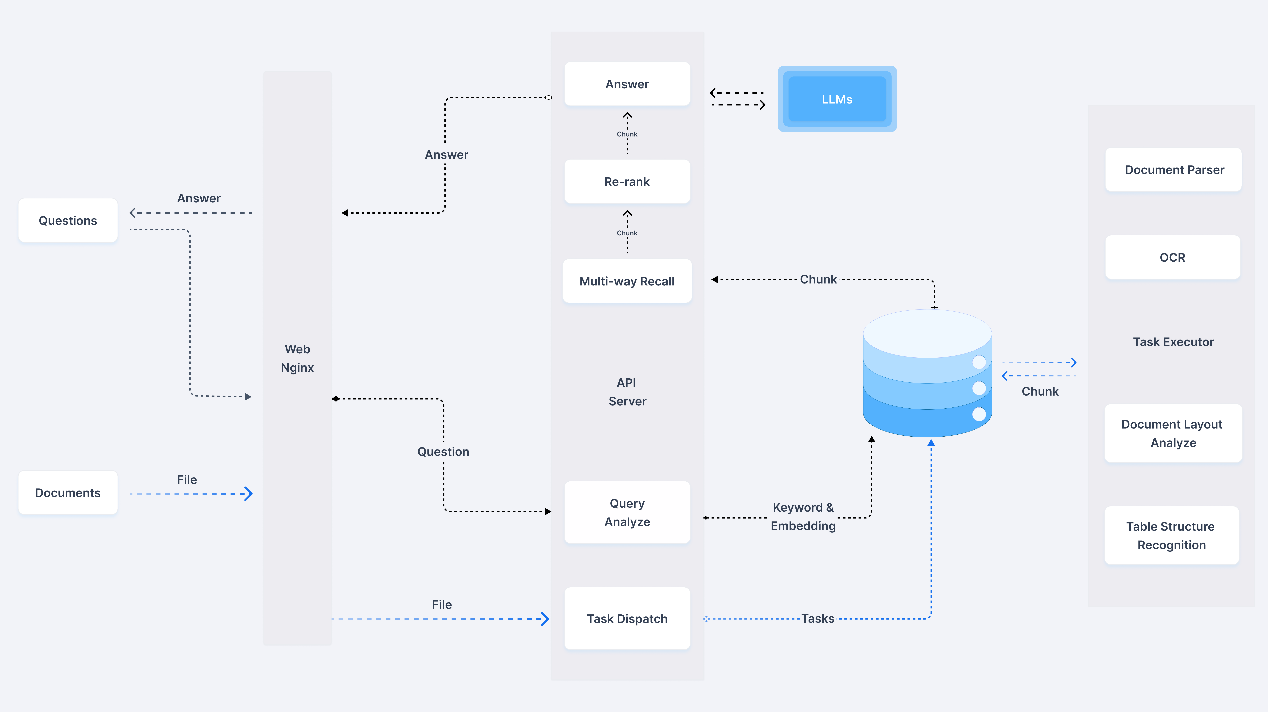
**项目报告PT.3**

**RAGFlow：开源RAG工作流框架**

RAGFlow 是一个开源的 RAG (Retrieval-Augmented Generation，检索增强生成) 工作流框架。它专门为构建和管理基于自身知识库的问答系统而设计。

RAGFlow系统架构

对于初步构建知识库的需求，RAGFlow 的优势主要体现在以下几个方面：

1. **端到端解决方案 (End-to-End Solution):** RAGFlow 提供了一个相对完整的流程，涵盖了从原始文档上传、处理（如文档解析、分块）、向量化、构建向量索引、到最终结合大模型进行检索和生成答案的全过程。不需要自己从零开始搭建整个 RAG Pipeline，这大大降低了技术门槛和开发时间。
2. **简化知识库构建过程:** 它提供了用户界面 (UI)，让我们可以方便地上传各种格式的文档（如 PDF, DOC, TXT, Markdown, 网页等），系统会自动进行处理。这比手动编写代码来解析、分块和向量化大量文档要高效得多。
3. **支持多种数据源和格式:** RAGFlow 通常设计来处理多种非结构化数据格式，这对于整合来自不同渠道的知识非常有帮助。
4. **向量数据库集成:** RAG 的核心是利用向量数据库进行高效检索。RAGFlow 框架内置或易于集成主流的向量数据库，帮你管理文档块的向量表示和索引。
5. **灵活的检索策略:** 可能支持不同的检索方法和参数配置，让你能够根据知识库的特点优化检索效果。
6. **易于与大模型集成:** RAGFlow 的输出（检索到的相关文档块）会作为上下文输入给大语言模型。它设计了与不同大模型接口对接的能力，无论是本地部署的模型还是云端 API。
7. **可视化管理和测试:** 通过其用户界面，可以直观地管理上传的文档、查看处理状态，并且可以方便地进行测试，输入问题，看看模型是否能根据知识库给出正确的答案，便于调试和优化。
8. **专注于 RAG 核心功能:** 它专注于解决“让大模型利用外部知识”这个问题，而不是成为一个通用的 LLM 平台。这使得它在 RAG 领域更加专业和高效。
9. **开源:** 开源项目通常意味着免费使用、社区支持、可定制性强以及潜在的快速迭代和改进。

总的来说，对于想要快速、便捷地构建一个基于自身文档的智能问答或知识检索系统的人来说，RAGFlow 提供了一个开箱即用的框架和工具集，大大降低了从概念到落地的难度。

对于本课程项目的需求，我们可以先通过微调模型，得到微调后的模型文件，再在RAGFlow中使用微调后的模型。需要注意的是，微调大模型需要准备高质量的数据集并采用适合的微调工具和框架如（LoRA等）进行训练。