生成框架 KG-RAG

Fig. 1 KG-RAG: knowledge graph-embedded retrieval-augmented generation framework

for geological prospecting.

由于地质资料的涉密性，在知识图谱构建与地质找矿知识推荐中，采用了本地部署 GLM4-9B 大语言模型辅助来实现两类任务。两类任务均采用思维链 (Chain of Thought, CoT) 进行任务逻辑拆分，以保证复杂任务能够有效进行分解，从而使得 9B 参数的 GLM4 大语言模型能够有效执行单次逻辑推理分析任务。同时对提示词模板进行参数化，以保证在知识图谱构建中的地质本体模式层约束，在知识推荐中提示词参数化则为知识图谱节点提供增强检索生成的入口连接。

2.1 大语言模型辅助构建知识图谱

地质找矿知识图谱本体模式层是对地质找矿领域概念体系的形式化规范，通过概念节点与关系边构建语义网络，为数据层实例的属性分解和关联推理提供理论框架。本研究从地理、地层、岩体、构造、地质年代、成矿要素、地球物理、地球化学等 8 类要素进行了总结，通过先验知识厘定各要素的层次关系、定义属性关系及语义关系，如图 2 所示。

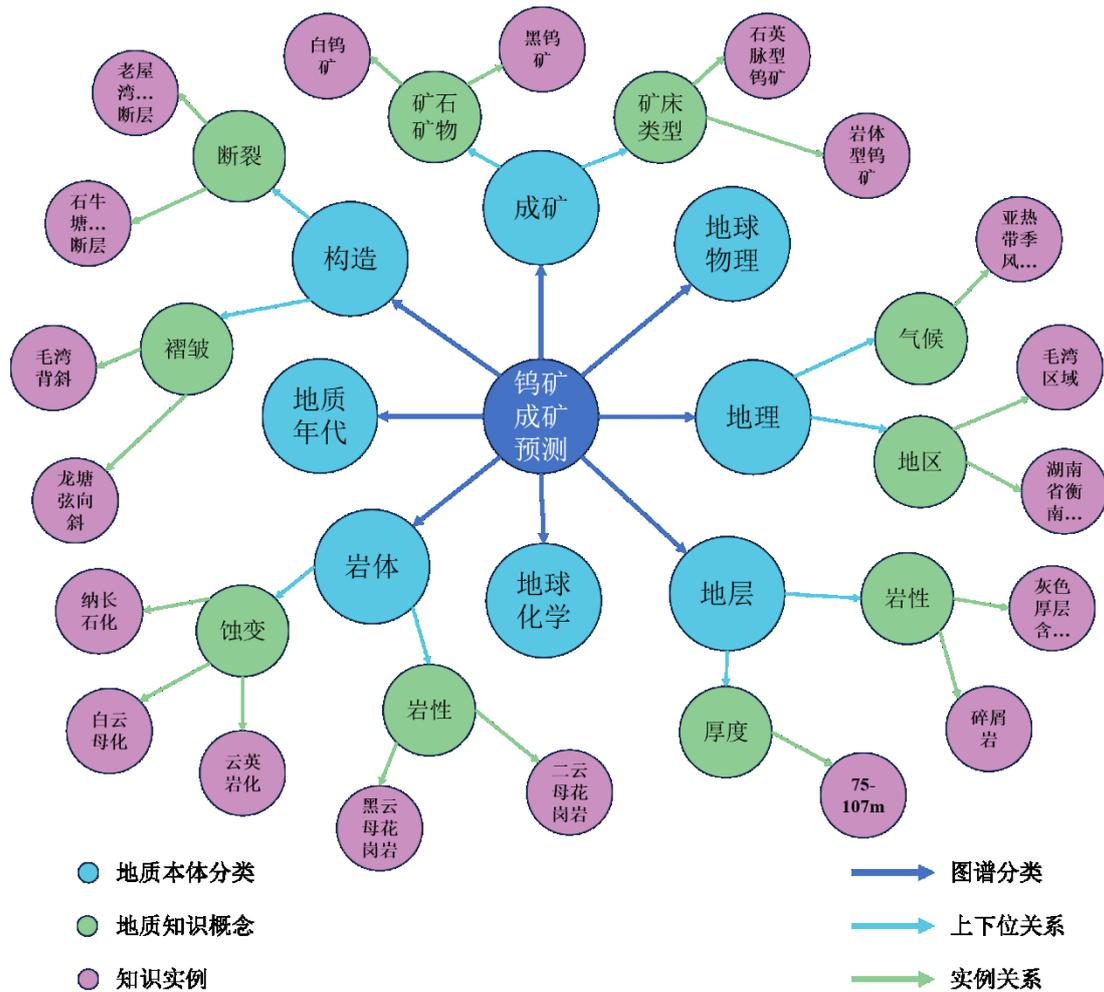


图 2 地质找矿知识图谱本体模式层

Fig. 2 Ontology Schema of geological prospecting knowledge graph.

在地质找矿知识的信息抽取时，需要利用 CoT 流程辅助大语言模型进行推理工程。思维链的意义是让大语言模型逐步参与，将一个复杂问题进行分解为一步一步的子问题并依次求解。CoT 允许垂直领域的大语言模型应用，将复杂推理任务进行分解，从而保证模型的有效性；同时，CoT 也在任务构建与调优过程中显式地表达任务中存在的问题与正确思考，如表 1 所示。提示词设计（图 3）分为模糊关键词匹配推理、三元组知识抽取和知识校验与规则迭代三部分。

表 1 基于思维链的三元组抽取伪代码

Table 1 Pseudocode of triple extraction through chain-of-thought.

算法 1. TripleExtractionByChainOfThought	
输入:	
Q:	用户问题
K:	领域知识文本
E:	few-shot 示例集合

Ontology: 本体定义 (实体、关系约束)

输出:

Triplets(Q, K): 抽取的三元组集合 [Subject, Rel, Object]

函数定义:

parse(Q): 对问题进行语义解析与关键词提取

augment(Q, E): 将 few-shot 示例引入问题, 提示大模型学习推理模式

align(entity, Ontology): 实体对齐, 保证与本体定义一致

extract(K, Q): 根据问题和知识文本抽取候选实体与关系

validate(triple, Ontology): 验证三元组是否符合语义和本体约束

normalize(triple): 对三元组格式进行规范化表达

程序:

1. keywords \leftarrow parse(Q)
// 问题解析, 提取出关键实体和关系词
 2. prompt \leftarrow augment(keywords, E)
// 将 few-shot 示例注入提示, 形成 Chain-of-Thought 引导
 3. candidates \leftarrow extract(K, prompt)
// 从知识文本中抽取候选实体和关系, 得到初步三元组集合
 4. for each triple \in candidates do
 - a. aligned \leftarrow align(triple.entities, Ontology)
// 将实体与本体对齐, 解决术语不一致
 - b. valid \leftarrow validate(triple, Ontology)
// 检查关系类型、实体层级是否合理
 - c. If valid then
Triplets \leftarrow Triplets \cup normalize(triple)
// 保留并规范化该三元组
 5. Return Triplets
//返回三元组
-

