

晋佳莉

15240034478 | 1014223067@qq.com |

中国科学技术大学 工商管理硕士 (全日制国际班-选拔制)

作品集: <https://jinjiali.bozipaopao.cn/>



个人优势

- **一人 + AI 端到端交付**: 以 Agentic Coding / Vibe Coding 方式独立完成从产品规划、交互原型、功能开发到测试验收与上线迭代的全链路工作, 累计交付 7 款 AI 产品, 覆盖 RAG 智能问答、多模态内容生成、Agent 任务执行三类核心场景。
- **AI 原生架构设计**: 熟悉 Context Engineering、Agent Harness、MCP / Tool Use、Agent Orchestration、Agentic RAG、Memory System、Guardrails 与 Observability; 能在产品设计阶段即完成 Agent 角色拓扑、任务链路、工具权限、Handoff 机制、Human-in-the-loop 边界与异常降级策略设计。
- **Eval 工程与 Badcase 闭环**: 以 Evaluation-Driven Development 方式从 0 到 1 搭建覆盖模型输出、检索、工具调用、任务轨迹与最终结果的多层 Eval 体系, 维护 Golden Dataset、回归集、对抗集与边界 Case, 评测手段含规则评分、Code-based Grader、LLM-as-a-Judge 与人工校准; 建立“线上监控 → Badcase 识别 → 聚类归因 → 评测集回流 → 回归验证 → 灰度发布”的持续闭环, 按 Prompt / Context / Retrieval / Tool Use / Memory / Workflow 归因, 关注任务成功率、Groundedness、Recall@K、工具调用准确率、幻觉率、P95 延迟与 Cost per Task。

实习经历

北京乐信圣文科技有限责任公司 | AI 产品经理

2026.5- 至今 (实习)

游戏出海 AgentOS —— AI 广告创意素材自动生产系统 (toB & toC)

项目背景: 休闲游戏与轻应用海外发行商, 买量 + 广告变现 (IAA) 驱动。负责把“人工做素材”改造成“Agent 批量产素材”, 产出直接进买量投放。链路: 素材灵感采集 → 三元组决策选材 → LLM 导演编排 → 分镜生成 / 配音 / 字幕 / 合成 → 自动质检 → 合规闸门 → 入库投放。

我的职责:

- **Agent 架构与编排**: 主导生产引擎侧 (Orchestration / 原子能力 / 质检 / Observability) 的产品规划, 与决策大脑侧 (采集 / 反馈飞轮) 划分 Agent 角色拓扑、Handoff 契约与工具权限边界; 设计 plan 驱动的任务链路与三态退出语义 (全成功 / 部分产出 / 全失败), 使“要 5 支只出 2 支”不会被判定为成功。
- **Guardrails 与 Human-in-the-loop**: 建立版本化合规判据库 (判据指纹可回溯) 与三道内容闸门 (玩法同族 / 视觉归族 / 残留检测), 统一 fail-closed 原则“判不了即拒”; 高复刻度产物强制标记并设人审卡点, 其余走自动准入门禁 (未验证组合不得排产)。
- **Eval 体系 (EDD) 从 0 到 1**: 搭建覆盖模型输出、选材检索、工具调用、任务轨迹到最终成片的多层评测: Code-based Grader (产线 lint 七项 + 维度生效探针)、LLM-as-a-Judge 两票一致、规则评分与人工校准 (抽帧逐字比对); 沉淀 Golden 组合矩阵、回归集、对抗集 (规则绕过样本) 与误杀集 (假阳性样本)。通过判据分层缓存 (昂贵事实与廉价判决分离) 使换目标 / 改阈值场景零新增模型调用, 重复评测 Cost per Task 降至近零。
- **Badcase 回流与数据对账**: 对采集端 → 决策端 → 生产端做全链路真值基线, 定位学习闭环断点与统计口径失真 (而非凭日志字面判断); 真实失败案例自动转为回归 Case, 同类问题上线前必跑回归。
- **Observability 与协同机制**: 把一次性脚本收敛为“有责任人、有台账、有回归测试”的产品能力: 一次任务在执行群只占一条原地刷新的聚合卡, 失败单独告警并携带错误原文; 跨算法 / 工程 / 投放 / 法务对齐口径, 命名契约与合规类决策收口团队评审。

工作成果:

成果一 · **Cbrain 全自动素材流水线**: 沿「采集—判断—生产—反馈」四段闭环建设: 实现三条采集链路累计入库素材约 2.1 万条, 沉淀 782 个素材类型、53,065 个结构化元素。生产端完成长视频稳定性攻坚: 剧情类由稳定触发 70 分钟超时中断、出片率 0%, 提升至稳定 38-42 分钟出片。

