Computer Science and Application 计算机科学与应用, 2026, 16(11), 184-192
Published Online November 2026 in Hans. https://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2026.1611335

Hans汉斯

融合知识图谱与检索增强生成的地方文化智能问答系统

——以嘉兴为例

胡力豪，梅传榆，杨铭熙，陆贝妮*

嘉兴南湖学院信息工程学院，浙江 嘉兴

收稿日期：2026年10月7日；录用日期：2026年11月6日；发布日期：2026年11月13日

摘　要

嘉兴地方文化资源丰富，但智能化传播服务相对不足，传统问答方式在跨资源关联查询能力和回答事实可信度方面存在明显短板。针对上述问题，本文设计并实现了融合知识图谱与检索增强生成(Retrieval-Augmented Generation, RAG)技术的“嘉言善行”地方文化智能问答系统。系统基于Neo4j图数据库构建覆盖名人、非遗、地方史志三大主题的知识图谱，基于Chroma向量数据库构建向量语义知识库，采用“图谱精确查询 + 向量语义召回”的双引擎混合检索策略，由大模型借助RAG技术生成自然语言答案，并辅以检索约束、答案清洗、来源标注三层防幻觉机制。系统采用Spring Boot 3与Vue 3的前后端分离架构，已建成MySQL结构化数据800余条、向量知识库文档532条、知识图谱实体135个与关系85条，设计测试问题30个与推荐专题问题20个，后端单元测试10个用例全部通过，前后端联调验证正常。实验结果表明，系统能够针对地方文化问题提供事实可靠、来源可溯的问答结果，为地方文化数字化传播提供了一套可复用的轻量化技术方案。

关键词

知识图谱，检索增强生成，向量检索，地方文化，智能问答

A Local Culture Intelligent Q&A System Integrating Knowledge Graph and Retrieval-Augmented Generation

—A Case Study of Jiaxing

Lihao Hu, Chuanyu Mei, Mingxi Yang, Beini Lu*

School of Information Engineering, Jiaxing Nanhu University, Jiaxing Zhejiang

*通讯作者。

文章引用: 胡力豪, 梅传榆, 杨铭熙, 陆贝妮. 融合知识图谱与检索增强生成的地方文化智能问答系统[J]. 计算机科学与应用, 2026, 16(11): 184-192. DOI: 10.12677/csa.2026.1611335

Received: October 7, 2026; accepted: November 6, 2026; published: November 13, 2026

Abstract

Aiming at the abundant local cultural resources of Jiaxing but the lack of intelligent dissemination services, this paper designs and implements "Jiayan Shanxing", an intelligent Q&A system integrating knowledge graph and retrieval-augmented generation (RAG) for local culture. The system builds a knowledge graph covering celebrities, intangible cultural heritage (ICH) and red culture over the Neo4j graph database, and constructs a vector semantic knowledge base over the Chroma vector database. Through a dual-engine hybrid retrieval of "precise graph query plus vector semantic recall", answers are generated by a large language model with RAG, supported by a three-layer anti-hallucination mechanism of retrieval constraint, answer cleaning and source annotation. Built with Spring Boot 3 and Vue 3, the system has established more than 800 structured records in MySQL, 532 documents in the vector base, and 135 entities with 85 relations in the knowledge graph; 30 typical test questions and 20 recommended questions are designed, and all 10 unit test cases pass with front-end and back-end integration verified. Experimental results show that the system can provide fact-reliable and source-traceable answers to local culture questions, offering a reusable lightweight technical solution for local cultural digital dissemination.

Keywords

Knowledge Graph, Retrieval-Augmented Generation, Vector Retrieval, Local Culture, Intelligent Q&A

Copyright © 2026 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

嘉兴是国家级历史文化名城，文化资源底蕴深厚，既有马家浜文化遗址、嘉兴子城等地标性资源，也涌现出茅盾、丰子恺等文化名人，还拥有多项非物质文化遗产。然而，上述资源大多以纸质或静态数字化形式存在，线上展示主要依赖标题与关键词匹配，用户进行跨资源关联查询时需在多个资源库之间反复检索，缺乏系统性的语义关联机制；另一方面，通用大模型的训练数据主要来源于互联网公开内容，对嘉兴这类地方性历史名人、非遗项目及地方史志的具体信息覆盖有限，直接用于问答易出现事实性错误或无法作答的情况。归结起来，传统的地方文化问答在复杂问题检索能力、回答可信度和准确度三个方面均存在较为明显的短板。