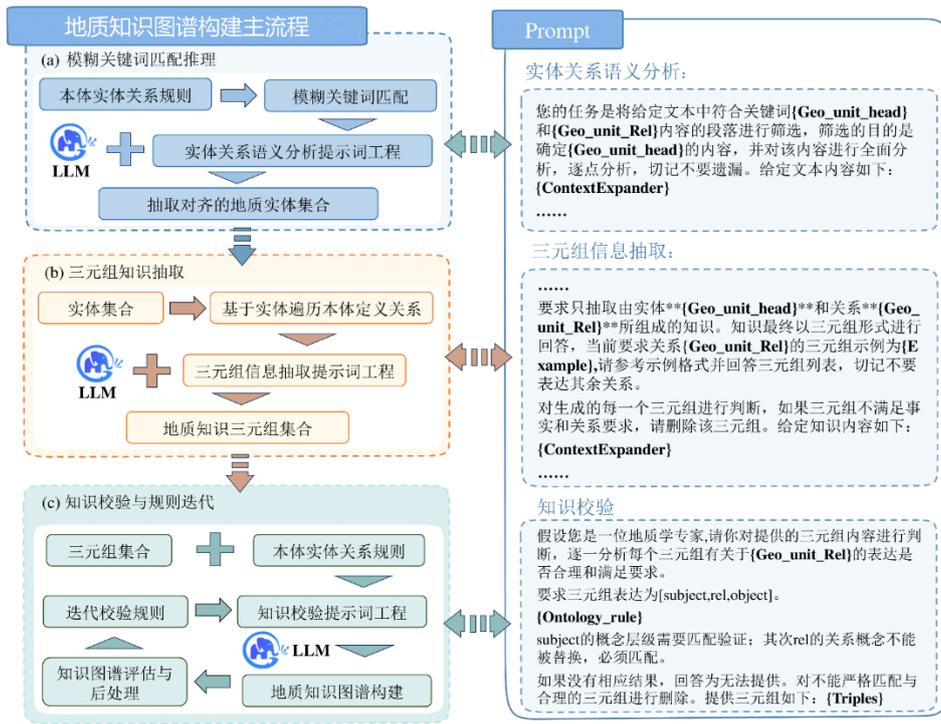


图 3 基于地质找矿本体模式层的三元组抽取提示词工程框架

Fig. 3 Prompt engineering framework for triple (head entity, relationship, tail entity) extraction constrained by the geological prospecting ontology schema.

(1) 模糊关键词匹配

模糊关键词匹配主要是利用本体定义的实体类型和关系规则进行信息抽取、获取相关语料，在语料的基础上进行信息分析与评价，从而为后续三元组知识抽取提供推理信息，其逻辑映射可以表达：

$$info = f(prompt_{info}, onto, txt) \#(1)$$

式中， $info$ 为大语言模型推理的知识分析文本， f 表示大语言模型映射函数， $prompt_{info}$ 表示模糊关键词推理的提示词模板， $onto$ 表示本体定义的实体和关系， txt 表示基于关键词匹配的上下文语料。

(2) 三元组知识抽取

三元组知识抽取是按照本体设计的上下框架进行从顶层到底层的实体节点抽取。在抽取对应层级的实体节点时，利用本体定义关系对其进行锚定，从而避免得到错误的结果。同时在提示词工程中，利用 Few-shot 的示例回答机制，将对应实体节点的实例 $Geol_{instance}$ 进行补充，从而让大语言模型对抽取的三元组与

正确三元组进行形式与内容上的比对，最终有效从上下文分析的语料中抽取三元组。Few-shot 不需覆盖全部预定义关系，而是按照专家经验设置少量回答模板来引导模型进行回答。

(3) 知识校验与规则迭代

本研究采用的是 9B 等级的本地大语言模型，尽管大语言模型大部分情况下能对提示词工程任务进行有效和合理的反馈，但是也会存在错误分类的情况，因此有必要在推理思维链中补充校正与规范化输出的过程。本研究将抽取得到的三元组与本体设计的规则以及实例进行匹配验证，让大语言模型判别三元组的合理性，从而建立补充规则以优化原始关系定义的描述，最终改善大语言模型推理的效果，得到结构化的三元组结果输出。

2.2 知识图谱增强大语言模型

检索增强生成是指结合文本嵌入模型和生成模型，利用外部知识库补充大语言模型的隐性知识。针对传统文档向量化的检索增强生成面临的地质文本数据集 Chunk 划分困难和长文本信息噪声导致的弱相关性问题，本研究设计了一个基于知识图谱的检索增强生成框架，旨在结合知识图谱和嵌入模型，通过多跳关系检索相关知识，以取代文档级检索增强生成流程，优化问答效果。实验使用的 Embedding 模型是 MiniCPM-Embedding，模型参数大小为 2.4B，结合知识图谱的检索增强生成流程如图 4 所示。以问题“川口区域的地层出露状况是什么”为例，首先，经过问题重构后将保留主要关键词“川口区域、地层、出露状况”等集合，通过对上述重组的关键字利用嵌入模型形成特征向量，同时从图数据库中将三元组知识进行图嵌入以形成特征向量；其次，在第一次相似性排序可以获取诸如<川口区域，出露的地层，泥盆系跳马涧组>等高度语义相似性的知识信息；最后，利用图节点进行多跳扩展，便可以从“川口区域”和“泥盆系跳马涧组”进行相关节点检索，从而得到诸如<泥盆系跳马涧组，岩性，巨厚层状石英砾岩>和<泥盆系跳马涧组，上段厚度，约 35m>的详细地层背景信息，从而摆脱原始 chunk 划分对于引申信息的弱关联性，并通过辅助回答模板对问答推荐进行生成。优化问题的输入，主要是提取问题的关键词从而弱化原始问题文本对回答的影响。问题重构不会强制依赖知识图谱，而会根据关键词属性关联图结构，即通过关键

