

孟祥龙

AI Agent 工程化 / 多智能体协作 / 后端架构

手机: 19879871236 邮箱: rayson951005@gmail.com 个人网站: raysonmeng.github.io

GitHub: github.com/raysonmeng 所在地: 深圳 工作经验: 9年+; 有 3-5 人团队管理、面试官、创业与产品设计经验

个人简介

9年+后端与平台工程经验, 前字节商业化审核质量方向负责人, 长期负责复杂业务流程、质量闭环、高可用服务和跨团队交付。近两段经历聚焦大模型与 Agent 工程化: 在 Klook 主导多语言内容中台从“机翻 + 外包校对”升级为智能工作流; 在 Picker 从 0 到 1 构建 POS、供应链、C 端社区与主动型 AI 店长 Agent。

持续开源多智能体协作与 Agent 工程化项目: AgentBridge 打通 Claude Code 与 Codex 会话内协作; HailMesh 抽象本地优先的 Agent 通信层; Agora 验证多模型讨论与匿名互评; Quilin Agent 从 0 到 1 实践通用 Agent 运行支撑体系。

核心能力

Agent 工程化: 运行循环、工具调用、模型接入、上下文组装、长期记忆、子 Agent 协作、人工确认、安全门禁、运行日志、问题定位、质量评测与回归。

大模型应用落地: 提示词设计、检索增强生成 (RAG)、向量检索、样例召回、结构化输出、富文本安全、模型路由、成本控制、质量评估。

后端与平台工程: Golang、Python、TypeScript、MySQL、PostgreSQL、Redis、ES、消息队列、Milvus、Docker、Linux、微服务、高并发、高可用、性能优化。

开源项目

多智能体协作与 Agent 工程化实践 | 2026 - NOW | Solo Maintainer

围绕“多 Agent 如何讨论、传话、组网、成为可治理节点”做连续实践: Agora 验证多模型讨论; AgentBridge 验证真实编码 Agent 会话协作; HailMesh 抽象通信与发现层; Quilin Agent 回到 Agent 节点本身, 实践通用运行支撑体系。

Quilin Agent | 通用 Agent 运行支撑体系

开源项目

- 从 0 到 1 构建通用 Agent 运行框架, 拆解大模型外围工程组件: 核心运行循环、流式模型调用、工具调度、会话状态、检查点、运行钩子、日志事件、子 Agent 生命周期和通信网络契约。
- 将工具调用、安全和可观测性作为一等边界: 实现工具路由、MCP 工具注册、沙箱策略、写操作审批、动作校验、外部上下文注入扫描、运行记录与审计事件。
- 接入上下文与记忆组件: 会话上下文组装、长期记忆桥接、时间上下文、检索增强记忆, 为多轮任务和跨会话延续打基础。

AgentBridge | Claude Code ↔ Codex 本地双 Agent 协作桥

GitHub 128 stars / 25 forks

- 打通 Claude Code 通信通道与 Codex 本地服务, 让两个独立编码 Agent 在同一开发会话内互相发送消息。
- 设计前台插件 + 持久后台服务 + Codex 代理架构, 支持断线重连、消息来源标记、防回环、长任务看门狗、会话回放、连接探活和失败消息兜底。
- 支持同机多组协作隔离; 截至 2026-06-09, npm 近一个月下载 598 次。

HailMesh | 本地优先的 Agent 通信与发现层

通信底座

- 抽象 Agent 能力卡片、消息信封、发现、寻址、发送、推送、请求响应、本地房间广播、在线状态、显式远端节点等核心能力。
- 采用 Go 后台服务 + TypeScript 适配器架构，将路由、身份、节点、连接、策略留在后台服务，将不同 Agent 运行环境隔离到适配器层。
- 明确不做任务编排框架，专注解决多 Agent 协作前必须存在的“能发现、能投递、能回复、能诊断”的通信问题。

Agora | 多模型讨论与匿名互评实验

前期实验产品

- 将复杂问题拆成独立分析、匿名互评、反驳、书记员总结四段流程，用结构化结果沉淀共识、分歧、建议和未决问题。
- 实现匿名身份映射、排除自评、交叉批评、轮间状态压缩、执行锁、观察者恢复和数据库状态管理，验证多方观点如何暴露分歧并收敛。