为缓解上述问题，本文将知识图谱与检索增强生成(RAG)技术相结合，面向嘉兴名人、非遗与地方史志三大主题，设计并实现了“嘉言善行”地方文化智能问答系统。系统基于 Neo4j 图数据库构建地方文化知识图谱，基于 Chroma 向量数据库构建向量语义知识库，通过“图谱精确查询 + 向量语义召回”的双引擎混合检索为问答提供证据，最终由大模型生成自然语言答案，同时实现来源溯源与防幻觉约束。本文的主要工作与贡献包括以下三个方面：一是提出知识图谱与向量检索双引擎协同的地方文化知识服务架构，实现精确关系查询与语义模糊检索的优势互补；二是设计检索约束、答案清洗、来源标注三层防幻觉机制，保障文化类问答的事实可靠性与可解释性；三是给出一套面向地市级文化资源的轻量化工

程实现方案，全部组件开源可复现，便于迁移与推广。

2. 相关工作

2.1. 知识图谱在文化遗产领域的应用

知识图谱以结构化方式描述实体及其语义关联，在文化遗产领域已有较为成熟的应用。“数字敦煌”项目将洞窟、壁画、人物、题记等资源组织为知识图谱，支持主题检索与知识挖掘[1]；王晓光等构建敦煌壁画叙词表并实现关联数据发布，为壁画资源的语义组织奠定了规范基础[2]。中国历代人物传记资料库(CBDB)依托结构化数据库支持人物社会关系分析与可视化查询[3]；欧洲 Europeana 平台聚合各国文博机构的数字藏品，基于语义网技术支持多语言检索与跨馆查询[4]。在方法研究层面，李永卉等以宋代镇江诗词为例，采用本体建模方法构建知识模型并存入 Neo4j 图数据库[5]；彭博等系统梳理了文化遗产信息资源知识图谱的构建方法[6]。上述项目多为国家级或大型文化遗产项目，投入较大、周期较长；面向地市级地方文化资源的轻量化知识图谱问答应用仍较为少见，为本项目留下了明确的研究空间。

2.2. 检索增强生成技术

大规模预训练语言模型在文本生成方面表现出色，但存在“幻觉”问题，即生成内容与事实不符。Lewis 等提出检索增强生成(RAG)技术，通过从外部知识库检索相关内容并将问题与检索结果一并输入模型，使生成过程建立在可核查的事实依据之上[7]。RAG 通常依赖向量数据库，经文档切分、向量编码与相似度召回实现知识检索，已在企业知识库问答、法律文书检索、医疗辅助诊断等垂直领域落地应用。在知识库问答领域，王守会等系统梳理了基于知识库的问答系统研究进展[8]，闫悦等从问答系统的技术演进角度进行了综述[9]；在大模型融合知识图谱方面，王帅等提出了融合知识图谱与向量检索的问答系统[10]，冯拓宇等对大语言模型增强的知识图谱问答研究进行了综述[11]，叶蕾等则探索了利用知识图谱实现多跳可解释问答的方法[12]。然而，在地方文化领域，尤其是地市级城市文化场景下，将知识图谱结构化知识与向量语义检索融合的 RAG 应用仍较为少见，本文在此方向上进行了工程探索。

