

李昱辰

27 届·意向岗位: Python Agent / LLM 应用开发实习

可尽快到岗·每周可实习 5 天·可连续实习 6 个月

邮箱: edllyc041202@gmail.com·GitHub: github.com/EDLlyc

教育背景

安徽大学 - 数字媒体技术专业 - 本科

2023.09 - 2027.06

- **科研与竞赛成果:** 以学生一作身份向 *IEEE TMI* 在投算法论文 1 篇 (外审), 提出基于拓扑学先验的连通性损失函数; 获美赛二等奖、数学建模省赛三等奖、三等学业奖学金。
- **外语能力:** 雅思 (IELTS) 7.0 分, 具备英文技术文档、学术论文阅读与复现能力。
- **成绩情况:** 均分 87.6/100, 系统修读数据结构、操作系统、计算机网络、数据库等计算机核心课程。

专业技能

- **Python Agent 工程:** 熟悉 FastAPI、AsyncIO、HTTPX、SQLAlchemy、LangGraph 状态图、Tool Calling、MCP、Plan-and-Execute/Replan、流式交互及工程化测试。
- **LLM / RAG / 多模态:** 具备 pgvector/Milvus/OpenSearch 混合检索、RRF/Reranker、父子切片、Embedding、多模态检索及 DeepEval/RAGAS 离线评测实践。
- **后端与基础设施:** 掌握 Java/Spring Boot、MySQL、Redis/Redisson/Lua、RocketMQ、Sentinel、Docker、Git, 具备并发控制、缓存治理和事务一致性实践。

实习经历

赛先文化 - AI Agent 应用开发实习生

2026.07 - 2026.08

项目技术: Python, FastAPI, LangGraph, Tool Calling, MCP, PostgreSQL/pgvector, 智谱 GLM, Qwen3-VL-Embedding, RAG, Docker

业务方向: 参与教育行业新媒体数字员工建设, 围绕新闻采集、事实治理、智能选题、品牌文案、配图审核与企业微信交付, 搭建可追溯、可评测、可恢复的内容生产链路。

主要工作:

1. **Agent 编排与 MCP 工具封装:** 基于 LangGraph 构建状态化 Agent 与有界工具循环, 将知识检索、事件详情、品牌 RAG 和文案校验封装为统一 Typed Tool Registry, 并通过 MCP Server 向 Agent 客户端提供标准化调用; 结合调用预算、严格 Schema 与确定性回退约束执行。
2. **智能选题与品牌 RAG:** 建立新闻快照、事实标签、事件版本和引用绑定, 设计“硬规则粗筛 + LLM 多维评分排序 + 确定性终审”; 针对 PPT/PDF 与访谈资料实现父子切片与 pgvector 混合检索, 脱敏离线评测 Recall@5 达 95%, nDCG@5 达 92.86%。
3. **多模态品牌素材检索:** 接入 Qwen3-VL-Embedding, 将文案与已审批品牌图片映射至统一 2048 维空间; 基于 pgvector 实现跨模态检索, 以“身份/角色/审批/完整性硬过滤 + 语义排序 + 规则回退”完成素材索引和自动配图。
4. **可靠性与工程质量:** 基于 FastAPI、AsyncIO、SQLAlchemy 实现异步任务、租约、幂等和故障恢复; 建设离线评测、类型检查、契约测试与迁移回退机制, 后端累计 1200+ 自动化测试, 覆盖率约 82%。

项目经历

12306 智能数字员工 (Python Agent / RAG 智能体研发)

2026.03 - 2026.06

项目技术: Python, FastAPI, LangGraph, LangChain, MCP, Redis, Milvus, OpenSearch, 智谱 GLM, DeepEval, RAGAS

项目描述: 面向 12306 规章问答与购退改查场景, 将智能体编排层重构为 Python 服务; Java 票务系统作为高并发交易工具内核, 由 Python Agent 负责意图识别、规划执行、RAG 检索、记忆管理与流式交互。

项目职责:

1. **LangGraph 编排与路由:** 基于 LangGraph 构建状态图主链路, 统一编排语义缓存、Query Rewrite、意图路由、上下文桥接、动作规划、工具执行与最终响应; 设计 ACTION / TICKET / RAG / CHITCHAT 多专家意图打分, 结合最高分、分差与置信度阈值完成分层路由决策。
2. **工具编排与安全执行:** 以标准 Tool Schema 封装查票、订票、订单与退票能力, 并接入天气 MCP 完成跨工具建议; 对购退票引入 Plan-and-Execute、显式确认、Replan 与副作用去重, 避免误触发和重复执行。
3. **混合检索与 Agent Eval:** 实现多格式规章文档解析、条款级切分和 hash 去重, 结合 OpenSearch BM25、Milvus、RRF/Reranker、Redis 语义缓存与长短期记忆; 构建 golden cases 并接入 DeepEval/RAGAS, 历史 context precision 0.881、context recall 0.935。

12306 高并发售票核心网关 (交易内核/后端开发)

2026.03 - 2026.06

项目技术: Spring Boot, Redis, Redisson, Caffeine, RBloomFilter, RocketMQ, Sentinel, MyBatis-Plus

项目描述: 作为 Python Agent 的票务交易工具内核, 面向 12306 高并发售票场景, 围绕查票、订票、退款等核心链路完成缓存治理、库存预扣、事务一致性与补偿闭环设计。

项目职责:

1. **缓存与高并发治理:** 基于 Caffeine + Redis + RBloomFilter 构建多级缓存, 结合 Redisson 分布式锁与基于 Completable-Future 的 singleflight 回源合并治理热点 Key; 在查票热路径压测中, 800 并发下达 5409.54 QPS、0% HTTP 错误率, P95 357 ms。
2. **库存与事务一致性:** 基于 Redis Lua 实现库存原子预扣, 设计 Segmented Stock 分桶模型分散热点; 通过 RocketMQ 事务消息串联 Redis 预扣、订单落库与 MySQL 扣减, 在 200 用户订票压测中, 订票前门受理吞吐为 510.94 requests/s、0% 错误率, 未观察到超卖。
3. **状态机与补偿闭环:** 设计支付、取消、退款状态机及业务幂等号, 结合延迟消息、状态条件更新和 Lua 幂等回补处理超时关单、MQ 重试与重复退票; 使用 Sentinel 完成限流和热点保护。