

我能做什么 WHAT I DO

从一个想法出发几小时搭出能跑的原型，也能设计评估方法判断它值不值得活下去。身份上在 Chase 定义的 Builder 和传统 PM 中间地带：能自己 Build 做验证，也能设计标准判断。

– 深度使用 **claude code**, **claude.ai**, **claude app**, **claude design**, 搭配 **graphify** 加强的 LLM wiki——把验证一个想法的周期 **从 3 个月压到 1 天**。

– 让评估维度从数据里长出来。AI 问答 trace 里浮现出 11 个维度（3 个是预设 rubric 想不到的）。

– 从判断到真上线。qibook 从查企业扩大到查“人”的新业务线，4 次碰壁每次用实验数据转向，反向蒸馏用户品味，**200 字三段论上线 qibook.com**。

教育 EDUCATION

蒙特雷国际研究院 MIIS

翻译与本地化硕士 (STEM)

2023.09 – 2024.06 GPA 3.9 / 4.0

2023–2024 走北外派出交换至 MIIS

北京外国语大学

会议口译硕士

2022.09 – 2025.06 GPA 3.8 / 4.0

厦门大学

英语文学学士 (Top 5%)

2018.09 – 2022.06 GPA 3.9 / 4.0

技能 SKILLS

AI 与 Agent

Claude Code / Claude App / Claude.ai 深度使用 · MCP 协议 · Multi-Agent 编排 · Prompt 工程 · 评估方法论设计 · Langfuse · Coze · LangGraph

产品

PRD · JTBD · A/B 测试 · 标注体系设计 · 意图分类 · 情景模拟 · 用户研究 · 跨团队推动 · 数据需求方法论

语言

CATTI 一级笔译 · 二级口译 · 中英双语原生

独立产品 INDEPENDENT

AI PMF Validator

基于 OASIS 开源平台搭 AI 虚拟焦点小组—验证产品想法商业可行性，为每一个人提供轻量商业咨询。

人生模拟器

平行人生模拟器让你探索那些“没走的路”。输入你的真实性格、经历和生活条件，观察自己平行人生的多种可能性。对话式人格采集 --> AI 人格档案生成 --> 5 层变量系统 --> 人生模拟与人生分支 --> 分层记忆与记忆可视化 --> 跟任意年龄的“世另我”对话/多版本互相对话

deep-decode 内容管线

选题到 四件套产出自动化: 骨架 → 草稿 → 信息图 → Word → TTS 播客。三源交叉验证防单一视角偏差。

完整故事、每个项目的 2000+ 字长文、设计决策复盘，都在 mingmingzi.com。

工作经历 EXPERIENCE

中数智汇 Chinadaas / AI 产品经理

JUL 2025 – PRESENT

企百科对话产品评估体系 自主发起

企百科上线了 AI 对话但没有上线标准——PM 说“答案不好”没数据依据，研发改完 prompt 无从验证变好变差。Rubric 保质期短：业务从“报告”变到“对话”就废一批维度。

用 **标注体系设计** (按工具调用链路分成简单商查 19 + 复杂商查 19 = 38 个涌现类目, 6–9 轮迭代冻结) + **三模型共识投票** (Qwen + GLM + Kimi 共识; Qwen3.6 + GLM-5.1 + DeepSeek-V3.2 多数票 2/3) + **LLM-as-a-judge** 自动化链路, 把评估从 PM 拍脑袋变成可复用流水线。Langfuse 上让评估维度从 10 条 Agent trace 中自然涌现出 11 个 (3 个是预设 rubric 想不到的)。

→ 发现 66.8% 满足率的瓶颈不在答案质量而在 68.2% 的误拒, 把产品团队的优化重心从答案质量翻到意图识别层, 单项潜力 +23%。

养一只虾: 从 chat bot 到会主动联系我的代理 OpenClaw + PM 分身 Amory

AI 工具在 terminal 和网页里、日常工作在飞书——两边不通, 灵感来了得切窗口找工具 (问题)。部署踩三层坑: API 限流、SSL 与权限、弱模型编排强模型会拆错任务 (编排者的天花板就是系统的天花板)。

