

钟斯慧

AI 产品经理 | Agent · RAG · AI Eval

5 年数据与 AI 行业经验 | 大模型算法 & 大数据背景 | 女 | 硕士

15920327925 | diarn137@126.com

代表成果

- 企业级 Agent 搭建：参与从 0 到 1 搭建任务路由、Plan-DAG、工具调用和可信检索闭环
- AI Eval：设计评测与归因体系，实现回答质量向 Planner、检索及工具节点回溯
- 智能客服：负责洗护家政客服场景产品方案，覆盖 10+ 服务品类、20+ 核心业务流程

专业技能

1. **产品设计与落地**：具备用户场景和业务链路拆解、MVP 定义、版本规划、PRD、原型、状态机及验收标准设计能力，能够推动业务、算法和研发协同交付
2. **RAG 产品方案与评测**：熟悉文档解析、切分、混合检索、多源召回、重排、生成和引用链路，能够结合 LLM-as-Judge、统计指标与人工复核设计 Eval 系统架构
3. **Agent 与工作流**：理解 Agent 任务规划、工具调用、状态管理和人工兜底机制，能够针对业务流程和任务类型设计 Plan / React / 订阅式 Agent 架构
4. **模型应用与选型**：理解 Transformer 及主流大模型架构，熟悉 Prompt、上下文工程、LoRA / QLoRA、SFT 及偏好对齐方法，能够根据场景权衡模型效果、延迟、成本和部署方式
5. **数据与系统架构**：理解向量库、图数据库、关系数据库、缓存系统的使用边界，熟悉 Chroma、Neo4j、MySQL、Redis、Doris 等的使用场景。拥有离线 / 实时数仓，数据治理经验
6. **数据分析**：使用 Python、SQL 完成数据分析处理、具指标体系、漏斗分析和 AI 链路观测能力

工作经历

源启开物（北京）科技有限公司

北京|2025.12-2026.05

AI 产品经理 | 2025.12-2026.05

- 负责面向企业员工的 AI 办公产品设计，一站式评测平台构建

深圳市丰巢网络技术有限公司

深圳|2021.07-2025.09

产品经理 (LLM) | 2024.12—2025.09

- 负责业务线社群运营系统产品设计，0-1 落地

大模型算法工程师 | 2022.09—2024.12

- 负责企业级大模型应用的构建与优化，涵盖模型训练、评估及业务落地效果验证

大数据开发工程师 | 2021.07—2022.09

- 负责广告实时数仓建设与指标体系设计，主导广告数据仓库重构方案制定与落地，支撑多系统广告数据分析与应用

深圳市蜜蜂互联网络科技有限公司

深圳|2020.04-2021.05

大数据开发工程师 | 2020.04-2021.05

项目经历

联想百应智能体 超级 AI

项目描述：

- 面向文件理解、联网检索、跨材料分析等复杂信息任务，从 0 到 1 构建统一 Agent 任务架构，将原有搜索整合/单次生成能力升级为可规划、可调用工具、可追踪执行结果的端到端 AI Agent

核心工作：

1. 基于办公写作、资料分析、实时查询和多轮追问等场景，构建“简单任务直答、复杂任务规划执行”的任务分成与路由策略
2. 采用“会话上下文继承与多轮查询改写 + 单次 Planner 生成 DAG + 静态执行”的方案。复杂任务通过 Planner 拆解为可执行步骤，DAG 支持无依赖步骤并行、依赖步骤串行及基于上游结果的查询改写
3. 将会话附件、知识库和联网搜索抽象为统一检索能力，按会话附件、知识库、联网的可信信息优先级选取结果
4. 按简单直答、单工具任务和多步骤任务分层设计指标，覆盖路由准确率、Plan 有效率、工具调用成功率、首轮完成率、TTFT 时延、引用覆盖率及用户追问率，推动模型效果从主观体验转化为可测试、可复盘的产品指标

项目成果：

1. 完成超级 AI 一期上线，形成“输入理解 → 任务路由 → Plan/DAG 执行 → 工具调用 → 结果汇总与溯源”的端到端 AI 任务闭环
2. 核心任务首轮完成率 84.7%、工具调用成功率 93.8%、简单任务 P95 响应时长 1.3s

联想百创平台 超级 AI Agent Eval 与能力治理系统

项目描述：

- 联想百创平台为面向企业用户，针对办公场景提供智能体搭建，自建工作流，知识库服务的平台
- 针对规划、工具与检索链路难以归因的问题，设计贯通质量评测与能力治理的产品方案

核心工作：

1. 规划“样本接入—自动评测—人工复核—结果发布”四阶段评测链路，统一定义样本、事件、评分与版本基线数据模型
 2. 建立答案正确性、问题解决度、合规安全、用户体验和服务效率多维质量标准，按维度匹配规则引擎、LLM-as-Judge 与行为信号，并输出结构化评分、置信度、失败标签与可解释证据
 3. 设计根 Trace 到节点 Span 的端到端埋点，覆盖改写、路由、Planner、DAG、检索、工具和最终回答
 4. 融合规则检测、LLM Judge 和行为信号进行交叉评测，输出分数、置信度、失败标签及可解释证据
- 规划 Skill 与 MCP 的能力快照、版本管理、最小权限和调用护栏，支持复杂任务的可审计扩展