词与节点属性的相似性进行关联。

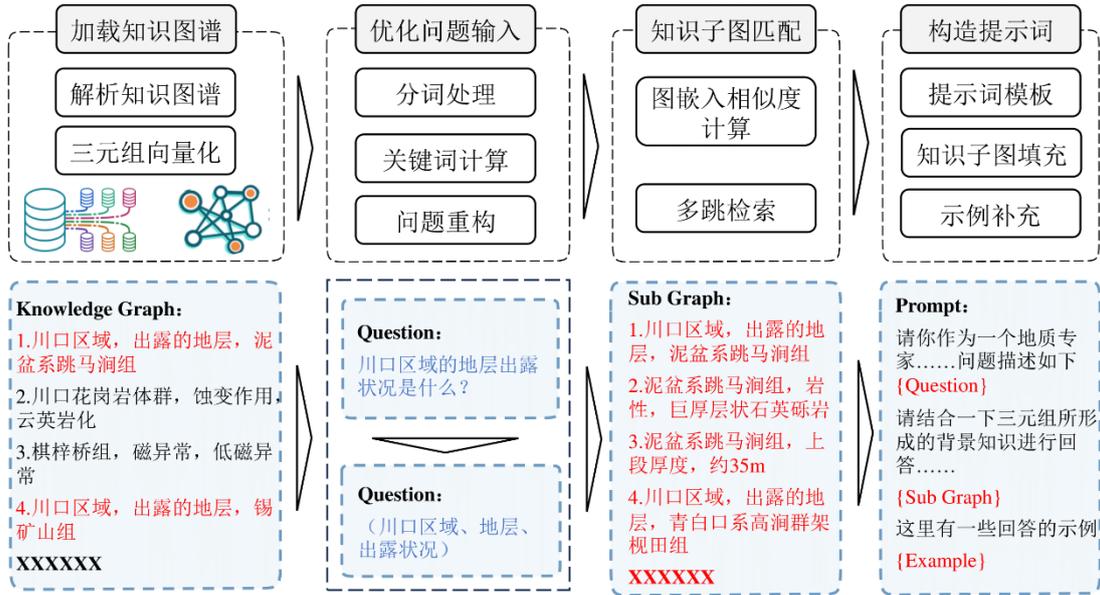


图 4 嵌入知识图谱的检索增强生成

Fig. 4 Knowledge graph embedded retrieval-augmented generation.

通过将问答模型的自然语言交互进行任务逻辑划分,从而让大语言模型能够构建一条逻辑推理,提示词工程设计如图 5 所示,分为任务描述/角色定义、找矿知识图谱利用和回答示例补充。通过该提示词确保 LLM 生成的回答既准确又符合专业领域的表达方式,同时也提高了可控性。

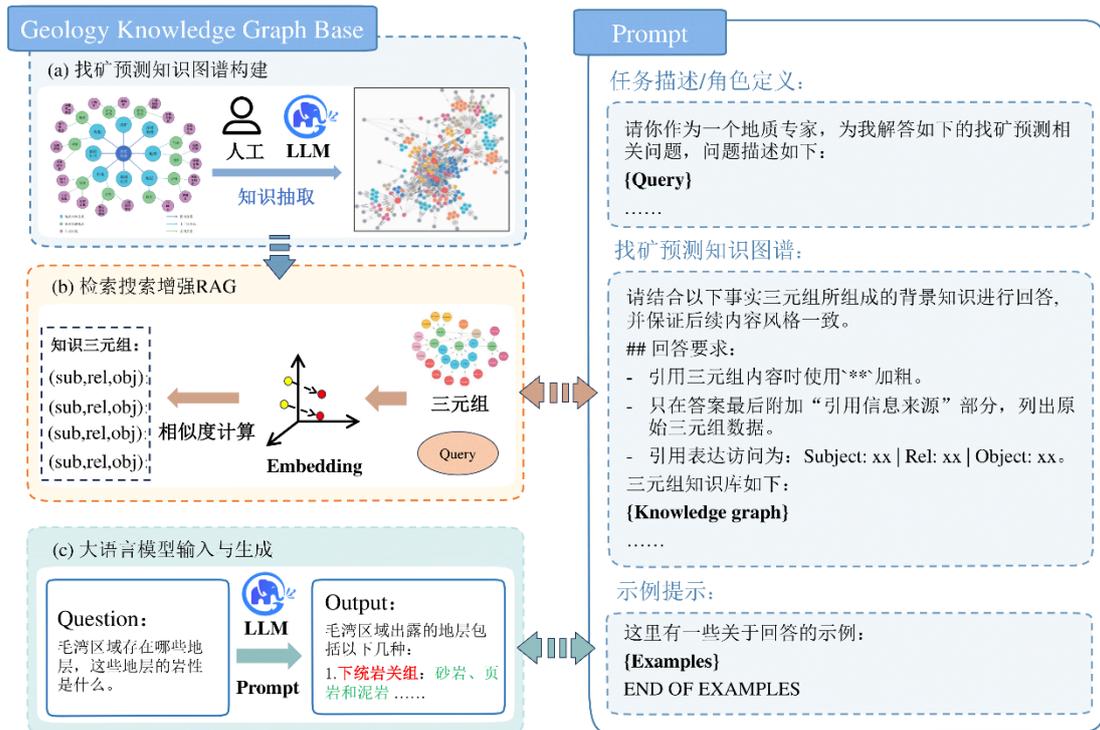


图 5 地质找矿问答模型提示词工程

Fig. 5 Prompt engineering for geological prospecting question-answering model.

(1) 任务描述/角色定义

通过提示词让 LLM 扮演地质专家角色，确保回答具有专业性和针对性。同时对原始问题进行输入，让 LLM 基于这个输入进行回答。同时强调回答的目标是“解答地质找矿相关问题”，而不是提供通用信息，避免回答泛化。

(2) 地质找矿知识图谱的利用

首要是明确 LLM 的回答，要求 LLM 的回答基于三元组提供的信息，而不是凭空生成。同时要求前后文风格一致性，以便 LLM 的回答能够按照后续示例进行优化，保证回答的规范性。还要通过列出所有使用的三元组来源，确保信息的可溯源性，并且采用标准格式进行输出，方便对回答三元组进行解析。

(3) 示例提示

通过提供参考领域问题和答案，帮助 LLM 理解理想的回答风格、格式和内容层次。在 Few-shot 的基础上，LLM 可以基于示例的结构进行模仿，确保输出风格与预期一致。同时面对不同类型的问题，让 LLM 学会如何灵活处理各种输入。

3 大语言模型赋能的知识图谱构建

3.1 语料库制作

本研究所采用的数据集源自湖南省衡阳市川口钨矿田的地质勘探资料，涵盖川口、白水、毛湾、塘江沅共四个主要钨矿床(蔡富成等, 2021, 宋宏邦等, 2002, Li et al., 2023)。目前在川口钨矿田已经完成多次地质勘探工作，并分析了其矿床类型与地球化学特征，对其成矿规律进行了总结(彭能立等, 2017, Qin et al., 2020, Li et al., 2021)。收集的文档资料主要包括钨矿普查总结报告、钨矿普查设计书、地球物理勘探报告、质量总结报告等 17 类地质技术文档，如表 2 所示。这些资料以 PDF 格式为主，辅以部分 Word 与 Excel 文档，其内容涵盖地质背景调查、勘查设计、物探成果、找矿结论等关键信息，全面反映了区域内的成矿特征与勘探工作进展。

表 2 研究区文档资料

Table 2 Geological prospecting documents of the study area.

