中文引用格式:李华,吴立舟,李新宏,等. 基于 DeepSeek 与 RAG 的事故调查报告知识生成模型[J]. 中国安全科学学报,2026,36(1):26-34.

英文引用格式:LI Hua, WU Lizhou, LI Xinhong, et al. Construction and application of intelligent question-answering model for accident investigation reports based on DeepSeek and RAG [J]. China Safety Science Journal,2026,36(1):26-34.

基于 DeepSeek 与 RAG 的事故调查报告知识生成模型*

李 华[1]副教授, 吴立舟[1,2], 李新宏**[1]教授, 张 越[1], 冯 垚[1], 覃紫芸[1]
(1 西安建筑科技大学 资源工程学院,陕西 西安 710055;2 西安科技大学 安全科学与工程学院,陕西 西安 710054)

中图分类号:X928 **文献标志码**:A **DOI**: 10.16265/j.cnki.issn1003-3033.2026.01.0840
资助项目:2025年国家级大学创新训练计划项目(202510703088)。

【摘 要】 为解决大语言模型(LLM)在安全工程领域应用中面临的语料资源有限、输入容量受限和数据隐私等制约因素问题,构建结合 DeepSeek 与检索增强生成(RAG)机制的本地化事故问答模型,以实现复杂文本的智能解析和知识服务,从而辅助安全管理决策。基于政府应急管理系统发布的事故报告与法律法规,构建语义特征语料库,融合 PaddleOCR、LayoutLMv3、YOLOv8 等技术完成文档结构重建与语义建模。模型涵盖文档解析、语义对齐、知识库构建和混合检索 4 阶段,具备因果链条提取、法规匹配与语义映射能力。结果表明:相较未使用 RAG 机制的 Deepseek-r1:32b 模型,模型问答自动评分提高 7.7%,人工评分提高 17.6%,响应速度与稳定性指标较对照模型呈现出更高的数值表现。模型运行仍受到本地参数规模与知识更新机制的影响,试验结果表明其在任务中可实现预期功能。

【关键词】 DeepSeek; 大语言模型(LLM); 检索增强生成(RAG); 事故调查报告; 知识库

Construction and application of intelligent question-answering model for accident investigation reports based on DeepSeek and RAG

LI Hua[1], **WU Lizhou**[1,2], **LI Xinhong**[1], **ZHANG Yue**[1], **FENG Yao**[1], **QIN Ziyun**[1]
(1 School of Resources Engineering, Xi´an University of Architecture and Technology, Xi´an Shaanxi 710055, China; 2 College of Safety Science and Engineering, Xi´an University of Science and Technology, Xi´an Shaanxi 710054, China)

Abstract: In order to address the constraints of limited corpus resources, restricted input capacity, and data privacy in applying LLMs to the field of safety engineering, a localized accident question-answering model integrating the DeepSeek with a RAG mechanism was constructed to enable intelligent parsing and knowledge services for complex texts, thereby supporting safety management decision-making. A semantic-feature corpus was built based on accident investigation reports and laws and regulations released by government emergency management systems, and technologies such as PaddleOCR, LayoutLMv3, and

* 文章编号:1003-3033(2026)01-0026-09; 收稿日期:2025-09-14; 修稿日期:2025-11-21
** 通信作者:李新宏(1991—),男,甘肃镇原人,博士,教授,博士生导师,主要从事能源安全工程、安全信息化技术、风险评估与控制方面的研究。E-mail:lixinhong@xauat.edu.cn。

YOLOv8 were incorporated to accomplish document structure reconstruction and semantic modeling. The model encompassed four stages—document parsing, semantic alignment, knowledge-base construction, and hybrid retrieval—and was designed with capabilities for causal-chain extraction, regulation matching, and semantic mapping. The results indicated that, compared with the Deepseek-r1:32b model without the RAG mechanism, the enhanced model achieved improvements of 7.7% in automated scoring and 17.6% in human evaluation, and the response-speed and stability metrics presented higher numerical performance than those of the baseline model. The model performance was still influenced by the local parameter scale and the knowledge-updating mechanism, yet the experimental findings demonstrate that it is capable of fulfilling the intended functions in the present study.

Keywords: DeepSeek; large language model (LLM); retrieval augmented generation; accident investigation report; knowledge base

0 引　言

事故调查报告作为明确事故原因、划分责任及制定整改措施的重要依据,在安全管理与执法监管中具有关键作用。但此类报告普遍篇幅冗长、结构复杂,涵盖大量技术细节与法规条文,人工解读分析效率偏低,难以适配事故快速响应与知识智能提取的需求。人工智能技术的快速发展,特别是大语言模型(Large Language Model,LLM)在语义理解与文本生成领域的突破,为复杂文本自动解析与知识服务提供了全新路径。