项目成果：

1. 形成统一事件、评测样本、人工复核与质量发布数据模型，明确跨 Agent 运行轨迹的关联规范
2. 输出从低分答案回溯至计划、工具、检索或权限决策的诊断链路，支撑版本比较和回归测试
3. 明确自动评分仅用于预标注和选样，人工确认后才进入金标集与质量基线的治理机制

项目描述：

- 面向特定健康公众号社群，针对非专业读者阅读免疫与功能医学文章时，专业名词理解门槛高、外部搜索打断阅读、通用大模型回答缺乏可靠出处等问题，设计垂直 AI 浏览器插件

核心工作：

1. 从 0 到 1 完成用户场景、产品定位和功能优先级设计，规划划词解释、段落提问、作者观点脉络、私人资料关联及联网深查的分级能力，明确公共知识库、私人数据和外部信息的使用边界
 2. 聚焦社群读者的阅读场景，定义“划词—快速理解—查看证据—延伸阅读”的核心任务
 3. 规划垂直知识库建设方案，定义文章与书籍的混合召回、来源隔离、引用校验及增量更新流程
- 建立覆盖检索召回、生成质量、引用真实性、隐私安全、响应延迟和调用成本的 AI 评测框架

项目成果：

1. 完成从用户需求、MVP 范围、AI 能力边界、知识库治理到评估发布的完整产品方案
2. 招募种子用户，建立首轮内测样本，围绕核心任务完成率、解释有效率、知识库覆盖率验证假设

丰巢洗护家政系统 智服小丰

项目描述：

- 面向洗护家政消费者及客服运营团队，建设覆盖售前咨询、服务推荐、订单查询、物流查询和售后处理的智能客服系统

项目亮点：

1. 设计“流程引擎+RAG/GraphRAG+LLM 生成+人工兜底”的混合方案，根据问题类型选择理链路
2. 针对下单、订单/物流查询和售后处理等任务，设计意图识别、槽位收集、状态流转的多轮业务任务闭环
3. 多层知识检索体系：针对商品知识、服务规则和洗护保养文档建设向量知识库；对于涉及商品、材质和服务关系的复杂问题，引入知识图谱和 GraphRAG，提升关联问题的检索与回答能力
4. 结合用户咨询内容、服务适用条件和优惠券信息，为用户推荐匹配的洗护或家政服务，完成咨询到转化的服务链路

项目结果：

1. 业务覆盖：系统覆盖洗护家政 10+ 服务品类和 20+ 核心业务流程，支持售前、订单、物流和售后等主要服务场景

丰巢洗护家政系统 智服小丰偏好对齐优化

项目描述：

- 对回复生硬、风格不一致和业务引导不足问题，通过评测体系增强后训练效率，优化模型回复质量

项目亮点：

1. 将优质客服回复拆分为业务正确性、安全合规性、任务有效性、信息完整性和表达自然度五个维度
2. 真实客服对话、模型低质量回复和异常场景中提取数据，覆盖模糊表达、多意图、口语化、信息缺失、售前推荐、闲聊和高风险回复等场景
3. 结合自动化指标、人工评审和线上 A/B 测试，从模型质量、用户体验、任务完成和业务安全多个层面验证优化效果

项目成果：

1. 通过 SFT + DPO 微调，模型在客服场景下的满意度提升 18%（业务侧调查）
2. 多轮任务型对话完成率提升 22%，闲聊场景不合规回复率下降至 0.5% 以下

算法项目

- 丰巢广告系统 标题生成

PyTorch + Transformers + CLIP + GPT2

- 基于多模态模型帮助广告主快速生成广告爆款标题，提升创意生产效率与营销转化率
- 丰巢广告系统 智能广告运营助手

LangChain + Qwen2.5 + Chroma

- 基于 RAG 架构的智能助手，统一知识检索与业务数据查询，辅助运营与销售快速响应客户需求
- 丰巢储物系统 存储物品品类识别

PyTorch + Chroma + Pillow + FastAPI

- 对用户储物前上传的物品图片进行识别与分类，自动检测物品类型，对潜在危险物品进行识别
- 丰巢洗护系统 智能地址修正

BERT + Transformers + PyTorch + RrapidFuzz

- 针对地址缺失、错别字或格式不规范导致的物流派送失败、通过 NLP 技术智能抽取与标准化修正

数据项目

- 丰巢广告系统 Flink 实时数据仓库

Kafka + WaterDrop + Flink + Doris

- 基于 Flink 开发统计分析平台，对广告领域行为及业务数据进行分析，促进用户增长与用户粘性增进
- 丰巢广告离线数仓

Hadoop + Hive + Spark + Atlas + Zabbit

- 负责广告业务的离线数仓与数据治理体系建设，构建指标体系，实现广告数据的高质量集成

教育经历

横滨国立大学，经济学，硕士	横滨 2017.04-2019.03
华南理工大学广州学院，日语，本科	广州 2010.09-2014.07

自我评价

- 具备良好的沟通表达能力、团队协作能力较强，和较强的学习意愿以及跨学科能力
- 熟练使用多种语言，英语（TOEIC 720，熟练）、日语（N1，精通）、粤语（良好）、韩语（熟悉）