序号	资料类型	资料字符数
1	矿区钨矿普查设计书	25096
2	矿区钨矿普查总结报告	46826
3	矿区钨矿普查附表	33239
4	矿区地球物理勘探报告	16323
5	矿区野外工作总结	20434
6	矿区工程测量报告	1507
7	矿田外围钨成矿规律及靶区预测研究	30260
8	矿区质量总结报告	6605
9	矿区实测地质剖面原始地质记录表	72
10	矿区实测剖面数据计算表	87
11	矿区实测剖面小结	2683
12	矿区外检样	63
13	矿区地质填图小结	4320
14	矿区内检查样	71
15	矿区钻孔采样登记表	338
16	矿区化学样分析结果	275
17	矿区情况说明	81

在对地质文档进行知识抽取前，需要对文档资料数据集进行非结构化数据转换，进行图文表分割以进一步处理。数据预处理流程分为以下 4 步：首先，数据输入主要负责数据的收集与归类；其次，转换与筛选主要是为文本数据集的字符串收集进行统一处理；再次，句子分割则是将字符串数据集进行合理的拆分，以保证输入大语言模型的语义连贯性；最后，校正优化则是补充自动化算法的遗漏或者校正错误。

采用 pdfplumber 进行图文表的划分，同时在数据拆分时能够有效去除噪音字符，如文档中的页码和页眉等信息。然后，通过集成 spaCy 的中文大语言模型包 zh_core_web_lg 对预处理数据进行句子分割以保证语义完整性，便于大语言模型能够有效地理解内容。对于文本 T ，每个词 t_i 被赋予一个标签 s_i ，表示该词是否是句子的边界：

$$s_i = \begin{cases} 1, & t_i \text{ 是句子结束点} \\ 0, & t_i \text{ 不是句子结束点} \end{cases} \quad \#(2)$$

式中主要采用 3 类规则进行句子边界约束，即标点符号规则 $R(t_i)$ 、依存关系规则 $D(t_i)$ 和机器学习概率模型规则 $P(t_i)$ 。

$$R(t_i) = \begin{cases} 1, & \text{如果 } t_i \in \{., ?, !\} \text{ 且满足上下文条件} \\ 0, & \text{否则} \end{cases} \quad \#(3)$$

$$D(t_i) = \begin{cases} 1, & \text{如果 } dep(t_i) = punct \text{ 且 } head(t_i) = Root \\ 0, & \text{否则} \end{cases} \quad \#(4)$$

式中， $Root$ 表示句子的主谓结构结束， $punct$ 表示句尾标点符号依附于主句。 $dep(t_i)$ 表示对 t_i 词互补性依存关系判断， $head(t_i)$ 表示对 t_i 词进行句法依存中的头节点进行判断。

$$P(t_i) = \sigma(W \cdot \phi(t_i) + b) \quad \#(5)$$

式中， W 表示权重向量， b 表示为偏置量， σ 表示激活函数， $\phi(t_i)$ 表示对词 t_i 的词性标注。

3.2 地质找矿知识图谱

经过上述提示词引导的大语言模型三元组抽取工作之后，最终累计获得了 973 条领域相关的三元组数据，其部分实例如表 3 所示。

表 3 地质找矿知识图谱三元组示例（部分）

Table 3 Some triple instances in the geological prospecting knowledge graph.

实体 A	属性	实体 B/关系
泥盆系跳马涧组	岩性	巨厚层状石英砾岩
泥盆系跳马涧组	岩性	砂砾岩
泥盆系跳马涧组	岩性	砂岩
泥盆系跳马涧组	岩性	含泥、钙质砂岩
泥盆系跳马涧组	上段厚度	约 35m
泥盆系跳马涧组	中段厚度	约 37m
泥盆系跳马涧组	下段厚度	约 46m
研究区域	气候类型	亚热带季风气候
花岗岩体群	岩性	中细粒二云母花岗岩
花岗岩体群	岩性	细粒白云母花岗岩
花岗岩体群	岩性	斑状黑云母花岗岩
花岗岩体群	形态	岩株
某背斜	核部地层	棋梓桥组
某断层	长度	17.35km
某矿区	矿产类型	岩体型钨矿

本研究从知识覆盖度和关联完备性两个维度评估了知识图谱质量，具体指标定义如下：

(1) 知识覆盖度(丰度指标)，知识覆盖度用于衡量图谱信息量的充分程度，其核心特征表现为：实体属性的完备性、关系连接的密集程度和概念层次的完整性。基于图论中的节点中心性理论，本研究用度中心性作为量化指标。具体而言，通过计算实体节点的平均连接度（包含入度和出度）来评估知识覆盖水平，平均度(aveD)的计算方式为：

$$aveD(G) = \frac{\sum_{i=0}^{|V|} D(v_i)}{|V|}, D(v_i) = \sum_{i=0}^{|V|} a_{ij} \#(6)$$

式中， $D(v_i)$ 代表每个顶点 V 的度。

(2) 链接预测 (Link Prediction)，链接预测通常是评估知识图谱推理能力正确性，直接衡量了知识图谱中隐藏信息的推理能力。知识图谱的推理能力指的是基于现有的三元组，推导出新的合理三元组的能力。这对于评估非监督知识图谱构建的有效性非常重要，因为推理能力直接与图谱的表现能力相关。通过这种方式，可以评估图谱的推理是否能够从已有的三元组推导出新的知识。本研究使用了准确性和召回率指标，具体定义如下：

(a) 准确性，准确性反映了知识图谱中存储的三元组是否正确地表达了实体之间的关系。在实际应用中，由于知识图谱构建通常依赖于自动化方法（如关系抽取模型、实体识别等），因此会存在一定的错误信息。通过比较知识图谱中的推理结果与已知的真实数据进行匹配，能够量化图谱的准确性。具体通过 MRR(Mean Reciprocal Rank)和 Hits@K(K=1,3,10)指标来量化预测的准确度。

$$MRR = \frac{1}{|T|} \sum_{i=1}^{|T|} \frac{1}{r_i} \#(7)$$

式中， r_i 是第 i 个预测的排名， $|T|$ 是测试集中三元组的数量。MRR 衡量的是所有预测中正确三元组的平均倒数排名，值越大表示模型预测准确度越高。

Hits@K 则是衡量在前 K 个预测中，正确答案是否出现在其中，即

$$Hits@K = \frac{\text{正确的预测数量}}{\text{总预测数量}} \#(8)$$

(b) 召回率，召回率衡量的是模型从知识图谱中检索到的正确实体和关系的比例。尽管模型在某些任务中具有较高的准确性，但往往可能会遗漏某些真实的实体或关系，因此召回率能够有效衡量模型的检索能力。为了评估召回率，本研究采用 Hits@3 和 Hits@10 等指标，测试模型在多个预测中能否识别出正确的尾实体。较高的 Hits@3 和 Hits@10 值表明，模型在大范围的预测中能够涵盖

