# 居浪

求职意向：AI 应用 / Agent 工程师



男丨25岁丨硕士 17314833596 langju123456@gmail.com
github.com/langju123456

## 个人优势

1. 圣路易斯华盛顿大学信息系统管理硕士（2025.12）；AI软件公司创始人兼技术负责人，主导LLM应用与AI Agent产品从0到1的研发与交付。
2. 需求到方案：从业务场景出发拆解问题，判断哪些环节交给模型、哪些由规则与人工确认兜底，把需求转成Agent流程、数据边界与验收标准，并对最终交付负责。
3. Agent与RAG：独立实现有边界的证据检索Agent（查询规划→检索→覆盖检查→有依据草稿）与向量RAG问答；精通Prompt与结构化输出、混合检索与RRF排序、引用核验与上下文预算控制。
4. 生产级可靠性与评测：持久化任务、幂等、租约、审批与回执分离，覆盖超时、崩溃与重复提交；自建Agent评测平台，在3个尺寸模型上采集776条真实轨迹做失败归因，以测试、CI与独立审阅保障交付。

## 工作经历

**Lang Horizon Technologies llc** 创始人 · AI应用全栈工程师 2026.01-至今

自营AI软件项目研发与交付。创立并经营公司，主导SoloScale AI OS、ForgeFlow、EvidenceLoop、BuildLog、美业服务Copilot、X Post Lab六项自研产品从0到1的需求、研发与交付，面向AI应用研发与客户需求落地。

1. 研发与集成：用Codex、Claude Code推进AI辅助开发，本人负责架构取舍、关键执行路径和代码审阅；集成多家云端与本地模型、OAuth及外部API，统一结构化输出、超时与错误处理。
2. 需求与范围：与真实店主做需求访谈和原型测试，把业务场景拆成可验收的流程与数据边界；按证据成熟度决定各产品的范围和下一步，未完成项在交付材料中明确列出。
3. 评测与质量：用真实模型运行和失败归因反查Agent与评测本身的缺陷；每个产品配自动化测试、类型检查与CI，关键版本另设独立审阅，关键技术决策写入决策记录。
4. 交付与部署：按客户实际设备与使用条件选择部署方式，完成门店本地轻量化部署、macOS桌面打包与云端GPU实验环境配置，配套部署说明与排障记录。

**南京康尼机电股份有限公司** 海外技术销售工程师 2023.12-2024.05

参与2-3个海外轨道交通投标及技术项目，连接海外客户与研发、生产、销售团队。

1. 将客户规格、交付要求和问题清单转化为技术方案、需求澄清材料及执行事项，支持英文客户沟通，协调跨部门事项推进投标。
2. 结合CRM/ERP业务流程识别人工操作、信息断点和数据管理问题，输出流程分析及改进需求。

## 项目经历

**SoloScale AI OS｜LLM应用与工程证据工作台** 2026.01-至今

1. RAG与检索：对授权的Git项目、文档和运行产物重叠分块、增量同步，保留稳定分块ID、来源与内容哈希；使用SQLite FTS5/BM25、元数据匹配、RRF融合排序和邻近分块扩展，支持中英文证据检索与来源追溯。
2. Prompt与Agent：BoundedEvidenceAgent按查询规划→检索→覆盖检查→有依据草稿执行；版本化Prompt与Pydantic严格结构化输出，限制轮次、查询数与上下文字节预算，核对模型实际可见片段及引用ID/哈希，配事实缺口提示与人工确认。
3. 模型集成与并发：适配Ollama、OpenAI-compatible、DeepSeek及托管网关，统一结构化响应与超时/错误处理；幂等键、FOR UPDATE SKIP LOCKED事务领取与租约，双worker完成20/20合成模板任务、无重复完成，中断后转人工复核。
4. 交付：Swift/SwiftUI界面与PyInstaller打包的Python sidecar经loopback与HMAC启动校验连接，交付macOS 0.4.2(9)候选；720项测试通过、10项跳过，Ruff与严格mypy通过。

