文章编号:2096-1472(2026)10-0001-04
DOI:10.19644/j.cnki.issn2096-1472.2026.010.001

基于图检索增强生成技术的面向离散生产的客诉智能问答系统设计

蔡家俊,鲍 敏

(浙江理工大学机械工程学院,浙江 杭州 310018)

✉ 2216646833@qq.com; mbao@zstu.edu.cn

摘 要:针对离散生产客诉服务中技术文档非结构化、业务数据割裂以及大模型在该领域易产生“幻觉”的问题,设计了基于图检索增强生成的智能问答系统。提出了面向质量追溯系统的生产缺陷知识图谱与领域知识库构建方法,并融合 Leiden 社区检测算法与混合检索策略,增强了问答系统的准确性与可解释性。实验采用某轴承企业的售后历史数据验证,在 RAGAs(Retrieval-Augmented Generation Assessment score)评估框架中较通用 GraphRAG 方法综合得分提升了 8.60 分,验证了它在专业领域知识抽取、组织与应用方面的可行性与实用性。

关键词:离散生产;领域知识库;图检索增强生成;大语言模型

中图分类号:TH186;TP183 **文献标识码**:A

Design of a Graph Retrieval-Augmented Generation-Based Intelligent Question Answering System for Customer Complaints in Discrete Manufacturing

CAI Jiajun, BAO Min

(*School of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou* 310018, *China*)

✉ 2216646833@qq.com; mbao@zstu.edu.cn

Abstract: Addressing the issues of unstructured technical documents, fragmented business data, and the hallucination tendency of large language models in discrete manufacturing customer complaint services, an intelligent question answering system based on Graph Retrieval-Augmented Generation (GraphRAG) is designed. A method for constructing a production defect knowledge graph and a domain knowledge base oriented to quality traceability systems is proposed. By integrating the Leiden community detection algorithm with a hybrid retrieval strategy, the accuracy and interpretability of the question answering system are enhanced. Experiments are conducted using after-sales historical data from a bearing enterprise. Under the RAGAs evaluation framework, the system achieves an overall score improvement of 8.60 points over the generic GraphRAG method, verifying its feasibility and practicability in professional domain knowledge extraction, organization, and application.

Key words: discrete manufacturing; domain knowledge base; Graph Retrieval-Augmented Generation; Large Language Model

0 引言(Introduction)

随着我国制造业加速向数字化与智能化转型,离散生产行业在全生命周期管理环节积累了大规模的多源异构数据。与流程制造相比,离散生产具有工序高度解耦、工艺路径复杂及变更频繁等特征,导致它的机理知识多以非结构化或半结构化形式散落在各类孤立的信息系统中。在薄壁轴承等高端装备领域,生产与售后环节的数据割裂问题尤为突出。当产品出现售后质量投诉时,工程师需在海量技术文档与历史记录中进行人工检索,因此面临知识追溯链条不完整的困境。

离散生产场景下的客诉描述往往具有高度的语义模糊性,传统工业知识检索方法多依托关键词匹配技术,难以理解复杂的故障表征与深层的因果诉求。近年来,以 ChatGPT、千问(Qwen)为代表的生成式大模型在自然语言处理领域表现出卓越的交互能力,为复杂客诉问题的智能化处理提供了新路径。然而,通用大模型缺乏特定工业垂直领域的知识沉淀,在面对严谨的工业逻辑时极易产生“幻觉”现象[1]。在一般民用场景下,幻觉仅影响用户体验,但在涉及质量追溯的工业场景中,决策必须依托确凿的证据与标准规范。假数值或错误的因果关

收稿日期:2025-12-01

www.rjgczz.com

联不仅会误导工程师决策，还可能引发严重的产品安全事故与质量合规危机。

为此，本文遵循国家军用标准 GJB 9001C—2017 和 IATF 16949 标准[2]，提出一套面向离散生产的基于图检索增强生成（Graph Retrieval-Augmented Generation，GraphRAG）技术的客诉智能问答系统。研究通过该系统解决大模型在垂直领域的应用难题，并以某轴承企业的售后历史数据为例，验证该方案的可行性与实用性。