成果二·Raphael 多产品多语言产线体系：cbrain 验证自动化可行性后，业务重心由「能生产」转向「规模化覆盖」，拉斐尔将素材生产沉淀为可复用的产线资产——每类创意玩法固化为一条标准化产线，以「product*lang*idea」为输入即可批量出片，并在合规、命名、准入三个维度约束产出。目前体系已实现覆盖 7 个自研产品、实现上游 idea 自动搜集判断产出 idea，已产出 18 条标准化产线。配套还搭建了文案种子库、美术资产库等供产线调用。

上海克莱音科技有限公司 | AI 产品经理

2026.01 - 2026.04（实习）

Machi —— 东南亚演出票务平台（toC）

项目背景：东南亚线下演出市场增长迅速，但票务长期被第三方平台主导，其 15%–25% 的高额抽佣与资金代收、结算滞后的模式，持续挤压主办方利润并拖慢回款。Machi 以此为切入点确立差异化定位：观众付款后由主办方自有 Stripe 直接结算，平台不充当收款主体、不占资金池，仅收取约 5% 平台费，以「把渠道抽成还给主办方」作为核心获客主张。

我的职责：

- **市场调研与商业模式定义：**调研东南亚票务市场与第三方平台抽佣结构，确立差异化定位：观众付款 → 主办方自有 Stripe 直接结算，平台不作为收款主体、不占资金池，平台费约 5%，以「把渠道抽成还给主办方」为核心获客主张。输出完整产品需求文档，逐条明确费用归属、跨境币种、退款边界等业务规则。成果：平台费 5% 对标第三方 15–25% 抽佣，形成可量化的主办方迁移理由；覆盖 6 城 8 类演出品类，沉淀 13 项经决策确认的业务规则。
- **三端产品设计与交易链路：**与 UI 协同完成三端信息架构与页面设计，打通「发现 → 选票 → 结算 → 电子票 → 核销」主链路；接入卡、微信支付、支付宝、Apple / Google Pay 多支付方式，完成 SGD 含 GST 结算与多币种展示合规适配；退款改期设计为「主办方可配置 → 受平台上下限约束」，在主办方自主权与平台风险敞口之间取平衡。成果：交付三端共 46 个页面，支持 5 种支付方式与 5 种币种；库存锁定、超卖防护、票据签发等交易关键路径全部具备服务端强校验
- **AI 能力产品化：**主导三项 LLM 能力的产品定义，重点在护栏而非能力演示：① 活动页草稿 —— 一段演出描述 → 生成标题、摘要、详情、日程、FAQ、建议票档与 SEO 标题，逐字段可接受 / 重生 / 手改，将人工校准嵌入主流程；价格、时间、场馆等事实字段设零幻觉护栏。② 活动问答助手 —— 以单场活动数据为知识边界做检索域隔离，杜绝跨场次串答；置信度低于阈值 → 自动转人工，退款改期等具资金与合规后果的问题 → 强制人工，据此划定 Human-in-the-loop 边界。③ 自然语言搜索 —— 口语查询 → 结构化筛选条件 → 逐条可视化、可单独移除，使用户能修正机器的理解而非接受黑盒结果，纠偏动作同步沉淀为失败样本。成果：活动发布耗时数天 → 分钟级；助手支持三语并以自助解决率为效果指标，该指标直接对应主办方客服成本，构成其配置助手的动因。
- **上线质量治理与迭代：**建立上线前质量基线并推动缺陷根因治理。定位线上「半中半英」真根因：本地化兜底逻辑以子串替换填补未命中文案 → 缺失翻译被静默掩盖 → 等同关闭报警；据此删除兜底使缺失显式暴露，补齐词典并修复文案与词典的键值漂移，将命中率检测与回归测试固化进发布流程。成果：全站文案词典命中率由 33% 提升至全覆盖，新增 300+ 条译文；跨 32 条路由 × 8 种视口的可访问性与布局审计实现零失败。

美的集团 | 产品经理

2020.09 - 2021.05（实习）

IOT——万物互联

主导物联网产品线（设备接入、场景联动）需求规划与落地，协同软硬件研发与测试等跨职能团队，推动端云协同的敏捷开发机制落地，对产品交付质量负责。

工作经历

中国信通院江西分院 | 产品经理

2021.07 - 2025.08（工作）

政企 AI 产品线（政务 RAG 知识库）

依托智库研究优势，主导政企场景 AI 产品全生命周期管理，牵头两条产品线从 0 到 1 并规模化，支撑 10+ 省市级数字化转型与产业赋能项目。

- **政务 RAG 知识库（0 → 1）：**设计知识采集 → 文档解析 → 切分向量化 → 检索召回 → 大模型生成的完整 Agentic RAG 链路；以检索召回率、答案可溯源性、回答采纳率与幻觉率为核心质量口径，建立 Badcase 采集、聚类归因、优化验证的迭代闭环，把线上失败问答沉淀为固定评测集，版本上线前必须跑回归。