更多的真实实体。

本研究构建的地质找矿知识图谱评估如表 4 所示。当前知识图谱节点总数，小于 1000 个节点属于轻量级，因此丰度的合理区间为 2-5。本研究构建的平均丰度数表现为 3.05，这个结果表明图谱相对不是特别稀疏，但仍存在许多低度数节点。其中经统计发现度数为 1 的节点占比最高，说明知识图谱中有许多“孤立”实体，这部分的原因是因为部分实体节点作为属性表现出了“叶子节点现象”，比如“岩性，走向，厚度”等关系对应的属性实体。同时少数节点在度数呈现几十的数量级，这说明这些节点连接了大量其他节点，主要是矿区和地质单元节点，可能在矿区这个概念细分得不够彻底，导致该节点存在过度泛化的现象。本研究采用 TransE 作为预测模型，在 MRR 实验中计算为 0.631，符合 0.5-0.8 的合理区间之内，这一结果表明，模型在预测尾实体时，能够相对较好地找到正确的实体。

表 4 评估指标统计表

Table 4 Evaluation metrics of the constructed knowledge graph

<i>aveD</i>	MRR	Hits@1	Hits@3	Hits@10
3.05	0.631	0.476	0.741	0.933

3.3 知识图谱构建方法对比

为验证本文所知识图谱构建方法的有效性，将本文方法与基于纯语料的知识图谱构建方法（ChatGLM-9B）进行了比较。人工标记了 100 份语料（表 5），选取了精确率、召回率和可信度进行方法比较。

表 5 用于知识图谱对比的人工标注语料库

Table 5 Manually annotated corpus for KG comparison.

序号	语料类型	数量	内容描述
1	地理	10	描述区域、矿区等的地理位置、坐标范围等信息
2	地层	25	描述各地层单位的岩石特征、厚度、组成等
3	岩体	5	描述岩体类型、成分、结构等信息
4	构造	20	描述褶皱、断层、背斜等地质构造信息

序号	语料类型	数量	内容描述
5	地质年代	8	描述地质时代、期次等时间尺度信息
6	成矿要素	20	描述成矿系统、控矿因素、矿化特征等信息
7	地球物理	6	描述物探异常、物理场特征等信息
8	地球化学	6	描述化探异常、元素分布等信息

(1) 精确率(Precision), 用于评估在模型预测的正例中, 实际为正确的比例。在面对地质垂直领域的知识图谱构建时, 高精确率意味着可以有效避免生成大量无效或者错误的三元组, 从而衡量有多少实际相关的地质知识被成功抽取出来。

$$Precision = \frac{TP}{TP + FP} \#(9)$$

式中, TP(True Positives)表示模型正确预测的三元组数量, 即预测为地质知识相关三元组且实际也为地质知识相关的三元组; FP(False Positives)表示模型错误预测的三元组数量, 即模型预测为地质知识相关但实际与领域主题无关的三元组。

(2) 召回率(Recall), 用于衡量模型是否能捕捉到所有的正例三元组, 即所有实际相关的三元组中有多少被模型正确识别出来。高召回率意味着模型能够从大量的语料中提取出更多的潜在知识。对于地质找矿知识图谱构建任务, 较高的召回率有助于确保图谱中包含尽可能多的知识信息。

$$Recall = \frac{TP}{TP + FN} \#(10)$$

式中, FN(False Negatives)为模型漏掉的三元组数量, 即预测为不相关的三元组但实际是相关的三元组数量。

(3) 可信度(F1-score)是精确率和召回率的调和平均数, 用于综合考虑精确率和召回率的指标。对于知识图谱构建任务, F1-score 能够平衡精确率和召回率, 避免只关注一方面导致的偏差。

$$F1 - score = 2 \times \frac{Precision * Recall}{Precision + Recall} \#(11)$$

根据上述评估方法, 地质找矿知识图谱构建的实验评估结果如表 6 所示。

表 6 地质找矿知识图谱构建评估表

Table 6 Evaluation of knowledge graphs for geological prospecting.

	Precision	Recall	F1-score
KG-RAG	0.807	0.833	0.819
ChatGLM-9B	0.537	0.768	0.632

4 地质找矿智能问答模型构建

4.1 地质找矿问答模型

地质找矿问答模型允许人工输入自然语言问句与大语言模型进行交互,大模型根据自然语句执行知识图谱检索增强生成(KG-RAG),从而得到领域内的专业结果。例如,对于问题“毛湾区域存在哪些地层,这些地层的岩性是什么?”的回答(图6),KG-RAG能够正确总结区域地层及岩性信息,而不使用知识图谱的大语言直接回答,缺乏地质背景信息产生幻觉,进行了信息捏造。