1 相关工作（Related work）

作为缓解大模型预训练数据滞后性及“幻觉”问题的核心技术，检索增强生成（Retrieval-Augmented Generation，RAG）已成为当前的研究热点。吴璇等[3]指出，RAG 通过检索外部知识库中的相关片段补充上下文，有效降低了模型对参数化记忆的过度依赖。然而，RAG 在处理结构复杂的工业文档时，易因文本切分碎片化导致语义关联丢失。RAG 这种“点对点”的检索模式难以支撑复杂的跨段落逻辑推理，成为制约它性能提升的关键瓶颈。

为了弥补向量检索在结构化逻辑表达上的不足，学术界引入知识图谱为 RAG 注入显式的语义约束。Mahalingam 等[4]提出了一种混合检索架构，通过整合结构化知识图谱与语义分块技术，显著提升了局部上下文的检索完整性。针对全局信息理解匮乏的问题，Edge 等[5]提出基于社区检测的全局摘要生成方案，利用图聚类算法表征节点间的语义簇，构建起从局部到全局的层次化认知。Pan 等[6]的研究表明，知识图谱能够显式建模实体间的演化路径，为大模型提供具备确定性推理轨迹的先验知识。周俭航等[7]明确指出，在处理涉及多跳逻辑的复杂查询时，GraphRAG 凭借子图检索与路径推理技术，检索精度与逻辑连贯性显著优于常规 RAG。目前，GraphRAG 的效能已在部分垂直领域得到初步验证，刘兆年等[8]基于通用 GraphRAG 构建油藏地质问答系统，利用地质实体的关联关系提升了专业问答的准确性，为 GraphRAG 在专业领域的应用提供了可参考的实践范式。

尽管 GraphRAG 已在特定场景取得突破，但将它迁移至离散生产的客诉服务场景仍面临诸多挑战：一方面，离散生产中的多源数据本体之间存在语义鸿沟，标准化图谱构建难度较大；另一方面，现有方案侧重于语义相似度，忽视了工业场景对质量可追溯性的严格要求。因此，如何依据行业标准构建高质量图谱并设计满足工业要求的检索追溯策略，仍是 GraphRAG 在该领域落地过程中亟待解决的重要问题。

2 系统总体架构设计（System architecture design）

为应对离散生产售后场景中技术文档结构化程度低、业务链条数据碎片化以及大模型易产生“幻觉”等挑战，本文提出一种面向轴承生产客诉领域的智能问答架构。该架构以通用 GraphRAG 框架为基础，通过知识图谱构建与混合检索推理，实现了领域知识的准确表示与高效调度。

（1）根据轴承生产与售后的历史数据[9]异构特征，提出基于业务语义的逻辑切分与自动化建图方法。利用大模型自动化抽取实体与关系，构建面向生产缺陷的知识图谱。通过整合向量索引与图谱结构化索引，形成双重知识表征，为精准检索提供全局语义支撑。

（2）在问答阶段，通过混合检索协同提取知识子图与上下文块，驱动大模型进行逻辑整合与推理。引入多模态反馈机制，生成基于原文溯源的文本解答，同时关联直观的图表数据，以显著增强故障诊断结果的可信度与可解释性。系统框架如图 1 所示。

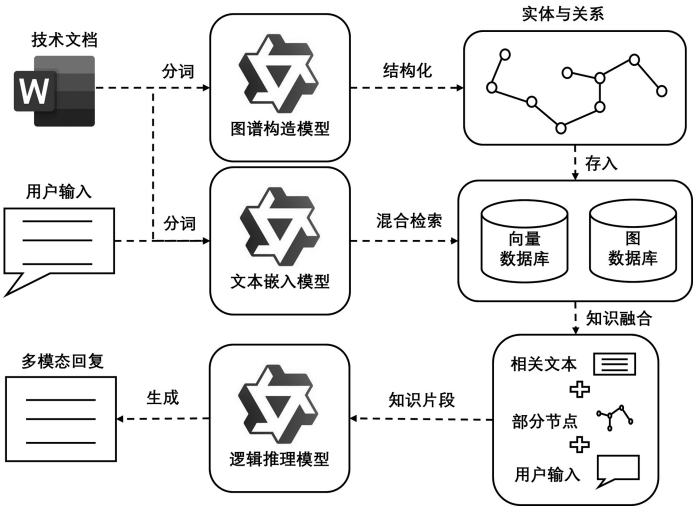

图 1 系统框架

Fig. 1 System framework

3 领域知识库与图谱构建（Construction of domain knowledge base and graph）

3.1 基于业务语义的数据预处理与逻辑切分

离散生产领域的技术文档（如 GJB/IATF 标准文件、8D 质量整改报告、轴承工艺卡片）专业性强且格式异构，包含章节、段落、图表等多层次组织，并广泛使用多样的术语体系。基于固定窗口的文本切分方法难以适应技术文档中的并列、嵌套及跨句表达，易导致语义割裂，影响实体识别和图谱一致性。