刘瑞祥等[1]提出流程简易表示法,降低结构信息形式化的复杂性,提升大模型任务聚焦能力。检索增强生成(Retrieval Augmented Generation,RAG)技术引入外部文档,缓解数据滞后与语料不足问题,提升生成准确性与可靠性[2]。大模型在信息采集、理解与推理中表现优越,可为应急管理提供科学支持[3]。林鹏等[4]指出安全智能化正向数据融合、知识驱动、人机协同等方向发展。向量数据库结合大模型语义检索,优于传统关键词匹配,提升事故知识库检索效率与泛化能力[5]。大模型在命名实体识别等语义处理任务中展现出重要优势[6]。

为提升智能制造问答效果,WAN Yuwei 等[7]提出融合知识图谱与向量检索的混合 RAG 框架,缓解领域适应与知识时效问题。WU Chengke 等[8]设计 RAG 施工管理框架,融合文档解析、检索优化与偏好学习,改善施工管理中的交互体验。WONG Saika 等[9]构建结合领域知识的合同审查模型,提升风险识别能力。SUN Jianwen 等[10]提出自适应多级检索增强框架,无需调参实现教学问答的控制与一致性。REN Tengfei 等[11]将 RAG 用于航空事故因果识别,提升准确性与鲁棒性。SILVA 等[12]通过生成查询增强训练数据,提升低资源语义搜索能力。LEE 等[13]对比 RAG 与微调方法,验证其在建筑安全知识检索中的效果与边界。ZHANG Haiyang 等[14]提出结合法学知识的免调优合同审查模型,在实测中展现高准确率。JEON 等[15]构建 LLM 对话监控系统,联通自然语言与数控数据,准确率达 93.3%,显示其在智能制造中的应用潜力。尽管 LLM 与 RAG 技术在多个领域取得进展,但现有研究多聚焦于通用场景,缺乏对事故调查报告中复杂语义结构与专业知识的深入建模,检索机制融合度不高,模型验证也缺乏行业落地支持。

鉴于此,笔者拟聚焦于 DeepSeek 语言模型与 RAG 机制的融合应用,构建面向事故调查语境的本地化问答模型,围绕语义特征建模、知识库构建、混合检索设计与问答性能评估等关键环节展开模型设计与实证测试,以期为构建面向安全领域的专业智能问答模型提供技术支撑与实践依据。

1 语义建模与 LLM 相关技术

1.1 事故调查文本与法规语义特征分析

事故调查报告与法规文件承载着事故事实还原与规范约束的双重功能,是智能辅助决策的数据基础。鉴于前者强调时序与因果逻辑,后者侧重条款严谨性与责任界定,模型需重点解决因果链条提取、法律实体识别及跨文档语义映射难题,以实现“事故-法规”的深度联动。

基于 MinerU 构建包含文档解析、版面分析、语义理解与结构化重建的 4 阶段建模框架,如图 1 所示。该框架融合光学字符识别(Optical Character Recognition,OCR)、LayoutLMv3 及 YOLOv8 等技术,精准提取多模态特征并重构复杂版面,最终将文档转化为标准 Markdown 语义单元,显著提升检索与问

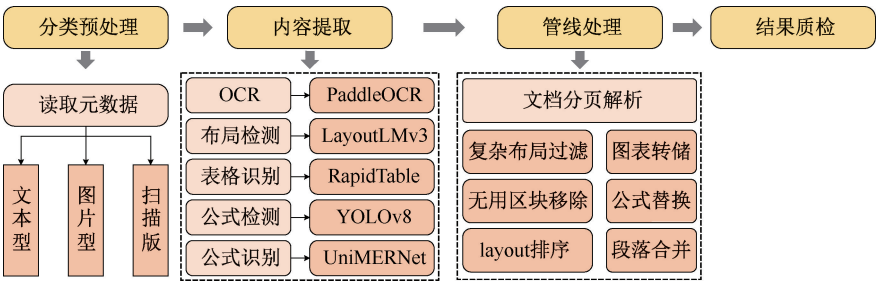


图 1　MinerU 语义建模处理流程

Fig. 1　MinerU processing flow

答模型的解析精度。

事故报告的因果演变逻辑与法规文件的责任界定规范互为补充。基于此双重特征,开展语义对齐与结构建模,构建精准适配的知识库框架,为事故智能问答与决策提供核心支撑。

1.2　LLM 与 RAG 机制

基于 Transformer 架构的 LLM 凭借卓越的上下文理解与泛化能力[16],为危机管理提供智能化新范式。结合罗伯特·希斯的 4R 模式(缩减 Reduction、预备 Readiness、反应 Response、恢复 Recovery)可深度赋能突发事件的全生命周期管理,在风险隐患排查、应急响应决策及灾后恢复总结等关键环节展现出广阔的应用前景[3]。

针对事故调查报告与法规文本兼具时序因果与规范约束的复杂特征,传统自然语言处理(Natural Language Processing,NLP)难以精准解析,而通用大模型面临知识幻觉与缺乏可追溯性的挑战。为此,引入 RAG 机制,通过句法切分、实体识别与向量化编码构建领域知识库。采用稀疏与稠密混合检索策略,协同 DeepSeek 模型形成“检索+生成”闭环(图 2),有效提升问答的准确性与可解释性。