2.3. 多智能体协作与图检索增强生成

近年来，将多个智能体组织起来协作完成复杂任务的思路在问答系统中受到广泛关注。多智能体框架可将复杂问题分解为若干子任务，由专门智能体分别执行后汇总结果，在处理跨领域、多跳推理问题时较单智能体更具条理性。与此同时，图检索增强生成(GraphRAG)作为一种将知识图谱结构引入 RAG 流程的技术路线，通过实体关系图实现全局语义检索与摘要生成，在处理复杂查询时表现出优于纯向量 RAG 的能力[13]。李倩等以图书馆领域为例，构建了多智能体协作框架下的图检索增强生成式问答系统，经实体识别、知识提取与 Neo4j 图存储构建知识图谱，本地化部署大模型并链接图谱，由多智能体协作完成问题分解与执行，实验表明加入多智能体框架后 BLEU-4、ROUGE-L、METEOR 三个指标分别提升 11.7%、11.4%、15.7% [14]。不过，此类方案通常需要本地化部署大模型和较为复杂的智能体调度机制，部署成本与运维门槛较高。本项目面向地市级地方文化资源的轻量化传播需求，未采用完整的多智能体调度框架，而是通过“图谱精确查询 + 向量语义召回”的双引擎并行检索实现类似的问题分解效果，在保证问答质量的同时将部署成本控制在较低水平。

3. 系统设计与关键技术

3.1. 总体架构

系统整体架构自底向上分为数据层、知识构建层、检索引擎层、生成应用层与交互层五层，如图 1 所

示。数据层负责多源数据的采集、清洗、去重与标注；知识构建层运行知识图谱与向量知识库两条并行构建流水线，分别完成三元组抽取导入与文档切片向量化；检索引擎层提供图谱精确查询与向量语义召回两类检索服务，并对检索结果进行融合与去重；生成应用层基于检索上下文调用大模型生成自然语言答案，同时执行答案清洗与来源标注；交互层面向用户提供问答对话、知识图谱可视化与资源检索等功能。架构设计遵循四项原则：一是数据与检索解耦，知识图谱与向量库独立构建、统一调度；二是检索与生成解耦，检索结果以结构化上下文形式传入生成模块；三是降级可用，任一外部服务异常时系统自动回退至轻量本地方案；四是来源可溯，每条回答均附带来源标识。