针对以上问题，本文设计了融合领域特征的语义切分策略。基于标题风格和编号模式构建领域正则表达式模板（如匹配“^(\d.)+[磨削|热处理]工艺”或“^5.1 振动检测结果”），自动识别文档的层级骨架。同时结合业务关键词（如“故障现象描述”“不合格品审理结论”）进行单元过滤与语义对齐，避免内容与标题脱节。通过 docx2python 库提取段落内容，并借助样式特征区分正文与图表说明。

在此基础上，文档被切分为语义统一的结构化单元，需要为每个单元标注来源。该处理保证了切分结果在语义上的一致性和层次清晰度，保留段落索引和路径信息，从而支持问答系统中的内容溯源与结果验证。数据处理流程如图 2 所示。

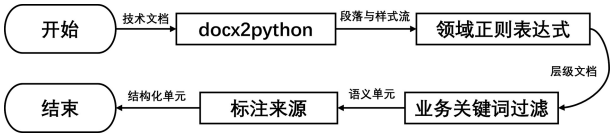

图 2 数据处理流程

Fig. 2 Data processing process

3.2 领域图谱的自动化构建方法

在完成基于业务语义的文档逻辑切分后，需要进一步将文本单元转化为可推理的结构化知识，以支撑质量追溯分析。在离散生产场景下，面对海量且动态更新的技术文档与客诉数据，传统知识图谱构建依赖人工标注，存在人力成本高昂、构建周期冗长等显著局限性。为适配高效数据处理与快速更新需求，本文提出一种基于大模型驱动的知识图谱自动构建[10]方法。该方法包含实体关系抽取与知识图谱生成 2 个阶段：

（1）针对离散生产领域存在大量专业术语和实体表述依赖上下文语境的痛点，设计了领域定制的提示词模板（如识别文本中的设备、工序及检测指标，以三元组形式列出）。通过预定义产品对象、生产工序、检测方法、缺陷类型、质量指标及处置

www.rjgcz.com

措施等 6 大核心本体,引导模型实现关键知识的精准识别与三元组提取。考虑到技术文档来源的多样性导致同义异构现象普遍,为避免抽取阶段直接消歧造成的语义丢失,本文采取先抽取后对齐的策略,将实体消歧与规范化延后至社区检测阶段统一处理,以提升全局对齐效率。为验证抽取效果,对 100 个随机样本进行人工校验,结果显示该阶段实体关系抽取整体准确率达 92%。关系抽取流程如图 3 所示。

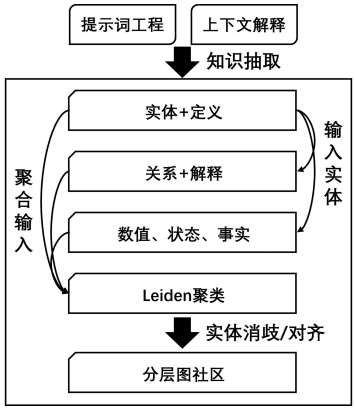

图 3　关系抽取流程

Fig. 3 Relationship extraction process

(2)在完成三元组抽取后,对原始结果进行结构化校验与规范化映射,构建基于属性图模型的领域知识图谱。图谱中节点代表实体,边代表语义关系,封装来源文档及段落等元数据属性以实现知识溯源。引入基于本体约束(如"工序"必为"影响"关系的主体而非客体)构建校验提示词,对生成的违反领域常识的关系进行判别与二次修正,显著提升了图谱的工程可用性。