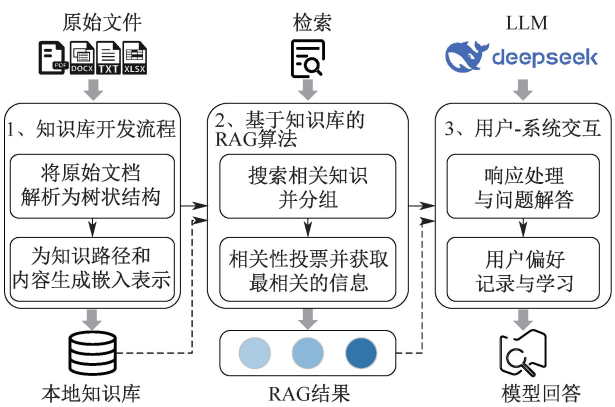


图 2　LLMs 与 RAG 组合结构

Fig. 2　Combined structure of LLMs and RAG

1.3　大模型本地部署与知识库构建方法

为兼顾数据隐私保护与低成本运行需求,基于 Windows 环境与 Ollama 工具,实现 DeepSeek R1 模型的本地化部署。选用模型为 DeepSeek-r1:32b,其经由 Qwen-7B 架构优化,具备卓越的逻辑推理与代码理解能力,有效保障了离线环境下的智能交互体验。在此基础上,构建可扩展的私有语义知识库以支持 RAG 流程。通过格式解析与清洗 Word 与 Markdown 文档,利用中文 Embedding 模型将文本转化为高维向量并构建索引,最终实现基于语义相似度的高效检索与增强问答。

2　知识库构建与问答模型设计

2.1　数据预处理与知识库构建

为构建高质量本地知识库,统计分析了 1 631 份核心文献,包含 629 份法规与 1002 份事故报告,涵盖成因分析与责任认定等关键信息。针对数据异构性,直接提取 845 份结构化文档文本,并引入 MinerU 工具深度解析剩余 786 份含复杂排版或扫描图的文档,完成多源文档的结构化重建与语义重构。全文数据可访问链接 https://www.scidb.cn/s/mQbQby 获取。

MinerU 文档处理包含 4 阶段:①文档解析,利用 PaddleOCR 提取图文位置信息以生成基础版面。②版面分析,融合 LayoutLMv3、YOLOv8 与 UniMERNet 等模型,精准识别表格、公式及复杂结构。③语义标注,微调 LayoutLMv3 以完成语义单元划分与逻辑构建;④结构化重建,依据自然阅读顺序重组内容,最终输出语义清晰的 Markdown 文档。

基于清洗后的结构化数据,利用 RAGFlow 框架构建本地事故语义知识库。该框架集成分块、嵌入与检索功能,首先,对文档进行语义切分以保障信息完整性;然后,利用预训练 Embedding 模型生成高维向量索引;最后,知识库与本地大模型实现无缝集成,通过领域知识的精准调用,显著提升问答系统的

准确性与上下文关联能力。

2.2　问答模型集成与流程设计

图 3 为模型以本地知识库为核心,集成问题理解、语义检索与答案生成流程,通过模块化设计实现高效协同,提升事故调查问答的响应速度与专业匹配度[16]。

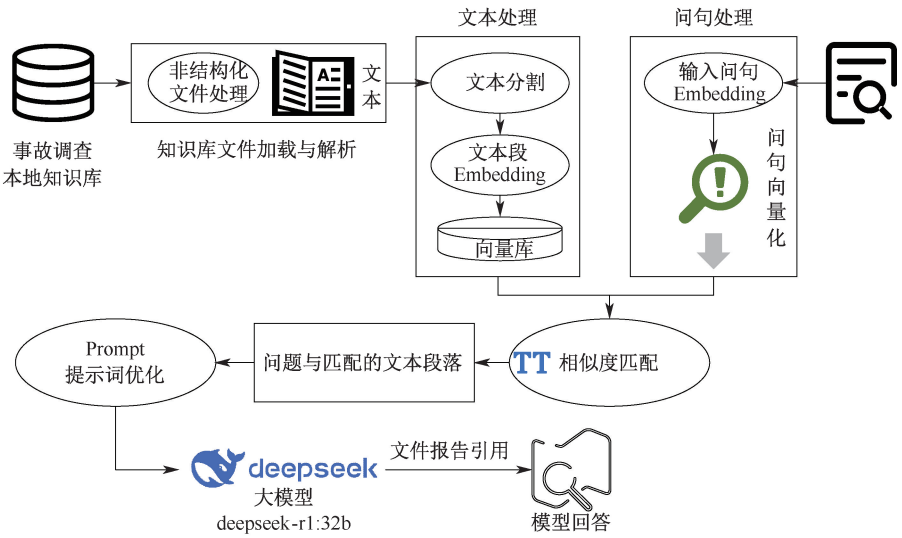


图 3　事故调查报告知识生成模型整体流程

Fig. 3　Overall process of the knowledge generation model for accident investigation reports