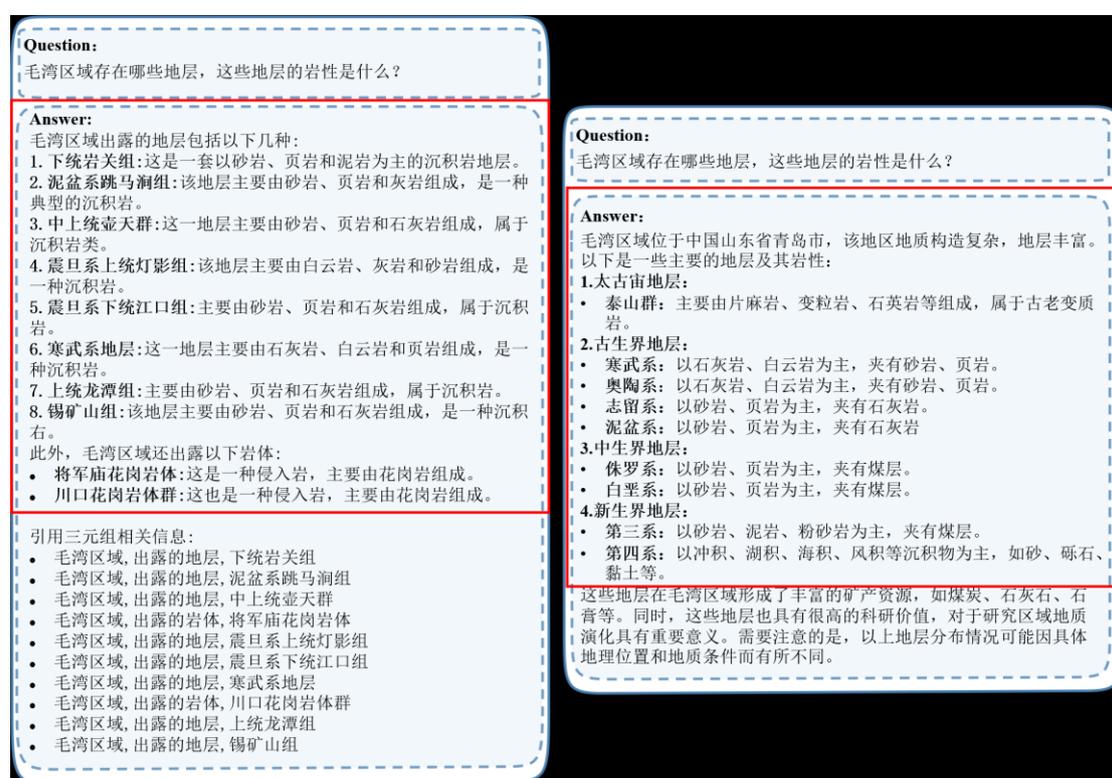
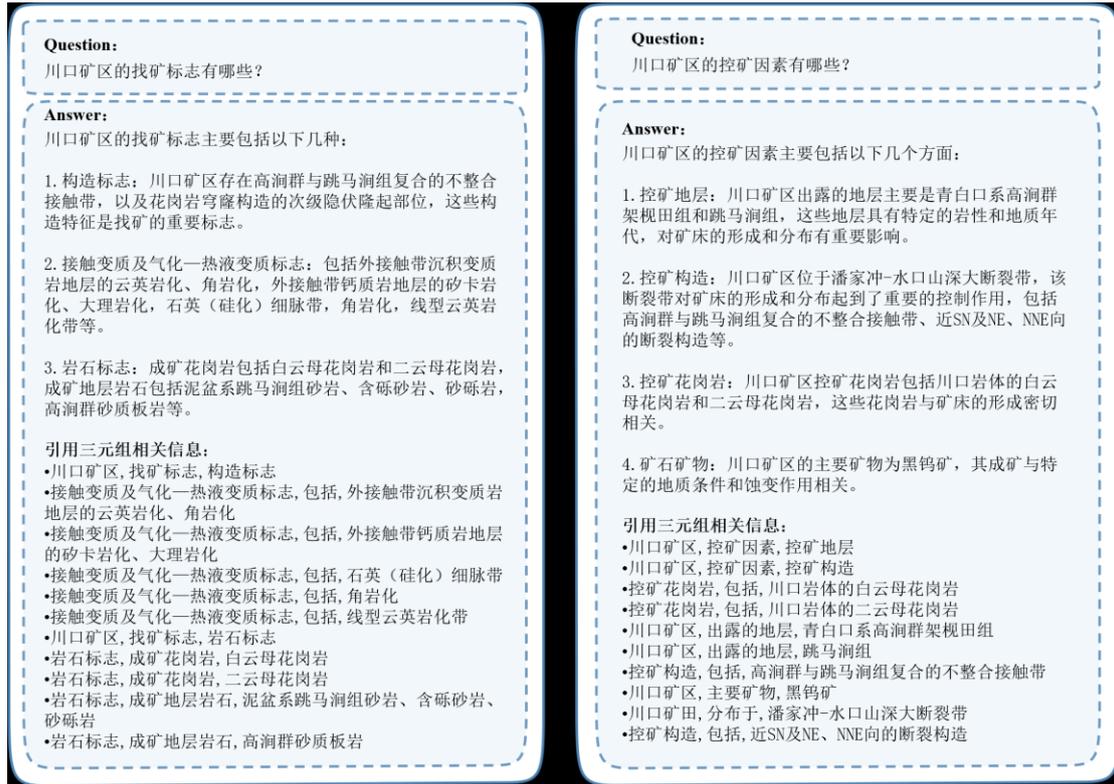


图 6 知识图谱检索增强生成的知识回答示例:(a)使用 KG-RAG 的大语言模型,(b)大语言模型直接回答

Fig. 6 Instances of question-answering by (a) KG-RAG and (b) LLM directly.

此外,本研究分别从矿区的找矿标志和控矿因素进行提问,问答模型均能从众多提供的背景信息三元组中选择相关的成矿信息进行总结回答,如图7所示。



(a) (b)
图 7 研究区地质找矿知识回答结果: (a) 找矿标志, (b) 控矿要素

Fig. 7 Question-answering instances of geological prospecting knowledge in the study area:
(a) prospecting clues and (b) ore-forming conditioning factors.

4.2 检索增强生成方法对比

为验证本研究所提出 KG-RAG 模型的有效性, 分别与微软的 GraphRAG (<https://github.com/microsoft/graphrag>)、基于 LangChain 构建的文档级检索增强生成 (Doc RAG) 模块以及不采用 RAG 方案(No RAG)的单纯大语言模型进行比较, 实验共设置 100 组问答对。选取了检索增强生成任务中常用的召回率和可信度来评估模型的检索性能与生成质量, 并增加了准确率作为整体输出效果评估指标。

(1) 召回率 (Recall): 通过将检索内容与正确回答所涉及的事实内容进行比较从而确定是否检索到有效信息, 在文档级 Doc RAG 中, 为统计检索到的 Chunk ($Chunk_{retrieval}$) 中包含标准答案 Chunk ($Chunk_{true}$) 的比例; 在知识图谱级 GraphRAG 和 KG-RAG 中, 则统计检索三元组 ($Triple_{retrieval}$) 中涵盖正确答案 ($Triple_{true}$) 的比例。召回率计算公式为:

$$\text{Recall} = \begin{cases} \frac{\text{Chunk}_{\text{retrieval}}}{\text{Chunk}_{\text{true}}}, \text{ for Doc} \\ \frac{\text{Triple}_{\text{retrieval}}}{\text{Triple}_{\text{true}}}, \text{ for KG} \end{cases} \quad \#(12)$$

(2)可信度(Faithfulness):用以评估大语言模型的生成是否基于检索结果,不添加无关信息或歪曲事实。可信度能直接衡量生成内容是否严格遵循检索到的证据,从而减轻幻觉或篡改事实。本研究将检索生成事实(Chunk_{generation}或Triple_{generation}),与检索正确事实(Chunk_{retrieval}或Triple_{retrieval})进行对照,如果检索生成结果无法与事实匹配,则视为幻觉生成。可信度指标计算公式为:

$$\text{Faithfulness} = \begin{cases} \frac{\text{Chunk}_{\text{generation}}}{\text{Chunk}_{\text{retrieval}}}, \text{ for Doc} \\ \frac{\text{Triple}_{\text{generation}}}{\text{Triple}_{\text{retrieval}}}, \text{ for KG} \end{cases} \quad \#(13)$$

(3)准确率(Accuracy):用以衡量采用RAG方案前后的模型输出正确与否,本研究采用预测回答和标准回答的语义相似性值进行判定,从而确定预测回答(Answering)是否为正确回答(Answering_{true}),准确性计算公式为:

$$\text{Accuracy} = \frac{\text{Answering}_{\text{true}}}{\text{Answering}} \quad \#(14)$$

根据上述评估方法,问答关联检索的实验结果评价如表7所示。实验结果表明,KG-RAG在各项指标上都取得了最好的表现,尤其是召回率和准确率方面,验证了结构化的知识能够显著提升检索和生成的效果。

表7 问答关联检索实验结果评价

Table 7 Evaluation metrics of question-answering models.