为评估上述机制的有效性,本研究选取 6 个典型技术章节(约 5000 字符)作为实验语料进行图谱质量测试。实验过程中,模型累计消耗约 10 万 Tokens。共提取出 152 个实体及 255 条关系,并据此构建了一个小型知识图谱。该图谱的平均度数达到 3.4,表明实体间具有较强的语义联结度。从关系分布来看,因果关系占 40%,影响关系占 35%,包含关系占 25%,符合离散生产领域质量追溯的逻辑特征。通过误差分析结果发现,少数抽取错误(如将"内圈磨削工序,影响,游隙稳定性"误识别为"游隙测量,影响,磨削工序")主要源于长文本中的实体距离过远导致的逻辑主体偏移。图谱质量评估[11]结果见表 1。

表 1　图谱质量评估结果

Tab. 1 Results of graph quality assessment

衡量指标	说明	数值/%
Precision@10	前 10 条高置信度三元组准确率	90
人工抽样准确率	随机抽样 100 条三元组准确率	91

3.3　知识库的存储与组织

在完成自动化知识抽取与图谱生成后,如何高效管理规模庞大且语义迥异的工业知识成为核心挑战。通用的 GraphRAG 方案通常通过扁平化手段将结构化关系降维为纯文本存储,这在离散生产场景下极易导致工艺参数及因果链条等关键逻辑发生语义降解与关联断裂,难以支撑深度的质量追溯需求。为实现对轴承生产与客诉服务领域异构知识的高效管理和深度检索,本系统构建了向量数据库与图数据库[12]的协同部署架构:

(1)利用分布式向量数据库(如 Milvus)承载领域知识库[13]中的动态知识层,包括技术手册、故障报告及客诉记录。原始文档经由逻辑分块与向量化嵌入处理后,它的高维特征向量与原文索引被同步存入。该层利用向量库对近似最近邻检索的支持,确保在海量数据环境下实现对工业术语及模糊语义的快速更新与高效初步召回。

(2)针对上游生成的具备强逻辑关联的知识图谱,采用基于属性图模型的图数据库(如 Neo4j)进行持久化管理。与扁平存储不同,图存储模式完整保留了知识的拓扑结构,能够通过节点与边的显式关联支持复杂的多跳查询与隐式关系挖掘。这种结构化关联为质量问题的跨工序追溯提供了拓扑层面的推理能力。图数据库查询结果如图 4 所示。

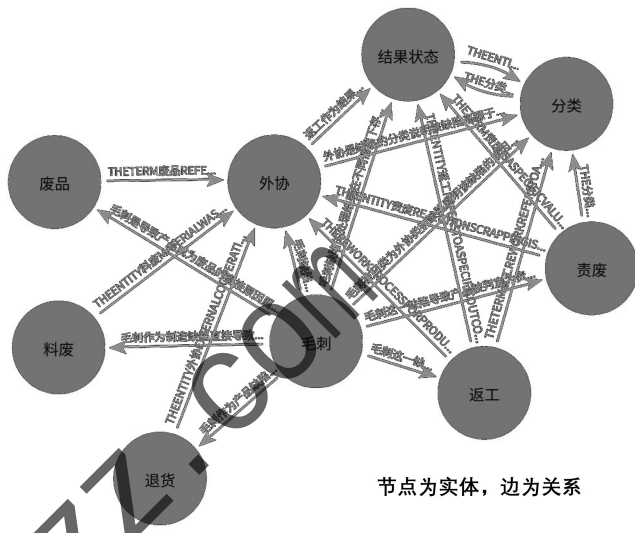

图 4　图数据结构

Fig. 4 Graph data structure

4　基于图增强的检索与生成机制(Graph Retrieval-Augmented Generation mechanism)

4.1　社区发现与摘要生成

生产缺陷知识图谱具有复杂的拓扑结构与异构的节点属性,而传统社区检测算法(如 Louvain 算法)在迭代过程中因缺乏对子图连通性的硬约束,易产生结构离散的非连通社区,导致检索阶段出现逻辑断层。为此,系统引入 Leiden 社区检测算法[14]。该算法通过改进局部移动机制与递归检测子社区,确保了划分出的每个社区在拓扑上均具备强连通性,从底层结构上规避了因检索空间碎片化而诱发的生成"幻觉"风险。

