

羊鸣天

AI 应用开发 / LLM 应用研发

电话 13678386848 | 邮箱 yangmingtian3777@gmail.com | GitHub github.com/Yorushikamimimi | 作品集 yoruming.cn
技术关键词: RAG Engineering | Langfuse | pgvector | FastAPI | Spring Boot | 文档智能

教育背景

重庆邮电大学 | 计算机科学与技术 (本科)

2023.09 - 2027.06

英语六级 536 分 | 专业排名前 30% | 青春重邮新媒体中心 (校团委) 主任

实习经历

重庆市移动互联网数据应用技术工程中心 (邮莱特) | AI 应用开发实习生

2026.04 - 2026.07

智能标书锚点解析系统 | 文档智能 / LLM 信息抽取

- 面向政府采购招投标文件, 负责 AI 解析链路、锚点抽取、评审证据定位、页面验证与交付收口; 支持 PDF / DOC / DOCX 解析, 覆盖资格审查、符合性审查、评分办法、废标条款等评审锚点。
- 设计 LLM-first 信息抽取 pipeline, 结合 OCR、PageIndex、规则 fallback 与多模型调用处理多格式差异; 围绕真实样本定位 OCR 识别、证据页映射和 LLM 输出稳定性问题, 并沉淀回归检查。
- 建立覆盖 22 份历史文档、2,747 条历史锚点记录的回归基线; 审计基线构建与样本回填逻辑, 识别历史匹配率口径偏差, 区分回归稳定性与真实准确率, 并沉淀错误归因和专家复核边界。

中国移动通信集团四川有限公司 (德阳经开区分公司) | 网络部实习生

2026.02

- 参与 4G / 5G 盲区路测、干扰排查及 PTN / SPN 传输网和机房环境跟进, 了解一线排障与运维协同流程。

项目经历

My Knowledge Base | RAG Engineering / Java + Python

- 基于 Spring Boot 3.3.4 / Java 17、FastAPI、React/Vite、PostgreSQL/pgvector 与 Ollama 构建本地 RAG 知识库; Java 负责权限、文件元数据与异步入库任务, FastAPI 负责解析、Embedding、向量检索与生成, 完成上传、入库、问答及 Sources 展示闭环。
- 接入 Langfuse, 为 RAG Query 的 Embedding、Retrieval、Generation 建立分层 Trace, 记录模型、Similarity Score、Token Usage 与 Latency; 验证 tracing 不可用时后台失败不阻断主业务返回。
- 基于 12 篇真实技术笔记构建 30 条 Gold V1.1 检索用例 (22 positive / 8 negative), 实现 production-faithful Retrieval Benchmark; 通过 controlled experiment 将 chunk 450/80 调整为 200/40, 解决 embedding context overflow, 实现 12/12 文档完整入库。
- 在 held-out chunk test 上取得 Section MRR 0.70、Section Hit@5 1.00、Wrong-Context 0; 对 min_score 做 DEV / Fresh Holdout 验证, 发现候选阈值 0.625 不泛化, 未上线并保持 production 0.35。

门诊患者-处方-库存管理平台 | Java 后端 / 全栈

- 负责患者管理、处方发药、药房库存、收费结算等核心模块, 打通“患者-就诊-库存-结算”业务链路, 并完成 Vue 2 + Element UI 前后端联调。
- 采用 PostgreSQL JSONB + GIN、UPSERT、Redis 分布式锁与事务回滚处理动态病历检索和关键流程一致性, 并完成 JAR + Nginx + systemd 私有化部署。

专业技能

- AI 应用工程: 具备 LLM API 接入、Prompt 调试、OCR / 文档解析、结构化信息抽取、规则 fallback、多模型调用与输出稳定性治理实践。
- RAG 与评测: 具备 pgvector、Embedding / Retrieval / Generation、Langfuse tracing、Retrieval Gold Set、Hit@K / MRR、Holdout 与 controlled experiment 实践。
- 服务端与数据: 熟悉 Java、Spring Boot 3.x、Spring MVC、MyBatis、Python / FastAPI; 熟悉 MySQL / PostgreSQL、Redis、索引、事务与并发控制。
- 工程化: 了解 React/Vite、Vue、TypeScript, 可完成基础页面与前后端联调; 熟悉 Linux、Docker、Nginx / systemd 部署与基础排障。