毕业设计（论文）开题报告

1. 选题的背景与意义

随着企业、学校和组织积累了大量的内部知识（如文档、指南、操作手册、技术资料等），如何有效地获取这些信息成为一个挑战。传统的检索方式难以满足用户快速获取信息的需求，问答系统可以提供更直观、便捷的方式来获取信息。

近年来人工智能技术快速发展，基于知识库的智能问答系统成为了信息检索和知识管理的重要工具。然而，目前多数知识库问答系统依赖网络连接或只能支持特定格式的文件，难以满足用户在断网环境下快速、准确检索本地知识的需求。

QAnything 是一种支持多格式文件的本地知识库问答系统，用户只需将本地的 PDF、Word、图片、电子表格等文件导入系统，即可获得准确、快速的问答体验，无需联网。通过本课题，基于 QAnything 的架构实现一个本地知识库问答系统，不仅可以帮助用户更高效地管理和利用本地数据，还能探索多模态数据处理与知识提取的关键技术，为知识库问答系统的多样化应用提供借鉴和支持。

本课题的意义在于为用户提供一个稳定、高效的知识管理工具，支持多种文件格式的导入和解析，进一步提升本地数据的利用效率，满足企业、学校和个人对离线知识问答的实际需求。

1. 课题研究的基本内容与拟解决的主要问题

本课题旨在开发一个基于 Qanything 实现的本地知识库问答系统。该系统主要解决以下几个问题：

1. 文件管理与知识库构建：如何将本地文档（如 PDF、Word、Excel 等）导入到系统中，并解析其中的关键信息，形成知识库。
2. 智能问答功能：如何利用自然语言处理技术，使用户能够通过自然语言提问并从本地知识库中准确获取答案。
3. 系统性能优化：如何设计一个高效的检索算法，确保在处理大规模文档时能够快速响应用户查询。

具体研究内容包括：

1. 文件管理模块：支持用户上传多种格式的文件（PDF、Word、Excel、图片等），并实现文件的预览与解析。
2. 知识库构建模块：自动解析文件内容，提取关键知识点，构建可供查询的本地知识库。
3. 智能问答模块：用户输入自然语言问题，系统能够从知识库中快速检索相关答案，并提供答案出处和可靠性标注。
4. 系统界面与用户体验：设计简洁直观的用户界面，提供友好的操作体验和实时反馈，确保用户能够高效地进行知识查询。

3、课题研究方法、设计方案或论文撰写提纲

研究方法：

* 需求分析与调研：首先进行市场调研，分析现有知识库问答系统的优缺点，明确用户需求，确定系统功能和技术要求。
* 技术框架选择：本项目基于 Qanything 构建本地知识库问答系统，采用本地化存储与数据解析的结合，同时结合自然语言处理（NLP）技术实现智能问答。
* 开发与实现：系统开发将采用前后端分离的架构，前端使用 React、Ant Design 等技术，后端基于 Node.js 或 Python 实现知识库管理与问答功能。
* 性能测试与优化：对系统的问答效率、查询准确性和响应速度进行测试，并根据反馈进行优化。

设计方案：

* 系统架构：系统采用典型的 B/S（浏览器/服务器）架构，前端通过 React 构建，后端使用 Node.js 或 Python 开发，利用 Qanything 提供的 API 实现对话和知识库管理。
* 数据库设计：使用本地数据库（如 SQLite 或 MySQL）存储知识库内容和用户信息，确保数据安全性和查询效率。
* 自然语言处理：集成 NLP 库（如 spaCy、NLTK、Transformers 等），实现对自然语言问题的理解与匹配。

论文撰写提纲：

1. 引言：介绍课题背景、研究目的和意义。
2. 文献综述：分析当前国内外关于知识库问答系统的研究现状及应用。
3. 系统需求分析与设计：分析系统功能需求，设计系统架构与模块。
4. 技术方案与实现：详细描述技术选型、系统实现和功能模块设计。
5. 实验与测试：进行系统的功能测试、性能测试及用户体验评估。
6. 总结与展望：总结研究成果，并对未来可能的改进和扩展进行展望。

4、总体安排与进度

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 日 期 | 教学周 | 设计（论文）各阶段名称或工作内容 |
| 1 | 2025.02.10-2025.02.16 | 1 | 前期调研，需求分析，市场调研，功能规划 |
| 2 | 2025.02.17-2025.02.23 | 2 | 方案设计，技术选型，架构设计 |
| 3 | 2025.02.24-2025.03.02 | 3 | 系统概要设计，确定开发框架与模块划分 |
| 4 | 2025.03.03-2025.03.30 | 4-7 | 系统前端开发，知识库管理模块实现 |
| 5 | 2025.03.31-2025.04.20 | 8-10 | 系统后端开发，智能问答模块开发 |
| 6 | 2025.04.21-2025.04.27 | 11 | 系统测试，修复Bug，撰写使用手册 |
| 7 | 2025.04.28-2025.05.18 | 12-14 | 撰写并完成毕业论文，完善系统文档 |
| 8 | 2025.05.19-2025.05.25 | 15 | 整理相关材料，准备毕业答辩 |

5、课题研究准备情况（主要参考文献、资料和参考书，完成任务所具备的条件等）

（1）在校期间，具备扎实的软件开发基础，熟悉 Web 开发、前端框架、后端开发及数据库管理。

（2）具备一定的项目开发经验，能够独立完成系统设计与实现，理解人工智能、NLP 及知识管理系统的基本原理。

（3）已掌握 Qanything 等相关工具，能够熟练应用其 API 进行系统集成。

主要参考文献：

[1] 梅忆寒,王琳琳,王鹏飞,等.基于多模态与检索增强生成的数据库知识问答系统[J].计算机教育,2024,(12).

[2] 杜恒峰.DeepSeek-R1惊艳全球展示中国AI发展巨大潜力[N].每日经济新闻,2025-01-27(001).

[3] 穆肃,陈孝然,周德青.生成式人工智能赋能教学设计分析：需求、方法和发展[J].开放教育研究,2025,31(01).

[4] 刘永东, 王文涛, 胡鹏. 本地化知识库问答系统研究与实现[J]. 软件学报, 2022, 33(6): 1428-1439.

[5] 姜嘉伟.基于Langchain-LLMs框架的智能问答系统的设计与实现[D].延边大学,2024.

[6] 任海玉,刘建平,王健,等.基于大语言模型的智能问答系统研究综述[J/OL].计算机工程与应用,1-24[2025-02-06].

[7] Radeva I ,Popchev I ,Doukovska L , et al.Web Application for Retrieval-Augmented Generation: Implementation and Testing[J].Electronics,2024,13(7).

[8] Mansurova A ,Mansurova A ,Nugumanova A .QA-RAG: Exploring LLM Reliance on External Knowledge[J].Big Data and Cognitive Computing,2024,8(9):115-115.