工作经历

啤客 Picker | 联合创始人 / 技术负责人

2025.04 - 2025.12

主动型 AI 店长 Agent + 精酿零售全链路平台

- 从 0 到 1 构建 C 端小程序、运营后台、门店 POS、老板手机端、供应链 ERP，并将主动型 AI 店长 Agent 作为 POS 手机端核心入口。
- 基于 LangGraph 搭建“主管 Agent + 数据分析 / 库存采购 / 菜单运营 / 营销活动子 Agent”的协作架构，支持意图识别、参数补全、多轮追问、任务路由和方案迭代。
- 组合 MySQL 经营事实、Milvus 相似门店 / 商品召回、图关系归因链路和 Mem0 分层记忆，将实时数据、历史偏好和门店画像装配成结构化上下文。
- 将 POS、库存、供应链、会员、活动等业务能力封装为可调用工具；采购、调价、发券等敏感操作加入权限校验、审计留痕和人工确认。
- 每日自动生成经营日报，并以“今日 / 未来一周建议”作为对话起点，推动店长完成补货、上新、活动和会员运营等动作。

Klook 客路旅行 | 高级研发工程师 / 项目负责人

2024.07 - 2025.02

多语言内容中台智能化重构

- 主导全球多语言内容中台从“机翻 + 外包校对”升级为大模型智能 workflows，覆盖模型路由、检索增强生成 (RAG)、富文本安全、质量评估和回归体系。
- 用 Milvus 召回金标准样例，通过多样性控制稳定不同业务类目的语气、术语和格式，减少小语种质量波动。
- 基于 Golang html.Tokenizer 做流式标签掩码与还原，避免大模型翻译破坏富文本结构，HTML 结构故障率从约 15% 降到接近 0。
- 搭建质量评估工作台，支持热词 / 随机 / 对照抽样、SBS 盲测和回归验证，形成“调整 → 评测 → 回归”的闭环。
- 项目收益：外包校对从 20 人降到 2 人，年省约 10 万美元；术语匹配从 200ms+ 降到 <5ms；内容上线从 3 天缩短到 2 小时。

独立开发者 | 精酿品牌舆情雷达与竞品监控系统

2023.07 - 2024.05

- 独立完成多源内容采集、去重留痕、关键词规则引擎、事件归并、风险预警和证据检索系统。
- 将食品安全、质量问题、服务纠纷、价格争议等风险从“偶然刷到”变成可追溯、可分派、可处理的日常流程。

字节跳动 | 高级研发工程师 / 质量方向后端负责人

2018.08 - 2023.05

商业化审核质量平台与规则系统

- 从一线后端成长为商业化审核质量方向负责人，负责搜索广告审核链路、审核人员质量平台、规则管理平台、销售复审和客户咨询平台等系统建设。
- 从 0 到 1 搭建搜索广告商业化审核链路；通过机器审核、历史结果复用、异步化与下游接口优化，审核延时降低约 50%，状态同步延时从 2400ms 降到 700ms。
- 建设审核人员质量体系，覆盖质检、真实任务混入、模拟审核、考试认证、案例库和质量配置平台，形成从培训、检验、复盘到技能认证的质量数据闭环。
- 主导规则管理平台和查询服务重构；客户负反馈率从约 3% 降到 0.75%，查询延迟从 200ms+ 优化到 5ms 内，稳定性提升到 99.99%。
- 负责需求拆解、技术评审、排期、代码质量和跨团队协作，带领 4 人小组获得研发质量奖。

北京一土网络科技有限公司 | Python 开发工程师

2017.07 - 2018.08

- 负责线下学校家校沟通 App 的多个微服务模块接口开发与优化，参与 Flask、Django、MySQL、Redis 相关服务开发。

教育经历

哈尔滨理工大学 | 软件工程 | 本科

2013 - 2017

投递关键词

AI Agent 开发、Agent 工程化、智能体运行支撑体系、多智能体协作、工具调用、上下文工程、长期记忆、检索增强生成 (RAG)、质量评测、运行日志、故障定位、安全门禁、后端架构、平台工程、AI 辅助研发。