模型构建涵盖事故报告与法规的本地知识库,通过 Embedding 向量化建立语义索引。利用 DeepSeek-r1 32b 模型精准解析用户意图与实体关系,结合向量相似度与关键词的混合检索策略,筛选高相关内容构建精准语境上下文。

检索模块引入缓存与并行策略提升响应速度,将匹配文本注入 DeepSeek 模型生成专业回答。系统内置质量控制机制,从语义一致性与逻辑性维度优化输出,并强制嵌入引用标注,确保答案的专业性与可追溯性。

采用包含数据处理、语义检索及用户接口等 6 大单元的模块化集成架构。各模块通过明确的数据流协同,支持负载均衡与动态资源调度,有效保障多用户并发场景下的系统扩展性与运行稳定性。

生成端设置 Temperature 0.2、Top P 0.3 及惩罚系数(存在 0.4、频率 0.7),以抑制冗余与幻觉;检索端采用混合相似度加权(关键词 0.7、向量 0.3),设定阈值 0.2 并截取 Top 8 高相关片段,在保障上下文完整性的同时优化计算效率。

2.3　模型架构与功能模块实现

为满足语义理解与知识检索的综合需求,设计基于本地大模型与向量数据库的 4 层模块化系统架构。该架构包含数据处理层、知识构建层、问答服务层和用户交互层,分别承担数据清洗、索引构建、语义响应及界面呈现功能,具备良好的可维护性与扩展性。

数据处理层对 1 631 份文献(含 1 002 份事故报告和 629 份法规文本),利用 MinerU 平台完成 OCR 识别、版面还原与语义重排,准确提取复杂表格及非连续段落,将结果转化为结构化 Markdown 格式,为后续向量编码奠定基础。

知识构建层基于 RAGFlow 框架构建语义索引。模型以段落为单元,采用 DeepSeek 嵌入模型提取文本高维向量,推理过程由本地 Ollama 平台支持,通过构建向量索引库实现离线环境下的高效语义匹配。

问答服务层集成了意图识别、语义检索与回答生成模块。先解析用户输入以提取核心意图,通过向量相似度匹配相关文档片段并构建多轮上下文,并输入本地 DeepSeek 模型生成自然语言回答,保障输出质量。

用户交互层采用 Docker 容器技术实现 Web 前端与 Python/Java 后端的解耦部署。各模块通过环境隔离提升维护效率,前端通过端口 80 提供服务,并将 DeepSeek 服务端口 11434 映射至 80 以方便访问。结合异步通信与本地缓存机制,系统显著提升并发响应速度,全流程本地化处理有效保障了数据隐私与合规。

