

API层: Flask Web Server

RESTful APIS

API Token

Websocket

登陆、认证

核心业务层

RAG引擎

向量搜索

混合搜索

重排序(Reranker)

文档理解(DeepDoc)

PDF解析器

Docx解析器

OCR识别

Agent

画布

工作流引擎

MCP

基础设施层

LLM集成

OpenAI

Embeding(嵌入)模型

重排序模型(Reranker)

执行器 Executor

执行器

异步处理

队列管理

数据层

mysql

minIO对象存储

Redis(Redis Stream)

elasticsearch/infinity向量存储