殷浚航

求职意向　AI Agent 开发工程师 / 大模型应用开发

13983878019 | yinjunhang@outlook.com

GitHub　github.com/mkynyd | 技术作品集 · 随附 2 页



教育背景

重庆大学　计算机学院 · 信息安全 · 本科　　预计 2027 年 6 月毕业

专业课程　计算机组成与结构（90） · 计算机网络安全（85） · 数字逻辑（86）

实践课程　硬件综合设计（98）：CPU Cache 设计。
软件综合设计（86）：微信小程序与 H5 双端前端设计、开发。

英语成绩　CET-4：559 分 | CET-6：516 分

项目经历

LumenLab　受控 Agent 学习与研究工作台　　2026.06 - 至今

支持资料检索、工具调用审批与审批后续跑，并以时间线追踪执行进度。

个人项目 · 独立负责需求、架构与交付，以文档和测试组织 AI Agent 辅助迭代

github.com/mkynyd/lumenlab | lab.mkynstudio.top | TypeScript / Next.js

- **执行控制：**以自建 Harness 统一 3 家模型供应商的工具调用；限制循环轮数、阻断重复调用与无进展循环，按 Token 预算摘要历史并保护近期对话。
- **资料检索：**文档分块后接入 Embedding / pgvector，与关键词检索经 RRF 融合排序；Redis 缓存查询向量与检索结果，以项目版本更新失效检索缓存。
- **权限与恢复：**校验参数、资源归属及 Skill 白名单，按风险审批；租约 Worker 保存 checkpoint 和工具结果，审批后续跑并复用已完成结果。
- **界面与交付：**以 SSE 时间线、工具状态与审批卡呈现执行过程，支持断线事件回放及多端适配；完成腾讯云部署，以回归测试、UI / API 验收和发布回滚保障交付。

arcthis　面向 Agent 的压缩文件访问工具　　2026.08 - 至今

让人和 Agent 无需先完整解压到磁盘，即可浏览目录、定位并读取目标条目。

个人开源项目 · 独立负责需求、架构与交付，以文档和测试组织 AI Agent 辅助迭代

github.com/mkynyd/arcthis | arcthis.mkynstudio.top | Rust / CLI / MCP

- **归档访问：**以 Rust 应用层统一 CLI / MCP 接口，支持 12 类格式；条目索引缓存复用目录查询，持久化索引按源文件指纹失效。
- **访问与写入安全：**目录授权、路径校验与解码 / 响应预算约束访问；写操作拆分 plan / execute，核对计划摘要后暂存、校验、提交，覆盖失败回滚，提交成功后才删除源文件。
- **协议与验证：**MCP 提供 9 个只读及 6 个授权后启用的写操作相关工具；通过 68 项 CLI / MCP / 公共 API 集成测试，配置 macOS / Linux CI 构建，提供中英双语文档。

专业技能

- **语言与 Web：**Rust、TypeScript、React / Next.js；响应式界面与复杂任务交互。
- **Agent 系统：**Harness、Tool Calling、MCP、Skill、多模型适配、上下文与状态管理。
- **检索与存储：**文档分块、Embedding / pgvector、RRF、Redis、PostgreSQL / Prisma。
- **工程交付：**Git、Linux、腾讯云部署、CI/CD、自动化测试、类型检查与 UI / API 验收。