图 4 为模型整合了从 OCR 结构化、语义编码到交互推理的完整流程。该架构设计清晰、部署灵活且成本较低,为领域知识驱动的智能问答提供高效的技术路径。

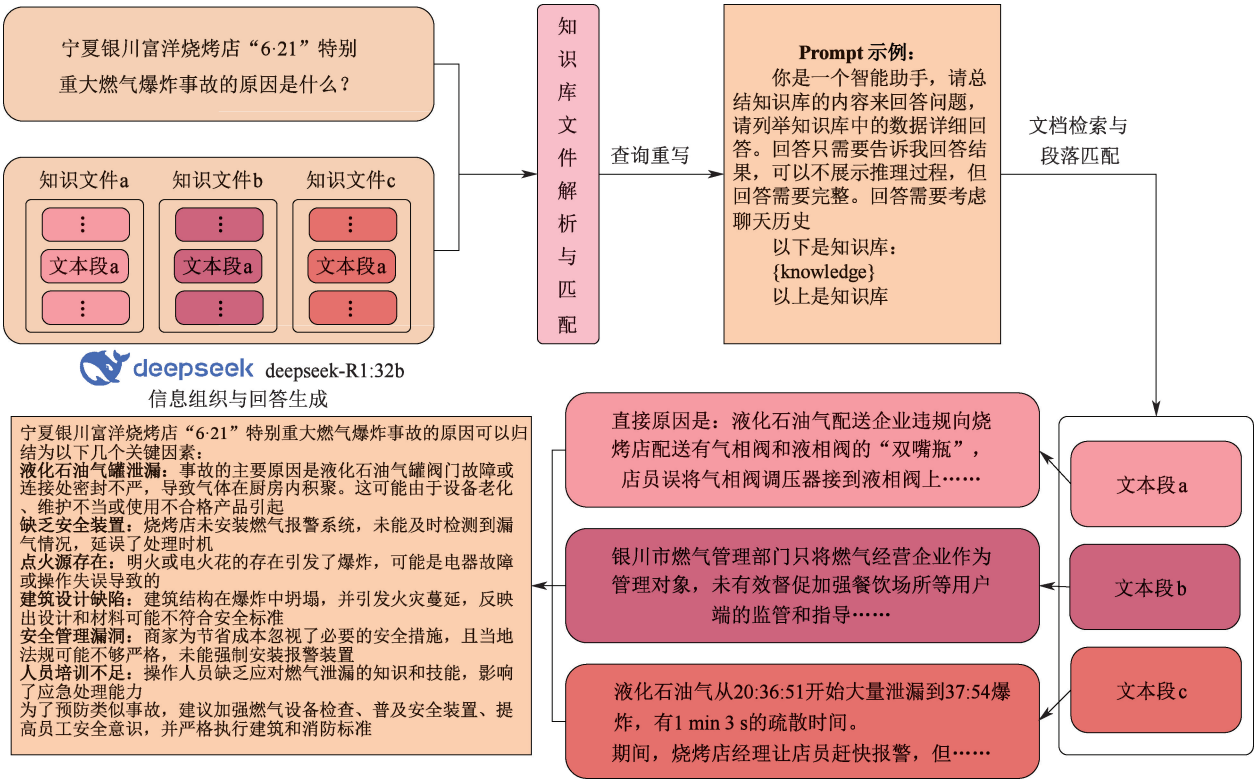


图 4　知识生成模型架构示例

Fig. 4　Example of knowledge generation model architecture

3　知识生成模型测试与评估

3.1　测试方案与评估方法

为验证模型生成内容的质量与实用性，设计包含自动与人工评估的综合评价体系，从正确性、完整性及表达效果等多维度进行验证。试验选取涵盖事故起因、过程复盘、责任判定及法规依据的 5 条典型问答任务（表 1），确保样本具备行业代表性与知识广度。过程设置对照组，分别采用基础模型（无检索）与构建的 RAG 模型生成回答，形成对比样本以分析检索增强机制的实际增益。

在自动评估方面，选用 Kimi、通义千问、豆包、文心一言及讯飞星火 5 款工具进行质量评价。构建标准提示词，要求模型扮演学术专家，针对“{input}”的回答从正确性、法规合规性、完备性、可读性、语义相关性 5 个维度进行评分（范围 1.000 0~5.000 0，保留 4 位小数）。利用大模型的高质量指令遵循能力，确保评价结果的稳定性与可复现性。

人工评估环节邀请 5 位具备安全工程、法律法规及应急管理背景的专家，独立评价 ChatGPT-4o、DeepSeek-V3、本地 Deepseek-r1:32b 及 RAGFlow 增强版 4 种模型的生成结果。试验采用双盲随机排序机制，覆盖事故处理与法规等多场景，依据同一指标

表 1　事故调查类问答任务问题类型设置

Table 1　Question type settings for accident investigation question-and-answer tasks

评估问题	问题类型
内蒙古阿拉善新井煤业有限公司露天煤矿“2·22”特别重大坍塌事故中，涉事企业未履行哪几项《安全生产法》规定的主体责任？请列举条款依据	法规依据与责任判定
宁夏银川富洋烧烤店“6·21”特别重大燃气爆炸事故的原因是什么	事故起因
请分析一个事故报告中关于“应急预案未执行”的判断逻辑，并指出企业因此可能违反的法律条文及处理建议	过程复盘与整改建议
请根据河南安阳市凯信达商贸有限公司“11·21”特别重大火灾事故，指出涉事单位未落实《消防法》中的哪几项要求？依据该事故调查结论，应追责的对象有哪些	责任判定与法规依据
事故报告中常提及“安全教育培训不到位”，请结合具体案例说明如何判断培训责任落实情况，并指出应依据的法律文件	责任判定与法规依据

体系打分(总分 25 分)。评审前已完成标准校准以保障公正性,最终形成“5 类问题×4 个模型”的 20 条基础评测数据。

3.2　知识生成模型性能与问答效果评估

大模型效果评估是 LLM 领域的重要研究问题,现有主流方法可分为特定任务指标、标准基准测试、自我评估和人工评估 4 类。生成的事故调查文本具有引用标识符、非唯一答案等特性,更具开放性与主观性,传统自动评估难以全面衡量其语义质量。因此,在自动化评分基础上,引入人工评估以补充验证。

1) 自动评估。为量化大模型在事故调查问答任务中的性能,构建自动评估体系,选取 Kimi、通义千问、豆包、文心一言和讯飞星火 5 种主流中文模型作为评估器,从正确性、法规合规性、完备性、可读性和语义相关性 5 个维度对 ChatGPT-4o、DeepSeek-V3、Deepseek-r1:32b 和 RAGFlow 的输出结果进行评分,评分区间为 1.000 0~5.000 0,保留 2 位小数增强分辨率。各问题得分见表 2,一致性检验结果表明:评分的 Cronbach's α 系数为 0.985,具备较高信度和稳定性。

表 2　自动化评估结果

Table 2　Automated Evaluation Results

测试模型	Kimi	阿里通义千问	豆包	文心一言	讯飞星火	Cronbach's α
Chatgpt-4o	114.25	115.10	106.00	106.10	106.70	0.985
Deepseek-V3	115.53	113.60	111.85	109.00	107.30	
Deepseek-r1:32b	100.22	102.35	91.40	93.70	87.00	
RAGFlow	106.65	109.70	97.20	100.15	97.50	