用 **心跳协议** (评分公式 $\text{impact} \times \text{confidence} - \text{cost}$) 让 Agent 自主扫 8 个数据源决定做不做 + 熔断机制防 token 烧穿; PM 分身用 **行为考古**——从 Claude Code 会话和 Redmine 日志归纳自己真实高频动作 (最高频是上传附件和更新需求字段, 不是写新需求), 筛出汇报 / 发需求 / 数据分析等场景做成 Skill。

→ 飞书 @ 一句话跑完以前 30 分钟手动汇报流程; 把我从被动切窗口改成被 Agent 主动推动。

qibook.com 个人工商简报 查“人”新业务线 · 0→上线

公司只能查企业, 但 7.8% 的用户需求是“拿人名查关联企业”——没人做过的新业务线 (问题)。纯 LLM 生成 100% 幻觉 (EO 实验验证), 22 版可视化无法收敛, 5 个银行部门定制版 PMF 5.25 分全 Dead End。

用 **反向蒸馏** (综合行业报告 + LLM 模拟银行客户经理视角, 把模糊认知翻成 6 类身份标签) 定义分类体系; 用 **data-req 四步数据需求方法论** (盘清 19 API / 1445 字段 → 从 1245 条会话做意图分类 → 缺口矩阵 → 数据需求文档) 打通数据到产品全链路。

→ 200 字三段论让银行客户经理 10 秒看懂一个人的商业角色和版图分布, 为公司从“查企业”扩出“查人”新业务线。

企百科 AI 搜索可见性 (GEO) 自主发起

在豆包搜“企百科”第一页只出应用商店——就绪度 F 级, 0.67/10。最根本障碍不是 SPA 空白渲染, 是网站对外关闭——说服 B2B 应用层团队“为什么要关心 AI 搜索”是组织问题不是技术问题。

写 SPA 语义注入方案 (JSON-LD + 语义 HTML + sitemap) 被研发采纳, 跨团队推动把网站对外打开; **搜索引擎逆向分析** (246 条 AI 搜索结果按域名统计) 发现外部占比悖论——千问引用中外部平台占 70%+; **autoresearch 方法** (Karpathy 框架, AI 自己跑 A/B 迭代评分) 优化软文 Skill, 把软文评分从 7.0 推到 12/12 满分。

→ SPA 语义注入 + 网站开放已落地, 让企百科从 AI 搜索完全不可见变成具备被索引的基础条件。

中数论坛 知识飞轮 + 方法论反思

公司内部知识散落在个人脑袋和飞书群聊里, 没有沉淀复用机制。冷启动是核心困难——上线后 8 用户 2 赞, 绝大多数帖子是我自己发。

设计 **Git 到 Agent 的语义协议** (普通虾 4 动作 push/comment/pull/star, 管理员虾 6 动作含 review/merge/rebase), Discourse 部署 + 论坛心跳 API 对接完成。

→ 真正的收获是一条方法论反思: 产品设计的正确顺序是用户真实行为 → 流程想象 → 产品设计, 从 UI 反推是死路。

标讯推荐 投标书 + 标讯 LLMwiki

同事要在 2 天内交一份 176 万投标项目的 9 章投标书初稿; 由此引出: 怎么找到合适的标? 现有剑鱼千里马都在做推荐, 纯拼“更准”没差异化空间。

用 Claude 一个 session 产出完整投标框架; 用 **情景模拟** (虚构用户老王的一天) 重新定位——用户不是“查标讯的人”而是“经营一家投标公司的人”; 设计 **画像自耕田机制** (决策因子 × 用户偏好双轴, AI bookkeeping 人 curation, 用户不断上传文件和提问, 标讯推荐越来越懂用户, lint 功能淘汰过期标/客户/竞对)。

→ 把标讯系统从信息推送重新定位为“经营助手”, PM 角色扮演用户从记忆出发设计产品。