- **AI 数据底座与知识工程**：主导 3000+ 家企业业务数据采集、清洗与治理，构建领域知识库、业务标签体系、农业数字经济指标体系与知识图谱本体模型，为 RAG 检索增强与模型微调提供高质量语料底座，也是后续做上下文工程的地基。
- **需求挖掘与方法论沉淀**：面向政务、智能制造、数字经济领域主导 300+ 家政企深度调研并输出诊断报告，将复杂业务流程拆解为 AI 可执行的原子任务，直接支撑产品定位与功能优先级；沉淀 100+ 份场景化 AI 解决方案与可复用交付方法论，交付效率提升 35%。

项目成果：完成 500+ 篇政策文件做结构化清洗与分块策略设计，以中文语义检索召回率为核心指标完成 Embedding 选型对比（BAAI/bge-large-zh vs. text2vec 系列），引入 Reranker 对 Top-K 候选精排，答案相关性提升约 18%，Embedding 召回准确率 82%；设计 Router 按用户身份分发链路（企业走政策匹配与申报指导、政府走政策解读与合规建议），匹配对应 System Prompt 风格与检索策略，并定义各链路输入输出规范与低置信度转人工兜底；建立检索召回率、答案相关性、用户满意度的分层评估体系，累计服务查询 10,000+ 次，政策申报辅助效率提升约 40%。

独立项目经历（一人 + AI 端到端交付）

EditAI 智能写作助手 | AI 产品负责人

2025.11- 2026.3

面向深度写作者“单一 AI 对话改稿缺乏结构化诊断、易丢失作者调性”的痛点，独立设计并落地本地优先、人机协作的多智能体写作工作台，跑通调研 → 写作 → 精修 → 事实核查 → 归档全链路。

- **Agent 角色拓扑与权限约束**：设计 Chief 主编 → Deputy 副主编 → Editor 编辑三层协作链路，定义各角色能力边界（Editor 设为只读、低温度、禁用外部工具调用），确保 Agent 行为可控可审计。
- **结构化质量评分与自动路由**：设计 Polish（文采）/ Logic（逻辑）/ Consistency（一致性）三维评分体系并定义通过线（ ≥ 0.70 ），Chief Agent 按评分自动路由：高分通过 → 中分重改 → 低分回退 → 超阈值升级人工，把改稿质量从主观感受变为可度量、可兜底的闭环。
- **流程规则与风格保护**：规划调研、写作、编辑、核查、分析、提取、归档等九大工作模式，确立阶段不可回退、关键节点自动保存、报告驱动 + 对话驱动双模式核心规则；设计写作风格指纹机制与 analyze / edit 双模式防止 AI 改稿覆盖作者原有风格，并制定版本冲突保护与全链路行为埋点方案。
- **成果**：已部署上线，兼容任意 OpenAI-compatible 模型，本地优先保障数据隐私；跑通“精修报告 → 对话改稿 → 质量路由 → 事实核查”完整人机协作流水线，内测用户改稿效率提升 60%。

AI 相机租赁智能客服系统 | AI 产品负责人

2026.4 - 2026.07

面向相机租赁业务人工客服响应慢、订单管理成本高的痛点，独立设计并落地覆盖 C 端服务与 B 端管理的智能客服助手，基于 RAG + Agent 架构实现从设备咨询到订单操作的全场景闭环。

- **双知识库与检索隔离**：构建设备知识库（参数、使用教程、故障处理）与租赁规则知识库（定价、押金、损坏赔偿），设计知识库间的检索隔离与协同策略，实现精准问答。
- **意图路由与 Tool Use**：设计意图路由区分“设备咨询 / 租赁问答 / 订单操作”三类意图并匹配处理链路：知识问答走 RAG 检索召回、订单操作走 Agent 工具调用；通过结构化指令与 Tool Use 打通订单系统读写，支持自然语言查询设备排期、更新邮寄状态、确认订单信息，把 AI 从问答延伸到业务执行，实现 AI 导购功能。
- **B / C 端权限边界**：区分 C 端客户服务与 B 端自主管理两类视角，分别设计对话路径与权限策略，确保 B 端管理操作的安全边界。
- **成果**：覆盖设备咨询、排期管理、邮寄跟踪三大场景，咨询响应从人工平均 30 分钟压缩至秒级，常见问题自动解答率 85%+，每周人工介入订单管理时长减少约 60%。

教育经历

中国科学技术大学 | 工商管理（全日制国际班-选拔制）· 硕士

2025.09 - 2027.07

GPA:3.75/4（专业前3%）-全英文授课及考试

管理信息系统（97 分）、电子商务（97 分）、数据、模型与决策（90 分）、管理经济学（94 分）

长江大学 | 国际经济与贸易 · 本科

2017.09 - 2021.06

GPA:3.5/4（专业前5%）

CET-6:515 分