自动评估结果如图 5 所示,从 5 个维度综合分析,大模型性能差异明显。DeepSeek-V3 整体表现优异,评分稳定,特别是在正确性与语义相关性方面接近满分,展现出较强的语言理解与泛化能力。RAGFlow 显著优于其基础模型 DeepSeek-r1:32b,说明检索增强机制在提升回答准确性和相关性方面效果显著,适用于专业性任务。

从 5 种自动评分工具的评价看,ChatGPT-4o 与 DeepSeek-V3 得分显著高于 2 款本地部署模型,分别为 109.63 和 111.46,在线模型表现更优。值得注意的是,RAGFlow 在正确性维度相较基础模型提升 7.31 分,说明外部知识注入能有效增强事实性任务表现,微调策略亦对综合性能提升具有积极作用。

由图 5 可知:ChatGPT-4o 与 DeepSeek-V3 在事故调查问答任务中综合表现最佳,RAGFlow 在提升专业准确性方面展现出一定潜力。自动化评估体系为后续人工评估与模型优化提供了稳定、可信的量化基础。

2) 人工评估。邀请 5 位长期从事事故调查或安全管理研究的专家参与主观评价。试验围绕 5 个问题,利用 4 种模型生成事故调查相关文本,并采用匿名盲审形式进行打分。专家依据 5 项一级指标,从 1—5 分评分,结果见表 3。人工评估的评分信度(Cronbach's α)为 0.664,属中等一致性水平。

人工评估结果如图 6 所示,RAGFlow 在全部维度均优于其基础模型 DeepSeek-r1:32b,特别在文本真实性与思辨性方面表现突出,显示出其在学术性

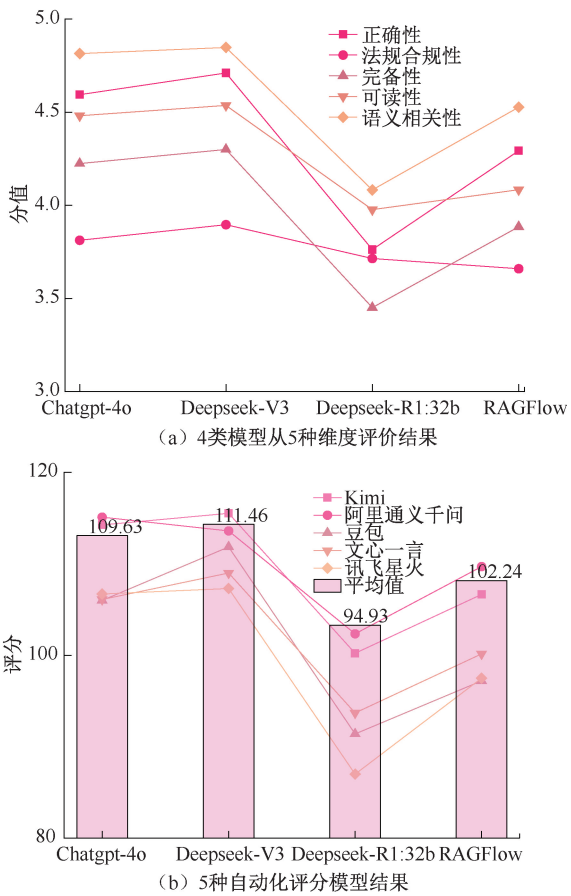


图 5　自动化评估结果

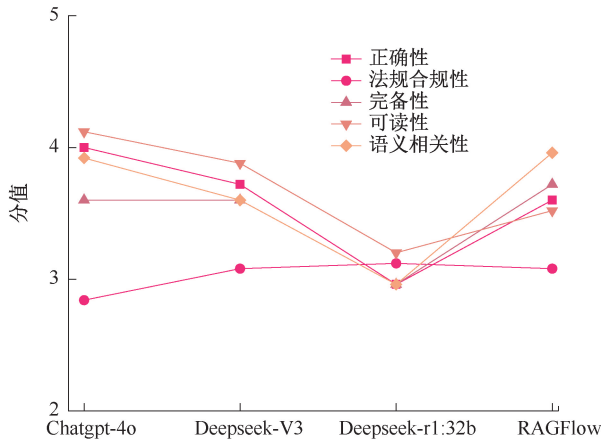
Fig. 5　Automated evaluation results

语言生成任务中的实际应用价值。微调后的模型在正确性与完备性方面提升明显,整体得分较基础模型提高 13.4 分,增幅达 17.6%。

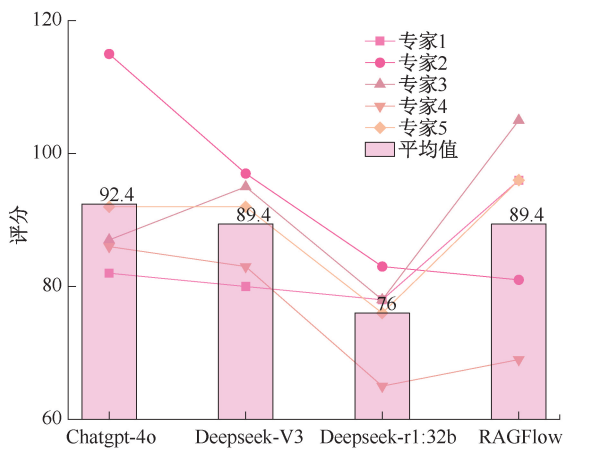
表 3 人工评估结果

Table 3 Manual evaluation results

测试模型	E_1	E_2	E_3	E_4	E_5	Cronbach's α
Chatgpt-4o	82	115	87	86	92	0.664
Deepseek-V3	80	97	95	83	92	
Deepseek-r1:32b	78	83	78	65	76	
RAGFlow	96	81	105	69	96	