系统通过多轮递归执行该算法,构建了一套层次化社区体系。这种基于结构一致性的聚类过程,在拓扑层面辅助实现了实体的隐式消歧与同义关联对齐,提升了图谱数据的纯净度。通过由宏观至微观的知识单元划分,为后续的社区摘要生成提供了高质量输入,有效增强了系统对长程逻辑关联的建模能力。算法流程如图 5 所示。

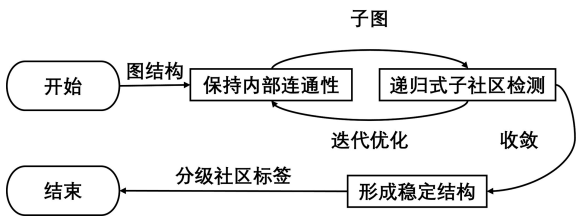

图 5　算法流程

Fig. 5 Algorithm flow

4.2　混合检索策略

在处理跨段落及跨文档的深度追溯任务时,基于单一语义

www.rjqcz.com

向量的检索方法难以捕捉工业领域复杂的逻辑关联。特别是在执行多跳关联查询时，由于缺乏拓扑结构的显式约束，检索过程极易引入无关噪声，产生偏离事实的“幻觉”现象，难以满足工业决策对结果确定性的严苛要求。

针对上述挑战，本文设计了一种融合语义匹配与图推理的混合检索策略。该策略严格遵循工作流，通过分层递进的知识处理机制，实现从广度召回到深度聚焦的精准知识定位。检索工作流如图 6 所示。

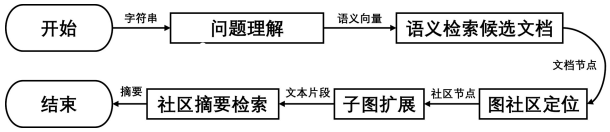

图 6 检索工作流

Fig. 6 Retrieval workflow

5 智能问答系统实现与评测(Implementation and evaluation of intelligent question answering system)

5.1 问答功能实现

在客诉服务场景下，故障判定往往依赖对复杂工况信息的精准解读，而传统纯文本回答模式难以承载多维工况数据的结构化表达，易出现信息歧义。针对上述问题，本系统设计并实现了多模态增强输出机制。该框架使系统能够提供多样的回答，通过图文、报表与原始文档溯源的协同展示(如参数表及散点图)，显著提升了推理结果的可信度与决策效率。系统界面如图 7 所示。

图 7 系统界面

Fig. 7 System interface

5.2 实验设置与指标

本节旨在评估本系统在真实轴承客诉场景下的实际效能。实验基于云服务平台构建计算环境，利用它的高可用计算资源与统一应用程序编程接口(Application Programming Interface，API)部署服务。在模型配置上，选用 qwen3-max 作为推理模型，text-embedding-v4 作为向量嵌入模型(Embedding)。鉴于对比的公平性，所有实验均在完全一致的模型配置、数据集及检索环境下进行。

为验证本系统的优化效果，设置两组基准模型作为对照组：(1)基础 RAG 模型。采用固定窗口(1200 tokens)对文档进行机械切分，不引入知识图谱，检索阶段仅依赖向量相似度召回切片，将结果直接拼接至提示词中。(2)通用 GraphRAG 模型。沿用固定窗口切分策略，但在预处理阶段利用通用提示词模板抽取实体与关系构建知识图谱，采用通用图谱检索逻辑生成回答。

实验构建了包含 30 组经专家审核的高质量问答对作为评测标准。数据集覆盖缺陷判定、故障溯源等典型业务场景(如请列出 XXX 轴承在振动检测中，缺陷的判定标准是什么?)。评估指标采用 RAGAs(Retrieval-Augmented Generation Assessment score)[15]评估框架，选取其中 4 个核心维度进行分析。鉴于工业故障诊断场景对事实准确性的严苛要求，采用加权平均法计算综合得分。权重分配为：真实性(F)占 40%、答案相关性(AR)占 30%、上下文精确率(CP)占 20%、上下文召回率(CR)占 10%。