演示：soloscale-showcase.vercel.app | 源码：github.com/langju123456/soloscale-ai-os

**ForgeFlow｜面向企业的Agent受控执行与审计平台** 2026.10-至今

1. FastAPI/PostgreSQL持久化run、step、event、approval、receipt，采用lease/epoch与幂等键；验证提交前SIGKILL回滚、重启恢复与跨进程同键回放。
2. 基于MCP SDK设定sandbox工具边界，配合Semgrep、Docker与工件哈希；模型提议、服务端判定、审批与回执分离。Supervisor、PM两个角色接入本地Qwen3:8B，5个场景共20次调用，修复评测器后离线重评5/5。
3. 主导一轮系统性审阅与修复（审批绕过、回放锚点、审批版本化、提议者与审批者分离），932项隔离测试通过。

**EvidenceLoop｜Agent评测与后训练数据平台（开源）** 2026.10-至今

1. 8个工具组成多步骤任务，建成550道冻结评测集并配独立oracle；验证器从数值、单位、来源、状态、权限五个维度判定，并检查最终回复与系统状态是否一致，错误路径表把失败归到具体原因。
2. 在Qwen3-0.6B/1.7B/4B上采集776条真实轨迹：1.7B的格式错误集中在嵌套工具参数，拆平接口后单位/状态/来源通过率由38%/43%/17%升至78%/85%/65%；修正验证器放过编造数值的漏洞；在Kaggle T4上搭建Qwen3-1.7B LoRA训练与vLLM评测流水线。

源码：github.com/langju123456/evidenceloop

**BuildLog｜AI内容与工作流任务服务** 2026.01-至今

1. API与独立worker解耦，持久化任务、产物与运行记录；幂等键拒绝重复载荷冲突，以事务领取、有限重试和超时恢复处理重复提交与中断；配请求ID、结构化日志与健康/就绪探针。
2. 集成LinkedIn与X OAuth、人工批准和重复发布保护，各完成1次受控文本发布并保留HTTP 201回执。

**AI Research Assistant｜RAG 与本地模型** 2025.04-2025.07

1. 解析PDF/DOCX，按900字符分块、150字符重叠并保留来源与chunk编号；SentenceTransformers生成嵌入，Chroma持久化知识库，保留FAISS（IndexFlatIP）余弦检索方案。
2. 统一封装OpenAI gpt-4o-mini与Ollama本地Qwen，组合文档检索与按需Web Search，返回answer、sources、grounding_status、warnings结构；FastAPI、Streamlit与Docker交付，来源契约测试覆盖无效chunk与冲突内容。

**Ardent Mills｜供应链数字化转型校企Capstone（项目负责人）** 2025.09-2025.12

带领6人团队，经14场客户、业务与技术相关方会议，澄清供应链B2B集成、主数据与数据治理问题；比较EDI、API与混合集成方案，向客户管理层及技术团队汇报分阶段路线图。

## 教育经历

**圣路易斯华盛顿大学** 信息系统管理（MISM） 硕士 2024.08-2025.12

**佛罗里达大学** 信息系统（B.S.B.A.） 本科 2019.09-2023.05

## 专业技能

**大模型应用：**LLM应用开发、Prompt工程、Function Calling / Tool Calling、结构化输出（Pydantic / JSON Schema）、上下文工程

**Agent：**AI Agent设计、ReAct / Plan-Execute、多Agent协作、Workflow编排、LangChain / LangGraph / LlamaIndex / smolagents、MCP

**RAG：**文档解析与分块、Embedding、Chroma / FAISS、BM25+向量混合检索、RRF融合、引用核验

**模型与部署：**OpenAI / DeepSeek / Qwen / Llama、Ollama / vLLM、本地私有化部署、LoRA微调

**评测与质量：**Agent评测、失败归因、可观测、pytest / Ruff / mypy、GitHub Actions CI/CD

**工程：**Python、FastAPI、PostgreSQL / SQLite / SQL、Docker / Linux / Git、AI Coding（Claude Code / Codex）

## 证书与语言

Hugging Face：Agents Course 卓越证书（涵盖 smolagents / LlamaIndex / LangGraph） | MCP Course | Deep RL Course | Robotics Course

DeepLearning.AI / Coursera：Machine Learning Specialization | Machine Learning in Production

语言：英语精通（听说读写） | 西班牙语熟练