arcthis

面向 Agent 的压缩文件访问工具

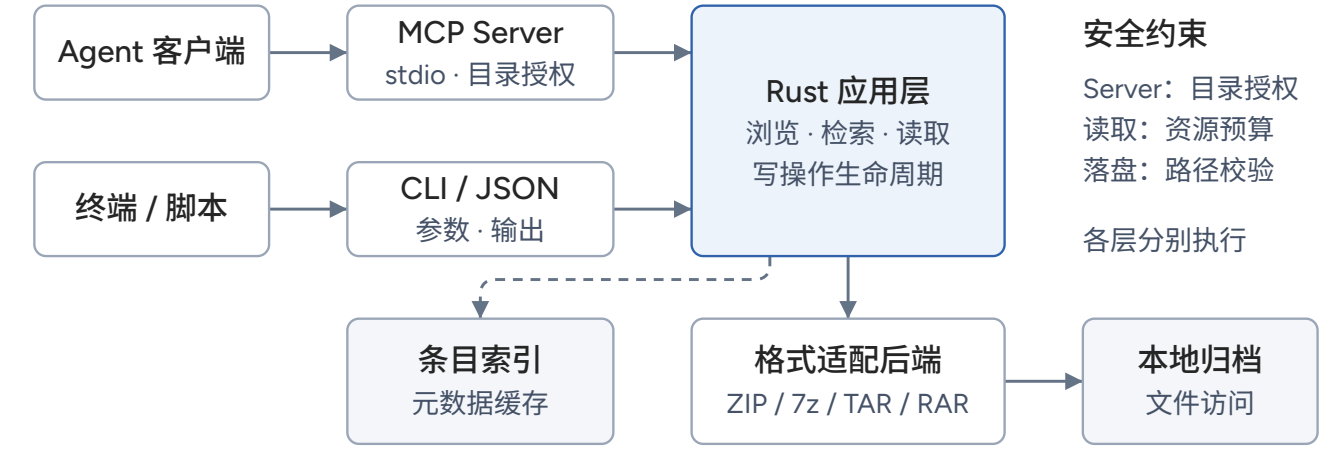
殷浚航 | AI Agent 开发 | 技术作品集

个人开源项目 · Rust / CLI / MCP · 2026.08 - 至今

项目网站

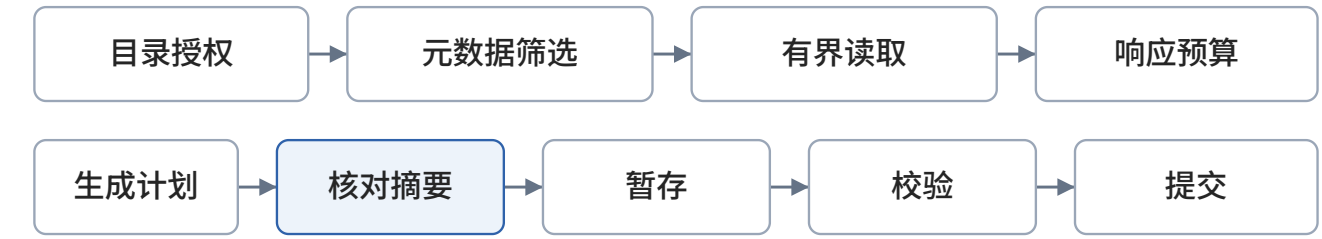
让 Agent 先浏览目录、定位条目，再按需读取；个人负责接口、安全策略与验证，借助 AI Agent 辅助迭代。

调用架构与职责边界



后端复用 zip / sevenz-rust2 / tar；RAR 经 compress-tools 调用 libarchive。元数据索引按源文件大小与修改时间失效，不缓存解码正文。

MCP 的读取与写入链路



实线：处理顺序。Server 暴露 9 个只读工具与 3 组 plan / execute；写工具需输出目录授权，执行前重算摘要，防止过期计划直接落盘。

格式支持与访问成本

格式（共 12 类）	访问方式	范围与限制
ZIP	按条目定位	正文读取支持 Stored / Deflate
7z	按块访问	solid 归档可能解码前序内容
RAR / RAR5	顺序访问	支持读取、提取；不创建 RAR
TAR / 单流（9 类）	顺序解码	TAR 及 GZ / BZ2 / XZ / ZST 变体； GZIP / BZIP2 / XZ / Zstandard 单流

设计取舍与工程验证

为避免归档条目越界，拒绝 ../ 等上级目录路径、绝对路径及驱动器 / UNC 路径，默认拒绝链接与特殊文件。为控制半成品和误删风险，暂存校验后提交，覆盖失败回滚，成功后才删除源文件。