5.3 结果分析

实验数据见表 2，对 3 组模型进行对比分析，可得出如下结论：

(1)对比基础 RAG 模型与通用 GraphRAG 模型，后者的真实性、答案相关性、上下文精确率和上下文召回率分别提高 9.0、1.0、7.5 和 1.7 个百分点，综合加权得分提升 5.57 分。这一结果表明，引入知识图谱能显著改善检索生成质量。基础 RAG 局限于局部语义相似性匹配，难以捕捉缺陷判定中隐式的因果链条关联，而图谱的结构化约束有效抑制了生成幻觉与上下文噪声，从而大幅提升了真实性与精确率。通用 GraphRAG 也暴露出局限性，它在答案相关性与召回率上的提升幅度较小，说明通用的实体抽取与检索策略在面对特定领域的复杂问题时，仍难以精准定位核心子图与实体。

(2)进一步对比通用 GraphRAG 与本系统，本系统的上述 4 个指标分别提高 6.0、12.0、7.7 和 10.6 个百分点，综合加权得分提升 8.60 分。这一显著差异验证了本系统针对轴承客诉问答场景优化策略的有效性。面向质量追溯系统的语义切分策略，使领域知识库在核心知识上的覆盖更加完整，提升了上下文召回率。领域定制提示词能优化实体抽取过程，保障图谱中的核心实体和关系，提供了可靠的事实支撑。Leiden 社区检测算法结合混合检索工作流，有效降低了无关上下文的引入概率，在提升上下文精确率的同时，确保检索过程紧密贴合用户查询的核心意图。

综上所述，工业级问答质量的突破不应仅依赖大模型推理能力的提升与通用框架的简单迁移，还需深耕高质量的领域图谱构建与检索策略。本系统成功解决了离散生产客诉服务场景下的推理难题，验证了它在轴承质量追溯领域的工程应用价值。表 2 中，F 为问答是否忠实于原文，AR 为答案不全或概念混淆，CP 为包含不相关的内容，CR 为是否遗漏关键实体，综合加权得分公式为$(4F+3AR+2CP+CR)/10$。

表 2 实验结果

Tab. 2 Experimental results

评估指标	F/%	AR/%	CP/%	CR/%	综合加权得分/分
RAG	85	80	60.9	85.5	78.73
GraphRAG	94	81	68.4	87.2	84.30
本文系统	100	93	76.1	97.8	92.90

6 结论(Conclusion)

本文针对离散生产客诉服务中技术文档非结构化、数据割裂以及大模型在该领域易产生“幻觉”的难题，设计了一套基于 GraphRAG 的智能问答系统。提出了面向质量追溯系统的知识图谱与领域知识库，有效解决了固定长度切分导致的语义割裂问题。融合了 Leiden 社区检测算法与混合检索策略，实现了从宏观社区到微观实体的精准定位。设计了多模态输出与溯源验证机制，增强了系统的可追溯性与可解释性。实验结果显示，该系统在 RAGAs 评估框架中较通用 GraphRAG 方法综合得分提升了 8.60 分，有效克服了通用大模型在工业垂直领域的应用缺陷，为离散生产企业实现客诉服务智能化转型提供了工程化解决方案。

(下转第 18 页)

www.rjgc.com

port[PP/OL]. V1. arXiv (2023-03-15) [2025-05-04]. https://arxiv.org/abs/2303.08774.
[8] Anthropic. Introducing Claude Sonnet 4.5[EB/OL]. (2025-09-30) [2025-05-04]. https://www.anthropic.com/news/claude-sonnet-4-5.
[9] Liu A, Feng B, Xue B, et al. DeepSeek-V3 technical report [PP/OL]. V1.arXiv (2024-12-30) [2025-05-04]. https://arxiv.org/abs/2412.19437.
[10] Bai J, Bai S, Chu Y, et al. Qwen technical report [PP/OL]. V1. arXiv (2023-09-28) [2025-05-04]. https://arxiv.org/abs/2309.16609.
[11] 鲁维海,魏华鹏,邓博元,等. 医学大语言模型中幻觉的防治[J]. 中国数字医学,2025,20(10):34-39,52.
[12] 陶立元,钟晓莹,代倩倩,等. 医疗保健领域大语言模型研究报告规范 TRIPOD-LLM 的应用范围解析与现状调查[J]. 中华医学杂志,2025,105(38):3379-3383.
[13] Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks [C/OL]// NeurIPS, 2020: 9459-9474. https://proceedings.neurips.cc/paper_files/paper/2020/hash/6b6e6f56826f826f90fb361c618d3608-Abstract.html.
[14] Hu E J, Shen Y, Wallis P, et al. LoRA: Low-Rank Adaptation of Large Language Models [PP/OL]. V1. arXiv (2021-06-17) [2025-05-04]. https://arxiv.org/abs/2106.09685.
[15] Arbelo E, Protonotarios A, Gimeno J R, et al. 2023 ESC guidelines for the management of cardiomyopathies: developed by the task force on the management of cardiomyopathies of the European Society of Cardiology (ESC)[J]. European Heart Journal, 2023, 44(37): 3503-3626.
[16] Liu X, Liu H, Yang G, et al. A generalist medical language model for disease diagnosis assistance[J]. Nature Medicine, 2025, 31(3): 932-942.