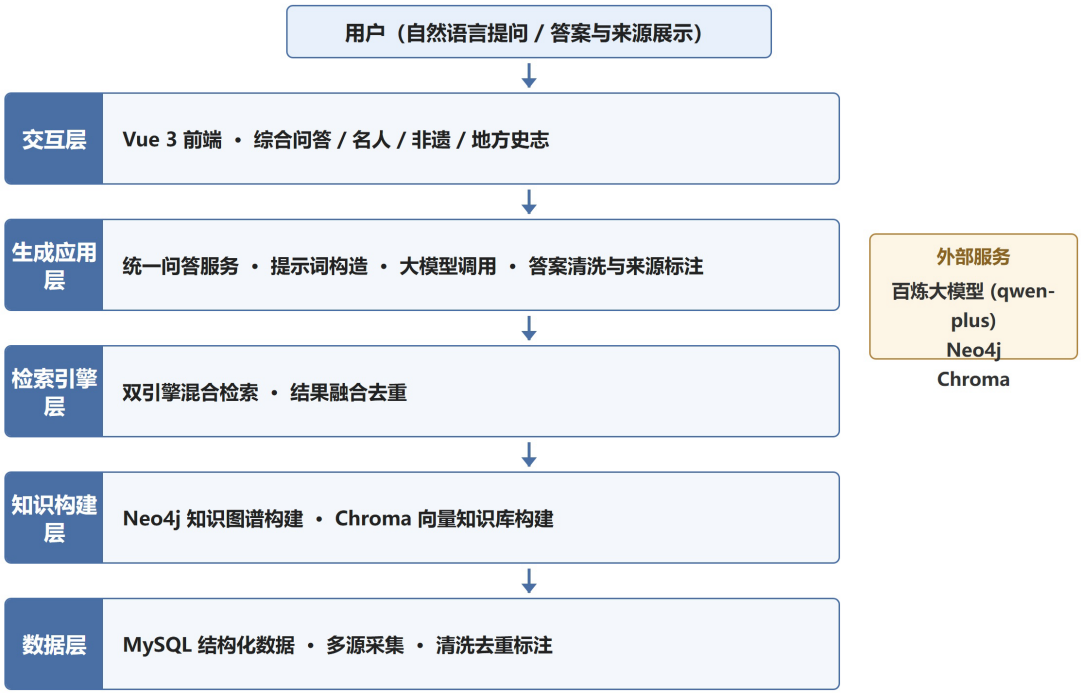


Figure 1. Five-layer architecture of the system
图 1. 系统五层架构示意图

3.2. 知识图谱构建

知识图谱构建分为本体设计、三元组抽取与数据导入三个环节。本体定义了名人、非遗项目、地方史志场馆、地方史志人物、历史事件、地区、作品与成果、传承人等实体类型，以及“属于主题”“位于地区”“关联人物”“代表作品”等关系类型，形成覆盖三大主题的本体模式。数据来源主要包括《嘉兴市志》等地方史志文献、浙江省非物质文化遗产名录官网公布的代表性项目信息、嘉兴市文化广电旅游局公开资料，以及经人工校验的百科类公开条目。三元组抽取采用人工校验与规则辅助相结合的方式：首先使用 jieba 分词与词性标注工具对原始文本进行预处理，基于人名、地名、机构名等词性规则识别候选实体，再通过预设的关系触发词(如“出生于”“代表作”“隶属于”)匹配实体间关系，最后由人工对歧义实体进行消歧与合并，经 Cypher 脚本批量导入 Neo4j。

在图谱访问层面，系统将独立的图谱服务组件进行封装，统一提供四类查询能力：一是按业务模块获取图谱节点与边，用于可视化展示；二是匹配图谱问题画像，支撑精准问答；三是根据关键词查找相关实体及其关联关系；四是查询实体的一跳或多跳邻居，支持关联推理。当 Neo4j 服务不可用时，系统自动回退至本地 JSONL 种子数据，保障基础查询功能可用。

以几条典型抽取规则为例：当文本中出现“×××，嘉兴××人”模式时，抽取三元组(×××，位于地区，嘉兴)；当出现“×××的代表作有《×××》”模式时，抽取(×××，代表作品，×××)；当出现“×××被列入第×批国家级非物质文化遗产名录”时，抽取(×××，属于主题，非遗项目)及(×××，级别，国家级)。上述规则以正则表达式实现，覆盖项目中约 70%的三元组，其余复杂关系由人工补充。