LumenLab

受控 Agent 学习与研究工作台

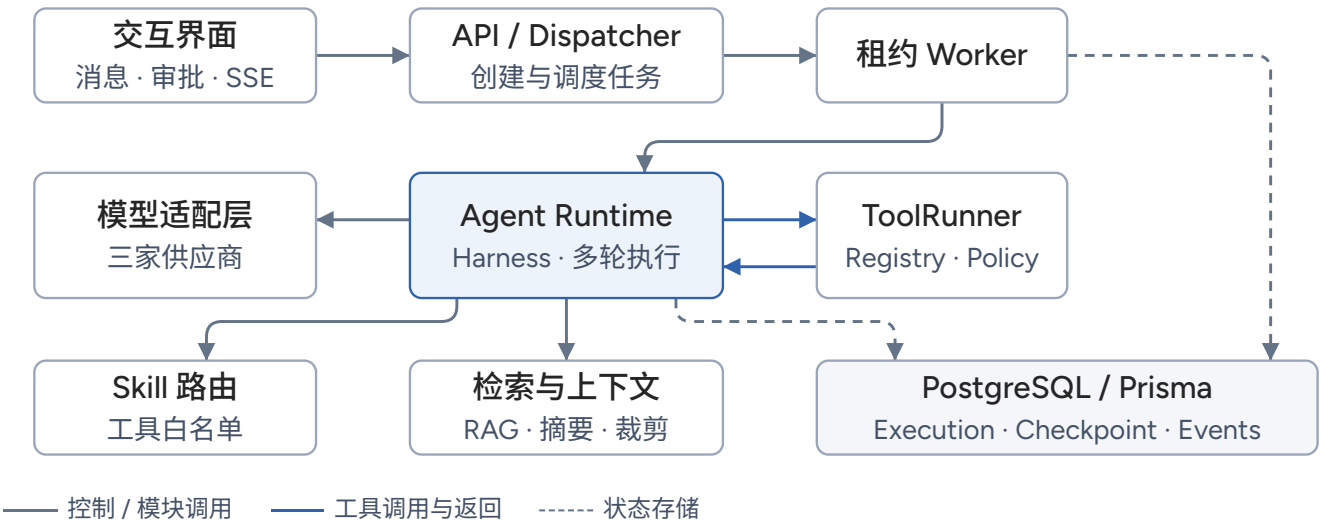
殷浚航 | AI Agent 开发 | 技术作品集

个人项目 · TypeScript / Next.js · 2026.06 - 至今

项目网站

面向个人资料学习与研究，独立设计检索、工具执行和进度交互，借助 AI Agent 辅助迭代。

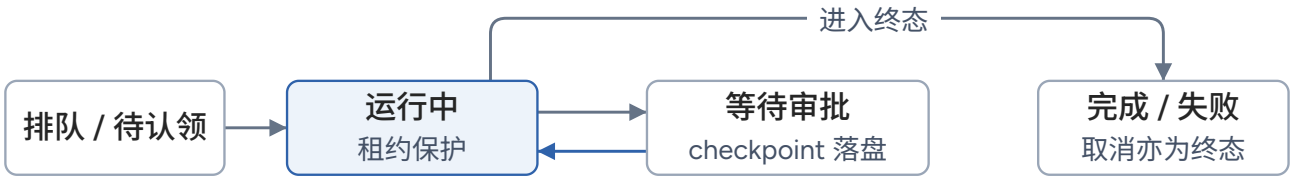
Agent Runtime 与持久化架构



检索、缓存与执行策略

RAG 检索	文档分块后生成 1024 维 Embedding，入库 pgvector；关键词与向量结果按 RRF 融合排序。向量不可用时降级关键词检索，保留来源。
缓存取舍	查询向量缓存 300 秒；检索结果缓存 60 秒，键包含项目、查询与文件范围。资料分块更新时递增项目版本，使旧检索结果失效。
上下文压缩	请求前按 Token 预算触发摘要，保留 system 消息与近期 12 条对话消息；摘要保留事实、约束和未完成事项，替换旧历史后重新检查预算。
循环停止	工具循环默认最多 8 轮，阻断重复调用；连续两轮未产生新信息时转入收尾。用户取消立即中断，等待审批时保存状态并暂停。

任务状态与恢复边界



灰线：状态迁移；蓝线：审批批准后续跑。ToolRunner 校验参数、资源归属与 Skill 白名单，高风险逐次审批；等待时保存 checkpoint 并释放租约，续跑复用已完成工具结果。

交互设计与验证交付

- **可追踪交互：**SSE 时间线、工具状态和审批卡适配多端；按用户归属过滤、按 Last-Event-ID 回放事件，断线恢复无需重跑工具。
- **工程验证：**腾讯云发布结合健康检查与回滚；本次压缩、检索、缓存、Embedding 与执行循环的 44 项专项测试通过，配合类型检查及 UI / API 验收。