	Recall	Faithfulness	Accuracy
KG-RAG	0.917	0.808	0.88
GraphRAG	0.785	0.808	0.73
Doc RAG	0.735	0.813	0.72
No RAG			0.38

5 结论

本研究围绕地质找矿领域的知识图谱构建、大语言模型问答优化展开,提出了一种融合地质找矿知识图谱的增强检索框架(KG-RAG),建立了从数据采集处理、知识抽取、知识组织、问答优化到实际应用的完整流程,不仅为地质找矿提供了新的智能问答模型,也为大语言模型与知识图谱的联合应用提供了借鉴。本研究取得的主要结论如下:

(1) 采用大语言模型辅助进行知识图谱的构建，避免了监督方法依赖人工标注数据的问题，同时面对大语言模型的数据信息噪声干扰、实体关系抽取准确率下降与知识融合不稳定问题，以地质找矿本体模式层作为指导，配合模糊关键词匹配算法，有效改善了地质文本语料中的信息冗余和数据噪声问题，实现了地质知识的自动化抽取和结构化组织；

(2) 利用本体模式层引导大语言模型提示词工程的思维链，通过实体关系分析、三元组抽取和知识校验三步走的思维链，改善了相较于通用大语言模型在地质找矿领域知识抽取的准确率下降问题和知识融合问题，从地质文献、报告、数据库等多种数据源中有效提取地质实体及其关系；

(3) 提出了知识图谱与检索增强生成结合的 KG-RAG 框架，搭建了地质找矿智能问答模型，利用知识图谱和思维链引导替代了传统文档向量 RAG 手段的 Chunk 划分，改善了在地质知识中长上下文 Chunk 划分困难与 Chunk 段落检索相关性低的问题，同时利用多跳检索算法，将问答模型进行深度与广度拓展，确保了回答检索信息的全面性与可靠性。

本研究提出的地质找矿智能问答模型的 KG-RAG 框架，仍存在一些需要进一步优化和拓展的方向：首先，大语言模型在地质找矿知识推理和生成方面仍存在“幻觉”问题，可以进一步引入基于知识图谱的事实验证机制，以及结合多轮对话增强，提高模型问答的可信度和逻辑一致性；其次，目前的知识图谱构建主要以文本数据为主，然而地质找矿领域的地、物、化、遥的数据形式丰富，如何将地质大数据的多模态信息有效地映射到知识图谱中，提高知识表示的多样性和完整性，也是未来的重要研究方向。

本研究的 Python 源代码（KG-RAG）可在 Github 上开放获取（<https://github.com/Geo3D-AI-CSU/KG-RAG>）。

致谢

中南大学邹滨教授和李沈鑫博士在研究中给予了热情协助，国家地理信息系统工程技术研究中心与中南大学共建“MAPGIS 实验室”为本研究工作提供了 MAPGIS 软件，在此一并表示感谢！

References

- Bizer C, Lehmann J, Kobilarov G, et al. 2009. DBpedia - A crystallization point for the Web of Data [J]. *Journal of Web Semantics*, 7: 154-165.
- Cai F C, Qin J H, Qin J N, et al. 2021. Geochemical characteristics and LA-ICP-MS zircon U-Pb dating of ore-bearing granite of Chuankou intrusion-related tungsten deposit, Hunan Province [J]. *China Geology*, 48: 1212-1224 (in Chinese with English abstract).
- Chen X, Liu Y, Han W, et al. 2025. A vision-language foundation model-based multi-modal retrieval-augmented generation framework for remote sensing lithological recognition [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 225: 328-340.
- Church K W, Sun J M, Yue R C, et al. 2024. Emerging trends: a gentle introduction to RAG [J]. *Natural Language Engineering*, 30: 870-881.
- De Almeida T D, De Oliveira N N, He C D, et al. 2025. Using Generative Pre-Trained Transformer-4 (GPT-4), ffmpeg, and Microsoft Azure to Aid in Creating a Text-to-Video Generation Tool to Improve Safety Shares and Incident Descriptions in the Mining Industry [J]. *Mining Metallurgy & Exploration*, 42: 1325-1343.
- Dong S C, Li Y, Lyu H R, et al. 2020. An Editing Platform of Geoscience Knowledge System [J]. *Geological Journal of China Universities*, 26: 384-394 (in Chinese with English abstract).
- Dreyer J 2025. China made waves with Deepseek, but its real ambition is AI-driven industrial innovation [J]. *Nature*, 638: 609-611.
- Floridi L, Chiriatti M 2020. GPT-3: Its Nature, Scope, Limits, and Consequences [J]. *Minds and Machines*, 30: 681-694.
- Fu Y, Wang M, Wang C, et al. 2025. GeoMinLM: A Large Language Model in Geology and Mineral Survey in Yunnan Province [J]. *Ore Geology Reviews*, 182: 106638.
- Guo F, Lai P, Huang F M, et al. 2024. Literature Review and Research Progress of Landslide Susceptibility Mapping Based on Knowledge Graph [J]. *Earth Science*, 49: 1584-1606 (in Chinese with English abstract).
- Hosseini S, Seilani H 2025. The role of agentic AI in shaping a smart future: A systematic review [J]. *Array*, 26: 15.
- Hu Y, Mai G, Cundy C, et al. 2023. Geo-knowledge-guided GPT models improve the extraction of location descriptions from disaster-related social media messages [J]. *International Journal of Geographical Information Science*, 37: 2289-2318.
- Jiang B, Yang J X, Yang C, et al. 2020. Knowledge Augmented Dialogue Generation with Divergent Facts Selection [J]. *Knowledge-Based Systems*, 210: 11.
- Jiang S W, Zhang J W, Hua L S, et al. 2025. Implementation of Meteorological Database Question-Answering Based on Large-Scale Model Retrieval-Augmentation Generation [J]. *Computer Engineering and Applications*, 61: 113-121 (in Chinese with English abstract).

English abstract).

Katz D M, Bommarito M J, Gao S, et al. 2024. GPT-4 passes the bar exam [J]. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*, 382: 17.

Li C L, Wang Z X, Lü Q T, et al. 2021. Mesozoic tectonic evolution of the eastern South China Block: A review on the synthesis of the regional deformation and magmatism [J]. *Ore Geology Reviews*, 131: 33.

Li H, Yue P, Tapete D, et al. 2024. ESDC: An open Earth science data corpus to support geoscientific literature information extraction [J]. *Science China Earth Sciences*, 67: 3840-3854.

Li H, Yue P, Wu H R, et al. 2025. A question-answering framework for geospatial data retrieval enhanced by a knowledge graph and large language models [J]. *International Journal of Digital Earth*, 18: 24.

Li N X, Zhang R-Q, Zhu L, et al. 2023. Tracing tungsten-tin mineralization processes with tourmaline geochemistry in the Wangxianling-Hehuaping district, Nanling Range (South China) [J]. *Ore Geology Reviews*, 163: 105806.

Liang J, Hou S, Jiao H, et al. 2025. GeoGraphRAG: A graph-based retrieval-augmented generation approach for empowering large language models in automated geospatial modeling [J]. *International Journal of Applied Earth Observation and Geoinformation*, 142: 104712.