图谱可视化效果如图 2 所示，图中展示了名人、非遗项目与地方史志实体及其关联关系。

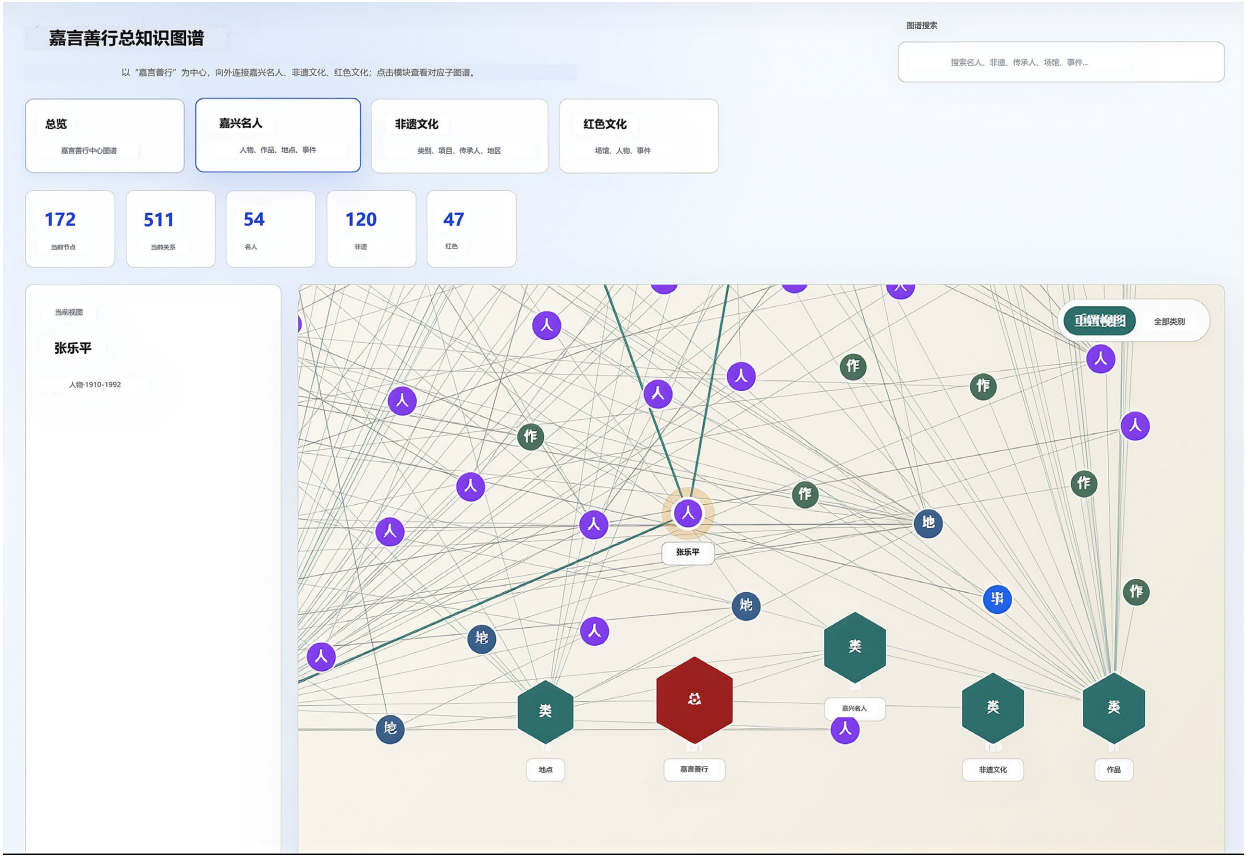


Figure 2. Knowledge graph visualization of Jiayan Shanxing
图 2. “嘉言善行”知识图谱可视化

3.3. 向量知识库构建

向量知识库用于支撑语义相似度检索。系统设计上采用多语言预训练嵌入模型将文本编码为 384 维向量并存入 Chroma，索引采用分层可导航小世界图(Hierarchical Navigable Small World, HNSW) (ef_construction = 200, M = 16)，在召回率与构建速度之间取得平衡。在实际联调过程中，为保障导入端与查询端的嵌入一致性，系统额外实现了基于 SHA-256 的确定性哈希嵌入(stableEmbedding)作为兜底方案：当预训练嵌入模型不可用或离线导入环境与在线查询环境不一致时，自动切换至哈希嵌入，确保同一段文本在不同环境下生成一致的向量表示。需要说明的是，哈希嵌入不具备语义泛化能力，仅作为一致性保障的降级手段，线上语义检索的主路径仍为预训练嵌入模型。知识库按名人、非遗文化、地方史志、综合模块组织，包含知识片段、问答提示、语义扩展、知识深化、查询路由等切片角色，每条记录携带 chunkId、标题、内容、来源名称、来源 URL、模块、实体类型等元信息。截至报告期，向量知识库共 532 条文档，覆盖三大主题的说明文本、讲解词、问答材料与来源信息。

3.4. 双引擎混合检索与 RAG 生成

面对用户提问，系统处理流程分为问题理解、并行检索、结果融合、答案生成与清洗五个步骤。问题理解阶段首先调用大模型对用户输入进行实体识别与意图判定，将识别出的实体与本体类型进行匹配，并依据预定义的问题模板生成对应的 Cypher 查询语句。例如，当问题包含“×××和×××有什么关系”模式时，生成 MATCH (a:Entity {name:$e1})-[r]-(b:Entity {name:$e2}) RETURN r 形式的查询；当问题为“×××的代表作有哪些”时，生成 MATCH (a:Entity {name:$e})-[:代表作品]->(b) RETURN b.name 形式的查询。随后并行执行两条检索路径。第一条为向量语义召回：将问题编码为 384 维向量，在 Chroma 中执行 Top-K (K 取 5~6)相似度检索，召回语义最相关的文本片段；当 Chroma 不可用或无召回结果时，自动回退至本地 JSONL 种子检索，依据关键词重合、标题匹配等规则返回候选片段。第二条为图谱结构化查询：将问题理解阶段生成的 Cypher 语句提交至 Neo4j 执行，获取实体节点、关系路径及关联属性；若 Neo4j 服务不可用，则回退至本地 JSONL 种子数据中的实体关系快照。