（a）4类模型从5种维度评价结果



（b）5种人工评分模型结果

图 6 人工评估结果

Fig. 6 Manual evaluation results

3.3 LLM 问答效果对比

为评估不同 LLM 在突发公共安全事件中的问答能力，以宁夏银川富洋烧烤店“6 · 21”特别重大燃气爆炸事故为案例，调用 ChatGPT-4o（在线）、DeepSeek-V3（在线）、DeepSeek-r1:32b（离线）及其增强版 DeepSeek-r1:32b-RAGFlow（离线）4 种模型进行问答测试。

在信息准确性方面，ChatGPT-4o 与 DeepSeek-V3 能准确识别事故的直接原因（液化石油气泄漏引发爆炸）与间接原因（安全管理缺失、员工操作不当、监管不到位等），内容与国务院事故调查报告基本一致，表现出良好的事实把握能力。DeepSeek-r1:32b 未能输出有效内容，仅停留在情绪表达层面，明显偏离问答目标。RAGFlow 通过嵌入本地文档，在离线环境下较准确地复现事故要点，虽措辞略有差异，但内容可靠。

在信息完整性方面，ChatGPT-4o 结构分明，区分直接与多层次间接原因，覆盖钢瓶老化、违规操作、安全监管等多个维度。DeepSeek-V3 内容简洁，涵盖关键要素如燃气泄漏、管理失职与追责等，逻辑紧凑。RAGFlow 结合本地文档推演事故发展，进一步识别建筑设计缺陷、应急预案不足等次生问题，信息层次更为丰富。相比之下，DeepSeek-r1:32b 内容空泛，缺乏结构与技术细节支持。

在专业表达方面，ChatGPT-4o 使用“软管老化、燃气积聚、点火源”等术语，体现较强的工程专业性；DeepSeek-V3 偏向通俗表达，呈现新闻报道风格；RAGFlow 除复现爆炸机制外，还涉及阀门故障、报警装置、疏散设计等专业细节，具备一定学术分析价值。DeepSeek-r1:32b 未体现任何术语或技术深度。

在生成策略方面，ChatGPT-4o 与 DeepSeek-V3 表现出“摘要提取+结构化组织”的应答模式，快速聚焦问题核心；RAGFlow 融合文档信息与生成能力，形成“分析+归纳”型文本，更贴近人类专家写作风格；而 DeepSeek-r1:32b 未能根据任务指令有效生成内容，暴露其在离线状态下上下文建模与知识支撑能力的不足。

综合评估结果表明：ChatGPT-4o 在准确性、结构完整性与专业性方面表现最为突出，适用于突发公共事件中的报告重建与原因分析等高精度任务。DeepSeek-V3 回答简洁、逻辑紧凑，适合在线环境下的信息提取与摘要类问答应用。DeepSeek-r1:32b-RAGFlow 在本地文档支持下展现出较强的分析与推理能力，适用于安全管理领域中依托知识库的本地闭环系统。相比之下，基础模型 DeepSeek-r1:32b 缺乏上下文理解与事实支撑能力，难以胜任独立完成的信息挖掘与问答任务。

4 结 论

1）基于 DeepSeek LLM 与 RAG 机制构建的本地事故问答模型能够高效解析事故调查报告与法规文件，实现因果链条提取、法规匹配与语义映射，显著提升事故类复杂文本的信息提取能力与智能问答质量。

2）模型在实际测试中表现出良好的问答准确

性和用户响应速度,问答质量相比未集成 RAG 机制的模型在自动评分与人工评分中分别提升 7.7%与 17.6%,验证了语义建模与混合检索策略在事故知识服务中的可行性与优越性。

3) 尽管受限于本地大模型的参数规模与知识更新频率,该模型仍展现出较强的适配性和实用性,为事故调查、应急响应与安全管理提供智能化支持,具备在更多安全场景中推广应用的潜力。

参 考 文 献

[1] 刘瑞祥,柳先辉,赵卫东,等. 基于大语言模型的业务流程自动建模方法[J]. 计算机集成制造系统, 2025,31(6): 2 001-2 014.
LIU Ruixiang,LIU Xianhui,ZHAO Weidong, et al. Automatic business process modeling method based on large language models[J]. Computer Integrated Manufacturing Systems, 2025,31(6):2 001-2 014.