Ma X 2022. Knowledge graph construction and application in geosciences: A review [J]. *Computers & Geosciences*, 161: 105082.

Memduhoglu A, Fulman N, Zipf A 2024. Enriching building function classification using Large Language Model embeddings of OpenStreetMap Tags [J]. *Earth Science Informatics*, 17: 5403-5418.

Peng N L, Wang X H, Yang J, et al. 2017. Re-Os dating of molybdenite from Sanjiaotan tungsten deposit in Chuankou area, Hunan Province, and its geological implications [J]. *Mineral Deposits*, 36: 1402-1414 (in Chinese with English abstract).

Qin J-H, Wang D-H, Li C, et al. 2020. The molybdenite Re-Os isotope chronology, in situ scheelite and wolframite trace elements and Sr isotope characteristics of the Chuankou tungsten ore field, South China [J]. *Ore Geology Reviews*, 126: 103756.

Qiu Q J, Wu L, Ma K, et al. 2023. A Knowledge Graph Construction Method for Geohazard Chain for Disaster Emergency Response [J]. *Earth Science*, 48: 1875-1891 (in Chinese with English abstract).

Song H B, Huang M X, Fan Z H, et al. 2002. Characteristics of the ore-controlling structures of the Sanjiaotan wolframite deposit and its relationships with ore formation in Chuankou, Hunan [J]. *Geotectonica et Metallogenia*, 26: 51-54 (in Chinese with English abstract).

Tong B, Yin Y P, Li B, et al. 2025. Review on artificial intelligence-based large language models for geological hazards [J]. *The Chinese Journal of Geological Hazard and Control*, 36: 1-12 (in Chinese with English abstract).

Vidivelli S, Ramachandran M, Dharunbalaji A 2024. Efficiency-Driven Custom Chatbot Development: Unleashing LangChain, RAG, and Performance-Optimized LLM Fusion [J]. *Cmc-Computers Materials & Continua*, 80: 2423-2442.

Wang C B, Wang M G, Wang B, et al. 2024. Knowledge graph-infused quantitative mineral resource forecasting [J]. *Earth Science Frontiers*, 31: 26-36 (in Chinese with English abstract).

Wang D H, Liu X X, Liu L J. 2015. Characteristics of Big Geodata and its application to study of minerogenetic regularity and minerogenetic series [J]. *Mineral Deposits*, 34: 1143-1154 (in Chinese with English abstract).

Wu H Y, Shen Z X, Hou S Y, et al. 2025. Large language model-driven GIS analysis: methods, applications, and prospects [J]. *Journal of Geodesy and Geoinformation Science* 54: 621-635 (in Chinese with English abstract).

Wu R L, Guo D H. 2025. Research on Evaluation Standards for Spatial Cognitive Abilities in Large Language Models [J]. *Journal of Geo-Information Science*, 27: 1041-1052 (in Chinese with English abstract).

Xie J L, Wang G Q, Zhang T, et al. 2025. LLaMA-Unidetector: An LLaMA-Based Universal Framework for Open-Vocabulary Object Detection in Remote Sensing Imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 18.

Xu C, Su M Y, Sun B. 2024. Tourism Knowledge Graph Construction Based on ChatGLM and Prompt-tuning [J]. *Science Technology and Engineering*, 24: 13484-13492 (in Chinese with English abstract).

Zhang W, Cai M, Zhang T, et al. 2024a. EarthGPT: A Universal Multimodal Large Language Model for Multisensor Image Comprehension in Remote Sensing Domain [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5917820.

Zhang Y, Wei C, He Z, et al. 2024b. GeoGPT: An assistant for understanding and processing geospatial tasks [J]. *International Journal of Applied Earth Observation and Geoinformation*, 131: 103976.

Zhang Z-J, Kusky T, Gao M, et al. 2023. Spatio-temporal analysis of big data sets of detrital zircon U-Pb geochronology and Hf isotope data: Tests of tectonic models for the Precambrian evolution of the North China Craton [J]. *Earth-Science Reviews*, 239: 104372.

Zhou Y Z, Zuo R G, Liu G, et al. 2021. The Great-leap-forward Development of Mathematical Geoscience During 2010-2019: Big Data and Artificial Intelligence Algorithm Are Changing Mathematical Geoscience [J]. *Bulletin of Mineralogy, Petrology and Geochemistry*, 40: 556-573 (in Chinese with English abstract).

附中文参考文献

蔡富成, 秦锦华, 覃金宁, 等. 2021. 湖南川口岩体型钨矿赋矿花岗岩地球化学特征及 LA-ICP-MS 锆石 U-Pb 定年 [J]. *中国地质*, 48: 1212-1224.

董少春, 李艳, 闫海荣, et al. 2020. 地球科学知识体系编辑平台 [J]. *高校地质学报*, 26: 384-394.

郭飞, 赖鹏, 黄发明, 等. 2024. 基于知识图谱的滑坡易发性评价文献综述及研究进展 [J]. 地球科学, 49: 1584-1606.

江双五, 张嘉玮, 华连生, 等. 2025. 基于大模型检索增强生成的气象数据库问答模型实现 [J]. 计算机工程与应用, 61: 113-121.

彭能立, 王先辉, 杨俊, 等. 2017. 湖南川口三角潭钨矿床中辉钼矿 Re-Os 同位素定年及其地质意义 [J]. 矿床地质, 36: 1402-1414.

邱芹军, 吴亮, 马凯, 等. 2023. 面向灾害应急响应的地质灾害链知识图谱构建方法 [J]. 地球科学, 48: 1875-1891.

宋宏邦, 黄满湘, 樊钟衡, 等. 2002. 湖南川口三角潭黑钨矿床控矿构造特征及其与成矿的关系 [J]. 大地构造与成矿学, 26: 51-54.

佟彬, 殷跃平, 李昶, 等. 2025. 地质灾害人工智能大语言模型研究展望 [J]. 中国地质灾害与防治学报, 36: 1-12.

王成彬, 王明果, 王博, 等. 2024. 融合知识图谱的矿产资源定量预测 [J]. 地学前缘, 31: 26-36.

王登红, 刘新星, 刘丽君. 2015. 地质大数据的特点及其在成矿规律、成矿系列研究中的应用 [J]. 矿床地质, 34: 1143-1154.

吴华意, 沈张骁, 侯树洋, 等. 2025. 大语言模型驱动的 GIS 分析: 方法、应用与展望 [J]. 测绘学报, 54: 621-635.

吴若玲, 郭旦怀. 2025. 大语言模型空间认知能力测试标准研究 [J]. 地球信息科学学报, 27: 1041-1052.

徐春, 苏明钰, 孙彬. 2024. 基于 ChatGLM 和提示微调的旅游知识图谱构建 [J]. 科学技术与工程, 24: 13484-13492.

周永章, 左仁广, 刘刚, 等. 2021. 数学地球科学跨越发展的十年: 大数据、人工智能算法正在改变地质学 [J]. 矿物岩石地球化学通报, 40: 556-573.