# 任彦舟 AI Agent 开发工程师

邮箱：2675528281@qq.com

## 教育背景

**电子科技大学** 软件工程 · 硕士研究生（在读） 2026 — 2029（预计）

**电子科技大学** 软件工程 · 本科 | GPA 3.93/4.0 2022 — 2026

2025 中国大学生计算机设计大赛国家一等奖 | CET-6

## 实习经历

**北京知道创宇信息技术股份有限公司（成都）** AI 研发实习生 · 多模态安全 2025.08 — 2026.01

- **模型服务化**：将多模态 Guard 封装为 HTTP 推理服务，统一 safe/unsafe 及 9 类风险标签输出，接入 DAG 检测节点；完善输入校验、健康检查及并发、OOM 异常处理。
- **模型微调与评测**：独立完成约 3.9 万条图文样本清洗、标签统一和 LoRA SFT，按误报、漏报及二级错分分析问题；在固定 5,082 条测试集上，主标签准确率 **70.01% → 87.66%**。
- **安全策略编排**：基于 DAG 并行执行关键词与轻量分类，按风险和置信度路由至语义、提示词注入检测；通过 YAML / Groovy 配置阈值、拦截与代答分支，支持策略热更新；补充工具调用中的命令注入与数据泄露扫描。

## 项目经历

**Roc · 科研 Agent 运行时与评测** 个人独立开发 2026.04 — 至今

GitHub: github.com/ryzqi/Roc

- **任务编排与恢复**：基于 LangGraph / deepagents 实现规划、工具执行与子智能体协作，使用 SQLite 保存检查点、事件和操作回执；通过幂等键与结果校验复用已完成操作，支持暂停、审批续跑和进程重启恢复。
- **上下文与记忆**：将长工具输出落盘，分阶段压缩历史并生成摘要，单独保留用户约束、待办和证据位置，支持按需回读；建立全局与项目两级记忆，处理去重、冲突更新和过期归档。
- **工具接入与授权**：整合文件、命令、检索、MCP 和 Skills，按任务与子智能体权限装配工具；将单次授权绑定任务、命令和工作目录，在主进程校验后执行，限制目录与网络访问。
- **回归与问答评测**：复用应用执行链路，建立 **94 项原生回归用例**，覆盖授权、工具副作用、取消与恢复；接入 QASPER 文本问答评测，完成 fixture 接线验证，真实模型质量基线尚未完成。

**CogniWeb · 多知识库问答与 Deep Research** 个人独立开发 2025.09 — 2026.05

GitHub: github.com/ryzqi/RAG-Research-Workbench

- **异步任务编排**：使用 FastAPI + Celery 解耦接口与文档入库、研究、导出任务，通过 SSE 推送执行进度。
- **Agentic RAG**：支持多知识库隔离与联合检索，结合父子分块、Milvus Dense + BM25 召回、全局 RRF 融合及可选 Rerank；基于 LangGraph 编排查询改写、HyDE 和问题拆解并行检索，回答复核不通过时改写查询并补检。
- **Deep Research**：基于 deepagents 按网页检索、论文检索、主张核验与章节写作分工协作；为证据绑定引用编号与原文位置，按会话保存计划、证据清单与报告，支持回溯和导出。
- **离线评测**：以 3 类知识库、68 份文档和 100 道人工构造题对比 BM25 + 向量基线，人工复核通过率 **54% → 79%**，引用命中率 **48% → 86%**；平均耗时 **8.4s → 16.0s**。

## 技术能力

| | |
|---|---|
| **编程语言** | Python、TypeScript |
| **后端与异步** | FastAPI、Celery |
| **Agent 与工具** | LangGraph、deepagents、MCP、Skills |
| **数据库** | PostgreSQL、Redis、SQLite、Milvus |