作者简介:

杨昌吉(1999-),男(汉族),汉中,硕士生。研究领域:人工智能应用。
张选德(1979-),男(汉族),银川,教授,博士。研究领域:人工智能应用。
贺惠新(1984-),男(汉族),厦门,高级工程师,博士。研究领域:智能信息处理、工业控制网络。
刘 婕(1993-),女(汉族),北京,住院医师,博士。研究领域:心肌病。
段岩峰(1980-),男(汉族),北京,高级工程师,博士。研究领域:智慧医疗。
黄 鑫(1981-),男(汉族),宁波,教授,博士。研究领域:人工智能应用。本文通信作者。

(上接第 4 页)

参考文献(References)

[1] 李铂鑫,鲁骁,张霄,等.面向检索增强式私有问答大模型的幻觉检测方法[J].软件导刊,2026,25(1):39-46.
[2] 李丽,陈克伟,王心瑶.军用车辆企业 GJB9001C 与 IATF16949 体系融合研究[J].重型汽车,2024,(3):35-36.
[3] 吴璇,付涛.检索增强生成技术研究综述[J].计算机工程与应用,2025,61(20):19-35.
[4] Mahalingam A, Gande V K, Chadha A, et al. SKETCH: structured knowledge enhanced text comprehension for holistic retrieval[C/OL]// GenAIK, 2025: 27-42. https://aclanthology.org/2025.genaik-1.4.pdf.
[5] Edge D, Trinh H, Cheng N, et al. From local to global: a graph rag approach to query-focused summarization [PP/OL]. V2. arXiv (2025-02-19) [2025-06-19]. https://arxiv.org/abs/2404.16130.
[6] Pan S, Luo L, Wang Y, et al. Unifying large language models and knowledge graphs: a roadmap [J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(7): 3580-3599.
[7] 周俭航,肖诗斌.图检索增强生成研究综述[J].人工智能与机器人研究,2025,14(2):402-413.
[8] 刘兆年,姜彬,王宁,等.基于图检索生成增强的油藏地质问答系统构建[J].石油科学通报,2025,10(5):1069-1082.
[9] 武天野.轴承制造业生产数据实时集成系统设计与实现[J].计算机时代,2022,(8):56-60.
[10] 徐浩,康振渊,张焱,等.面向电力变压器故障辅助决策的知识图谱构建及补全策略研究[J].中国安全生产科学技术,2025,21(5):46-54.
[11] Marchesin S, Silvello G, Alonso O. Large Language Models and data quality for knowledge graphs[J]. Information Processing & Management, 2025, 62(6): 104281.
[12] 周学文,王俊强,江荣.结合向量匹配的图数据库智能问答智能体设计与实现[J].信息与电脑,2025,37(23):51-53.
[13] 李靓果,薛志一,陈小红,等.结合大语言模型和领域知识库的证券规则规约方法[J].软件学报,2025,36(10):4671-4694.
[14] Traag V A, Waltman L, Van Eck N J. From louvain to leiden: guaranteeing well-connected communities[J]. Scientific Reports, 2019, 9(1): 5233.
[15] Es S, James J, Anke-Espinosa L E, et al. RAGAs: automated evaluation of retrieval augmented generation[C/OL]// EACL, 2024: 150-158. https://aclanthology.org/2024.eacl-demo.16.pdf.

作者简介:

蔡家俊(2000-),男(汉族),台州,硕士生。研究领域:智能制造,数字化及智能化。本文通信作者。
鲍 敏(1977-),男(汉族),杭州,副教授,博士。研究领域:智能制造,生产过程数据分析。

www.rjgcz.com