[2] 朱俊仪,朱尚明. 利用检索增强生成技术开发本地知识库应用[J]. 通信学报, 2024, 45(增 2): 242-247.
ZHU Junyi, ZHU Shangming. Development of local knowledge base application using retrieval augmented generation technology[J]. Journal on Communications, 2024, 45(S2): 242-247.

[3] 祝哲,周文倩,林婕,等. 大语言模型技术赋能的应急管理框架[J]. 指挥与控制学报, 2024, 10(6): 677-688.
ZHU Zhe, ZHOU Wenqian, LIN Jie, et al. Emergency management framework enabled by large language model technology [J]. Journal of Command and Control, 2024, 10(6): 677-688.

[4] 林鹏,向云飞,樊启祥,等. 基础设施工程安全智能化管控变革、挑战与思考[J]. 中国安全科学学报, 2024, 34(7): 8-19.
LIN Peng, XIANG Yunfei, FAN Qixiang, et al. Evolution, challenges, and thoughts on intelligent management and control for infrastructure engineering safety[J]. China Safety Science Journal, 2024, 34(7): 8-19.

[5] 唐海洋,刘振翼,陈东平,等. ChatSOS:基于大语言模型的安全工程知识问答模型[J]. 中国安全科学学报, 2024, 34(8): 178-185.
TANG Haiyang,LIU Zhenyi,CHEN Dongping, et al. ChatSOS: large language model-based knowledge Q & A system for safety engineering[J]. China Safety Science Journal, 2024, 34(8): 178-185.

[6] 王明达,赵宝熙,吴志生,等. 基于大语言模型的燃气事故调查报告实体识别[J]. 中国安全生产科学技术, 2025, 21(2): 139-145.
WANG Mingda,ZHAO Baoxi,WU Zhisheng, et al. Entity recognition of gas accident investigation reports based on large language model[J]. Journal of Safety Science and Technology, 2025, 21(2): 139-145.

[7] WAN Yuwei, CHEN Zheyuan, LIU Ying, et al. Empowering LLMs by hybrid retrieval-augmented generation for domain-centric Q&A in smart manufacturing [J]. Advanced Engineering Informatics, 2025, 65: DOI: 10.1016/j. aei. 2025. 103212.

[8] WU Chengke, DING Wenjun, JIN Qisen, et al. Retrieval augmented generation-driven information retrieval and question answering in construction management [J]. Advanced Engineering Informatics, 2025, 65: DOI: 10.1016/j. aei. 2025. 103158.

[9] WONG Saika, ZHENG Chunmo, SU Xing, et al. Construction contract risk identification based on knowledge-augmented language models[J]. Computers in Industry, 2024, 157/158:DOI:10.1016/j.compind.2024.104082.

[10] SUN Jianwen, SHI Wangzi, SHEN Xiaoxuan, et al. Multi-objective math problem generation using large language model through an adaptive multi-level retrieval augmentation framework[J]. Information Fusion, 2025, 119: DOI:10.1016/j. inffus. 2025. 103037.

[11] REN Tengfei, ZHANG Zhipeng, JIA Bo, et al. Retrieval-augmented generation-aided causal identification of aviation accidents: a large language model methodology[J]. Expert Systems with Applications, 2025, 278: DOI:10.1016/j.

eswa. 2025. 127306.

[12] SILVA L, BARBOSA L. Improving dense retrieval models with LLM augmented data for dataset search[J]. Knowledge-based Systems, 2024, 294: DOI:10. 1016/j. knosys. 2024. 111740.

[13] LEE J, AHN S, KIM D, et al. Performance comparison of retrieval-augmented generation and fine-tuned large language models for construction safety management knowledge retrieval[J]. Automation in Construction, 2024, 168: DOI: 10. 1016/j. autcon. 2024. 105846.

[14] ZHANG Haiyang, YAN Wei, HU Huicong, et al. An LLM-based knowledge and function-augmented approach for optimal design of remanufacturing process[J]. Advanced Engineering Informatics, 2025, 65: DOI: 10. 1016/j. aei. 2025. 103206.

[15] JEON J, SIM Y, LEE H, et al. ChatCNC: conversational machine monitoring via large language model and real-time data retrieval augmented generation[J]. Journal of Manufacturing Systems, 2025, 79: 504-514.

[16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: DOI:10. 48550/arXiv. 1706. 03762.

[17] 王文湖,韦昌法. 基于大语言模型和知识库的阿尔茨海默病智能问答系统构建研究[J]. 世界科学技术-中医药现代化, 2025,27 (3): 856-866.
WANG Wenhu,WEI Changfa. Research on the construction of an intelligent question-answering system for alzheimer´s disease based on large language models and knowledge bases [J]. Modernization of Traditional Chinese Medicine and Materia Medica-World Science and Technology, 2025,27(3): 856-866.

作者简介： 李 华 (1979—),女,陕西西安人,博士,副教授,硕士生导师,主要从事企业风险评估与安全管理、建筑安全监测与监控、公共安全与应急管理方面的研究。E-mail: lihua@ xauat. edu. cn。