两条路径的检索结果统一进入融合模块：向量检索按相似度降序取前 5 条，图谱检索取前 8 条三元组与前 8 个相关实体；对同一文本片段按 chunkId 判定重复命中并仅保留一条，随后将片段、三元组与实体分别标注来源后拼接为上下文，交由大模型生成答案。该策略以语义检索结果为主体、结构化证据为补充，不引入加权系数，以避免参数调优带来的额外开销。

提示词模板包含系统角色约束、检索上下文与用户原始问题三部分，结构如下：“你是嘉兴地方文化问答助手。请严格依据以下检索到的资料回答用户问题，不得编造资料中未提及的内容；若资料不足以回答，请直接说明根据现有资料无法回答。【检索资料】{context}【用户问题】{question}”。其中{context}为经上述去重后拼接的知识片段与图谱证据，{question}为用户原始输入。当 API Key 未配置或模型调用失败时，系统自动回退至基于模板的规则化回答，保证演示链路不中断。

3.5. 防幻觉与来源溯源

防幻觉机制从三个层面构建。第一层为检索约束：Prompt 中明确要求模型仅基于检索到的知识片段与图谱证据作答，证据不足时直接回答“根据现有资料无法回答该问题”，禁止依赖模型参数记忆编造事实。第二层为答案清洗：后端在返回前执行兜底清洗逻辑，移除模型可能泄露的内部证据段与提示词残留，确保用户获得面向公众的自然语言回答。第三层为来源标注：清洗后的回答统一附加“来源：×××”说明，来源名称取自向量片段的元信息，便于用户核实。

4. 系统实现与实验

4.1. 系统实现

后端基于 Spring Boot 3.3.5 构建，采用 Java 17，集成 MyBatis-Plus 3.5.7 访问 MySQL，通过 Neo4j Java Driver 访问图数据库，使用 WebFlux WebClient 访问 Chroma 与大模型的 HTTP 接口，缓存由 Redis 管理。后端按业务域拆分为 redculture、celebrity、heritage、inheritor、jiayanshanxing、region、graph、qa、search、vector、stats、news 等模块，统一封装 ApiResponse 响应结构与全局异常处理。为消除前端各模块重复维护问答接口的问题，系统设计了统一问答接口：通过模块归一化函数将前端传入的模块名称映射为内部模块编码(jys, heritage, celebrity, red)，所有问答请求均进入统一问答服务流程，返回统一的回答、向量来源、图谱三元组与相关实体结构；推荐问题列表由后端从知识种子中动态读取，数据源统一且便于扩展。

前端基于 Vue 3 + Vite + TypeScript 构建，状态管理采用 Pinia，路由采用 Vue Router，整体包含嘉言善行主平台以及名人风采、非遗文化、地方史志三个子平台，提供综合问答、知识图谱可视化、资源检

索与分类展示等功能。知识图谱视图支持按模块展示节点关系，节点按类别分区排列，有效控制线条重叠。前端所有数据均通过 API 从后端获取，未使用 mock 数据。

4.2. 数据规模

截至报告期，项目已建成三类数据资产，具体规模见表 1。申报书设定的量化指标为“数据集 500 条以上、图谱实体 40 个以上、关系 60 条以上、向量库 150 条以上”，当前规模均已达标。需要指出的是，本系统定位为面向地市级文化资源的轻量化方案，数据规模的设计原则并非追求海量覆盖，而是在保障三元组质量的前提下实现可持续扩展：知识图谱构建采用“人工校验 + 规则辅助”的半自动化抽取流程，新增实体与关系可通过 Cypher 脚本批量导入；向量知识库的切片与元信息规范已固化，后续接入新的文化资源仅需按既定流程执行数据清洗与向量化导入，无需调整系统架构。

Table 1. Scale of data assets
表 1. 数据资产规模统计

数据资产	存储	实测规模
结构化数据	MySQL	800 余条
向量知识库	Chroma	532 条文档
知识图谱	Neo4j	135 个实体、85 条关系

4.3. 测试问题集与问答效果

为验证问答效果，项目设计 30 个典型测试问题，覆盖非遗项目、地方史志、名人风采与跨域综合四个主题，分布情况见表 2；另设 20 个推荐专题问题(名人 7 个、非遗 7 个、红色 6 个)，用于前端展示与演示。上述问题的问答提示、图谱画像与向量数据均已入库。评估方式以人工评估为主，从事实准确性与来源可溯性两个维度对每个测试问题的回答进行判定。后端实现统一问答控制器测试、统一问答服务测试(含“模型泄露证据清单”回归场景)与地方史志服务测试共 3 个测试类，经 Maven 测试 10 个用例全部通过。

