

研究生专业实践报告

基于大语言模型的智能信息处理与研究辅助系统实践

一、实践背景与实践目的

随着大语言模型（Large Language Models, LLMs）在自然语言理解与生成任务中的快速发展，其在科研辅助、知识管理与智能决策等领域展现出广阔应用前景。然而，在真实工程与科研环境中，大模型往往需要与外部系统、知识库及业务流程进行深度结合，才能发挥实际价值。

在研究生阶段的学习与科研过程中，传统的信息检索与资料整理方式已难以满足复杂研究任务对效率与系统性的要求。因此，本次专业实践围绕**“大模型在科研与工程场景中的系统化应用”展开，结合 LangChain 编排思想、检索增强生成（RAG）、RPA 自动化与前端交互系统，探索如何构建一个可控、可扩展、贴近真实使用场景的智能信息处理与研究辅助系统**。

本次实践的主要目标包括：

1. 探索大语言模型在复杂研究任务中的工程化应用方式；
2. 理解并实践 RAG、Agent、流程编排等关键技术；
3. 提升系统集成、模块化设计与自动化能力；
4. 将理论知识转化为可运行、可迭代的工程系统。

二、实践内容与主要任务

2.1 大模型驱动的信息处理与研究辅助实践

在本次实践中，重点围绕大语言模型在研究辅助场景中的应用展开，系统性设计并实现了一个多阶段信息处理流程。该流程并非简单的问答系统，而是将研究任务拆解为多个阶段，包括需求解析、结构生成、资料检索与内容整合等。

在实践过程中，重点关注以下问题：

- 如何将自然语言研究需求转化为结构化任务；

- 如何限制大模型生成范围，避免无关扩展；
- 如何在生成内容中引入外部证据，提高可靠性。

通过该实践，加深了对大模型“能力边界”与“约束机制”的理解。

2.2 基于 RAG 的知识检索与管理实践

在知识获取层面，本实践引入了检索增强生成（RAG）机制，将外部文档、研究资料与大模型生成过程相结合。主要实践内容包括：

- 文档切分与向量化处理；
- 向量检索与关键词检索的组合使用；
- 检索结果作为上下文证据输入大模型。

在实践中发现，RAG 并不能完全替代研究过程本身，其核心价值在于为生成内容提供可追溯的信息来源。因此，在系统设计中，将 RAG 定位为“支撑模块”，而非系统核心逻辑。

2.3 LangChain 编排思想的实践应用

在系统实现过程中，参考并实践了 LangChain 的核心设计思想，即：

- 将复杂任务拆分为多个功能明确的组件；
- 使用链式或图式结构组织模型调用流程；
- 通过统一接口管理 Prompt、模型与工具。

虽然未完全依赖 LangChain 框架本身，但在实践中吸收了其**模块化、可组合、可扩展**的设计理念，并将其应用于研究任务拆解与多阶段执行流程中。这一过程显著提升了系统的可维护性与可扩展性。

2.4 RPA 与自动化流程实践

除大模型相关内容外，本次实践还结合了 RPA（机器人流程自动化）技术，用于实现数据处理与系统操作的自动化。相关实践包括：

- 基于 Python 的脚本自动化；
- Excel 文件的批量读取、计算与汇总；

- 多工具、多系统之间的流程衔接。

通过 RPA 技术，将原本需要人工重复操作的流程自动化，不仅提升了效率，也为后续将大模型能力嵌入实际业务流程提供了基础。

2.5 前端与辅助工具实践（侧边栏项目）

在系统交互层面，实践中实现了基于 HTML、CSS 与 JavaScript 的轻量级前端界面，并探索了侧边栏式工具的设计方式，用于展示系统状态、研究进度与中间结果。

该部分实践主要关注：

- 前后端接口设计；
- 流式输出与实时反馈；
- 用户交互体验的基本优化。

通过该实践，加深了对系统完整性的理解，使大模型应用从“接口调用”走向“可使用系统”。

三、实践过程中的关键问题与解决方法

在实践过程中，主要遇到以下问题：

1. **大模型输出不稳定、结构易发散**
→ 通过引入结构化 Prompt 与阶段划分进行约束；
2. **检索结果与生成内容关联度不足**
→ 将检索结果作为证据上下文而非直接拼接文本；
3. **多模块系统调试难度较大**
→ 采用模块化设计与清晰的状态管理机制；
4. **人工操作流程重复、效率较低**
→ 引入 RPA 技术实现自动化处理。

这些问题的解决过程显著提升了实践的工程深度。

四、实践成果与能力提升

通过本次专业实践，取得了以下成果：

- 构建了多个基于大模型的信息处理与研究辅助原型系统；
- 掌握了 RAG、LangChain 编排思想与 RPA 自动化技术；
- 提升了系统设计、模块拆分与工程实现能力；
- 对大模型在真实科研与工程场景中的应用有了更理性的认识。

在能力层面，实践显著提升了本人在以下方面的综合素养：

- 复杂系统分析与建模能力；
 - 跨技术栈集成与调试能力；
 - 将理论方法转化为工程实现的能力。
-

五、实践总结与反思

本次研究生专业实践不仅是一次技术实现过程，更是一次对大模型应用模式的系统性探索。实践表明，大语言模型并非“万能工具”，其价值需要通过合理的流程设计、技术组合与工程约束才能充分发挥。

未来，在进一步的科研与工程工作中，将持续探索大模型与传统算法、自动化系统及领域知识的深度融合，为科研与实际应用提供更加可靠的智能支持。