

欣禹行

求职方向：AI应用开发工程师

☎ 15061235115 | ✉ huali6641@gmail.com | 🌐 takagibit18 | 🏠 seanhomepage.top

教育经历

中央民族大学 (985)

2023/09 - 2027/06

计算机科学与技术 · 本科 · IELTS 7.0

实习经历

卓望信息技术有限公司（派驻国家数据集团） | AI 应用开发实习生

2026/06 - 2026/08

- **本体建模接入**：参与智能体侧对已有本体平台的能力接入，围绕政策规则建模任务补充 Session / Retry / Checkpoint，支持失败重试与断点恢复；并参与配置化 Skill 与动态 Tool Loading 的接入验证，使不同领域工具能够按配置加载。
- **政策检索优化**：参与已有 RAG 的检索侧优化，针对政策咨询中的口语化表达补充中文 Query Rewrite 与关键词扩展，并结合地区、生效时间、政策类型和版本信息过滤候选材料；配合离线检索评测分析 Bad Case，中文政策检索基线 Recall@5 达 0.95。
- **评测与优化**：基于 RAGAS 整理并运行 159 条自动化评测样本，围绕回答相关性、证据完整性及异常场景分析 Bad Case，并结合运行日志定位检索、路由与生成环节问题，为 Prompt 和问答链路调整提供依据。

项目经历

MergeWarden (AI Code Review Agent)

2026/03 - 至今

面向 GitHub PR 的证据驱动智能审查系统

面向 GitHub PR 的智能代码审查 Agent，以 Diff 为入口自主检索代码并补充证据，重点识别跨文件行为回归、边界条件与测试缺口，作为传统 CI 的补充。

- **Agent 核心架构**：设计可持续探索的 Review Loop，由模型根据当前证据自主决定继续检索、验证或提交结果；提供代码读取、搜索、符号定位等只读工具，并通过资源预算、沙箱隔离与 advisory-only 输出约束执行边界。
- **Graph 上下文检索**：针对大型 PR 中跨文件、多跳搜索成本高、同源问题证据分散的问题，构建代码关系 Graph，沿调用、引用及状态读写关系聚合关联 Context，辅助 Reviewer 完成跨文件推理与根因归并；在真实 PR 黄金集 A/B 评测中，相比 Agent Search 平均 Token 降低 27%、工具探索轮次降低 59%、审查延迟降低 54%。
- **上下文管理**：构建面向多轮 Agent 的分层 Context 管理机制，通过滑动窗口保留近期交互、压缩早期历史信息，并为最终结构化提交预留独立 Token 预算，在保留关键证据的同时控制长任务中的上下文膨胀。
- **索引持久化**：设计代码关系 Graph 的 SQLite 持久化索引与 Warm 复用机制，避免重复执行全量静态分析；将检索层与 Reviewer 推理解耦，使 Graph 与 Agent Search 等策略可独立替换和优化，不影响主审查流程。
- **评测与可观测性**：搭建真实 PR 驱动的自动化评测与运行追踪体系，统一记录模型决策、工具调用、Token、延迟、终止原因及证据来源，支持 A/B 对比、错误样本定位与回归验证。

专业技能

- **编程与后端**：熟练使用 Python，熟悉 TypeScript；掌握 FastAPI、Pydantic、asyncio 等后端开发技术，具备 RESTful API、异步任务与流式服务开发经验。
- **LLM / Agent 应用**：熟悉大模型 API、Prompt Engineering、Tool Calling 与 Structured Output，理解 Agent Loop、工具编排、上下文管理及异常恢复等核心机制；熟悉 LangGraph 等 Agent 工作流编排框架。
- **RAG / 评测**：熟悉 BM25、Query Rewrite、元数据过滤等检索优化方法；具备 RAGAS、Golden Set、A/B Test 与 Bad Case 分析实践，了解 Recall@K、MRR 等检索评测指标。
- **工程化**：熟悉 Git、Docker、pytest、GitHub Actions、SQLite 等开发工具，具备自动化测试、CI/CD、日志 / Trace 及 Token、延迟等运行指标监控实践。