Table 2. Distribution of test questions by theme
表 2. 测试问题集主题分布

主题	问题数量	示例问题
非遗项目	14	嘉兴有哪些代表性的非遗项目？海宁皮影戏有什么特点？
地方史志	8	嘉兴子城的历史沿革是什么？
名人风采	7	茅盾和乌镇有什么关系？王国维对中国近代学术有什么贡献？
跨域综合	1	从知识图谱角度看，嘉兴子城、嘉兴粽子、茅盾之间可能有哪些关联？

以“嘉兴粽子为什么能成为非遗文化代表？”为例，系统首先召回非遗知识库中的相关片段与图谱节点，构造提示词后由大模型生成面向用户的回答。回答从制作技艺、端午民俗、活态传承三个维度展开，提及选料浸米、裹粽烧煮等完整工艺流程，以及姚九华、冯月明等非遗传承人和五芳斋等老字号品牌，回答末尾标注“来源：嘉言善行二十问：非遗项目问答专题”。图谱联动层面，系统识别出嘉兴粽子、南湖区、传统技艺、端午文化、非遗传承人等核心实体及其关系，右侧面板同步展示文本数据来源与图谱关系依据。对该示例进行复测，回答正文能直接回应用户问题，来源标注与知识库片段一致，未出现内部证据泄露情况。问答过程与来源标注效果如图 3 所示。



Figure 3. Example of Q&A with evidence tracing
图 3. 问答示例与证据溯源

为进一步验证双引擎混合检索的有效性，项目在 30 个测试问题上开展了消融对比实验，分别考察纯大模型、仅向量 RAG、仅图谱查询、双引擎融合四种配置在事实准确性与来源可溯性两个维度上的表现。纯大模型配置不接入任何外部知识库，直接由 qwen-plus 生成回答；仅向量 RAG 配置关闭图谱查询路径，仅依赖 Chroma 语义召回；仅图谱查询配置关闭向量召回，仅依赖 Neo4j 实体与关系查询；双引擎融合为系统默认配置。经人工评估，结果如表 3 所示。双引擎融合方案在事实准确性与来源可溯性上均优于其余三组；仅向量 RAG 在语义模糊问题上召回率较高但存在证据不精确的问题；仅图谱查询在事实性问题上回答准确但覆盖范围有限；纯大模型在地方文化细节上易出现幻觉。上述结果验证了双引擎混合检索策略在精确性与覆盖率之间的平衡价值。

Table 3. Results of ablation experiment (manual evaluation, N = 30)
表 3. 消融实验结果(人工评估均值，N = 30)

配置	事实准确率	来源可溯率
纯大模型	52%	0%
仅向量 RAG	72%	65%
仅图谱查询	80%	90%
双引擎融合	90%	92%

4.4. 系统鲁棒性与降级机制验证

开发过程中，项目对系统的鲁棒性与降级机制进行了针对性验证，主要涵盖以下两类场景。一是数据服务异常场景：在非遗数据查询中模拟数据库连接被拒的故障，系统触发种子数据回退机制，返回预置的轻量化答案，验证了“降级可用”设计原则的有效性，修复访问配置后系统恢复正常。二是问答输出安全场景：开发初期发现大模型回答存在泄露内部证据段与提示词残留的问题，通过加强 Prompt 约束与返回前兜底清洗两层机制予以修复，并新增“模型泄露证据清单”回归测试用例防止复发，验证了“防幻觉”设计原则的可靠性。此外，针对知识图谱可视化中节点堆叠的问题，采用按类别分区布局、控制

节点相对大小与线条重叠的策略进行优化；针对多源数据重复问题，实现按名称去重的清洗逻辑并重新导入，保障 MySQL 与 Neo4j 数据一致性。

5. 结论

本文面向嘉兴地方文化数字化传播需求，设计并实现了融合知识图谱与检索增强生成的“嘉言善行”智能问答系统。系统以 Neo4j 知识图谱与 Chroma 向量知识库为双引擎，通过双引擎混合检索、大模型 RAG 生成、防幻觉约束与来源溯源等机制，实现了对嘉兴名人、非遗、地方史志三类资源的智能问答与可视化展示。数据资产规模超过申报书量化指标，单元测试与前后端联调均通过验证，为地方文化数字化传播提供了一套可复用的轻量化技术方案。

当前系统仍存在以下不足：向量嵌入的兜底方案采用确定性哈希算法，在预训练模型不可用时语义检索能力明显下降，且当前多语言预训练嵌入模型对中文地方文化术语的语义适配仍有不足，后续拟换用中文专用嵌入模型并重建向量库；大模型依赖外部 API，存在一定的成本与网络依赖；测试评估仍以人工评估为主、样本规模有限，尚未引入自动化评估指标；双引擎并行检索虽能处理大部分问题，但在需要多步推理的跨域复杂问题上，缺乏多智能体框架所具备的显式问题分解与任务调度能力。后续工作包括：引入本地化轻量模型与缓存策略、基于测试问题集建立 BLEU、ROUGE 等量化评估体系、探索引入多智能体协作框架以提升跨域综合问题的回答能力，以及将系统方法迁移至其他文化资源鲜明的城市，形成可复制的地方文化智能问答平台。

基金项目

本论文的工作受到浙江省大学生创新创业训练计划项目(省级，创新训练项目)“嘉言善行”——知识图谱与向量检索融合的嘉兴地方文化智能问答系统(项目编号：268519079)的资助。

参考文献

[1] 王晓光, 谭旭, 夏生平. 敦煌智慧数据研究与实践[J]. 数字人文, 2020(4): 11-23.

[2] 王晓光, 侯西龙, 程航航, 等. 敦煌壁画叙词表构建与关联数据发布[J]. 中国图书馆学报, 2020, 46(4): 69-84.

[3] 中国历代人物传记资料库(CBDB) [EB/OL]. https://cbdb.hsites.harvard.edu/, 2026-09-20.

[4] Europeana Foundation: Europeana Collections. https://www.europeana.eu/, 2026-09-20.

[5] 李永卉, 周树斌, 周宇婷, 等. 基于图数据库 Neo4j 的宋代镇江诗词知识图谱构建研究[J]. 大学图书馆学报, 2021, 39(2): 52-61.

[6] 彭博. 数字人文视角下的网络文物信息资源知识图谱构建[J]. 图书馆论坛, 2023, 43(2): 57-66.

[7] Lewis, P., Perez, E., Piktus, A., *et al.* (2020) Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Proceedings of the* 34*th International Conference on Neural Information Processing Systems*, Vancouver, 6-12 December 2020, 9459-9474.

[8] 王守会, 覃飙. 知识库问答系统研究进展[J]. 小型微型计算机系统, 2021, 42(9): 1793-1801.

[9] 闫悦, 郭晓然, 王铁君, 等. 问答系统研究综述[J]. 计算机系统应用, 2023, 32(8): 1-18.

[10] 王帅, 何文春, 王甫棣, 等. 大语言模型融合知识图谱与向量检索的问答系统[J]. 科学技术与工程, 2024, 24(32): 13902-13910.

[11] 冯拓宇, 李伟平, 郭庆浪, 等. 大语言模型增强的知识图谱问答研究进展综述[J]. 计算机科学与探索, 2024, 18(11): 2887-2900.

[12] 叶蕾, 张宇迪, 杨旭华. 利用知识图谱的多跳可解释问答[J]. 小型微型计算机系统, 2024, 45(8): 1869-1877.

[13] Edge, D., Trinh, H., Cheng, N., *et al.* (2024) From Local to Global: A Graph RAG Approach to Query-Focused Summarization. arXiv: 2404.16130. https://arxiv.org/abs/2404.16130

[14] 李倩, 黄斌. 多智能体协作框架下图检索增强生成式问答系统构建研究——以图书馆领域为例[J]. 图书情报工作, 2026, 70(16): 181-199.