



## 教育背景

- 北京大学-外国语言文学（国别和区域研究）博士 区域与国别研究院 全日制 2022年09月 - 2027年06月
- GPA: 3.7/4.0 | 获北京大学三好学生、优秀科研奖
- 中山大学-国际政治 本科 国际关系学院 全日制 2018年09月 - 2022年06月
- GPA: 3.9/4.0 | 获教育部国家奖学金、中山大学优秀共青团员、优秀学生一等奖学金（连续两年）、王老吉奖学金等

## 项目经历

- 全球商业洞察 Agent — 独立完成 2025年12月 - 2026年08月
- 项目概述: 从 0 到 1 独立完成面向企业出海决策的 AI Agent 产品，覆盖用户研究、产品设计、评测迭代与部署上线全链路。
- 用户需求研究:** 基于印尼 6 个月的长期田野调研及 30+ 出海企业负责人访谈，识别本地信息不足、跨国资料误用、结论无依据等痛点，并将其转化为国家匹配、证据支持、来源质量、行动建议四项产品指标
  - 产品架构设计:** 设计 Harness + Multi-Agent 架构，由 Harness 负责搜索路由、上下文管理和质量门禁，协同 Search Planner、Insight、Report、Critic 等 Agent 完成搜索规划、证据分析、报告生成与引用核验；搭建覆盖 195 个国家、2,367 个分层信源的 Market Pack，其中包括 1,265 个官方信源、1,102 个本地媒体，支持按目标国家调用当地语言来源。构建由外交部、商务部等 311 份官方文档、71,904 Chunk 组成的知识库，通过人工标注与 Top-K 评测确定 K=8，通过RAG设计补全基础信息缺乏。
  - 评测迭代:** 建立代码硬门禁与 LLM-as-Judge 七维评测体系，构建 10 题回归集和 5 题泛化集。经过 5 轮迭代，综合评分由 2.94/4 提升至 3.33/4 (+13%)，Actionability 提升 25%，Citation Quality 提升 15%，硬门禁通过率达到 100%。
  - 工程落地:** 基于 Vercel、Render Worker 与 Supabase 完成线上部署，实现多国家研究、多轮追问、企业知识库及运营监控看板。
- 知香水导购 Agent — 独立完成 2026年07月 - 2026年08月
- 项目概述: 从 0 到 1 独立完成香水导购 Agent，打通自然语言理解、知识图谱检索、真实商品推荐、评测迭代与部署上线全链路。
- 需求研究:** 通过 20+ 电商消费者访谈，识别香水知识门槛高、线上难试香、推荐缺乏解释等痛点，设计自然语言找香、多轮澄清、真实商品推荐及偏好记忆流程，确立“知识图谱负责理解、真实商品源负责供给”的产品边界。
  - 产品架构设计:** 设计领域路由、Orchestrator 对话编排及“意图解析—知识检索—证据生成—商品召回”推荐链路；构建覆盖 1,971 款香水、346 个品牌、504 个香原料及 12 类香调的知识图谱，并接入拼多多与淘宝联盟真实商品源；由 LLM 处理模糊语义与推荐解释，设计“精准款—香调方向—性别兜底—品类兜底”的分级召回策略。
  - 评测迭代:** 建立代码硬门禁 + LLM-as-Judge 多维评测体系（23 题基础回归集、23 题改写外推集、10 题香调 bad-case 专项集）；经三轮迭代，基础集综合评分 2.26/5→3.14/5 (+39%)，平均维度分 2.60/5→3.52/5 (+35%)。
  - 工程落地:** 基于 Taro / React / TypeScript / Node.js / Express / SSE / PostgreSQL 完成 H5 产品及线上部署，实现流式对话、商品卡片、历史会话、用户反馈与运营监控，并补充会话隔离、超时重试、数据源健康检查及评测断点续跑机制。
- 国际气候谈判 LLM 智能标注系统 — 主导完成 2024年01月 - 2025年05月
- 项目概述: 主导 AI 数据生产产品从 0 到 1 落地，成果入选 2025 Asian PolMeth Conference。
- 需求与路线验证:** 识别原数据库依赖人工标注、仅覆盖 1995-2013 年的问题；基于 6.2 万+ 历史标注数据完成 Prompt、Few-shot 与 SFT 路线验证，论证 SFT 在复杂关系抽取中的优势并推动项目转向。
  - 产品架构设计:** 协同北京大学王选计算机研究所设计 Expert-Augmented LLM Annotation 架构，将领域 Codebook、任务拆解、模型标注与专家校验整合为标准化 workflow；定义 Prompt、SFT 与人工复核的适用边界，兼顾模型效果、标注成本与数据质量
  - 数据闭环与模型迭代:** 组建并管理 11 人标注团队，建立“LLM 预标注—人工审读—一错误归因—训练数据回流—模型迭代”闭环；以 Codebook 约束 + 任务拆解将复杂任务 Recall 从 54.0% 提升至 87.9%
  - 落地成果:** 应用于 678 份国际气候谈判日报，自动生成 45,051 条互动数据，覆盖范围扩至 1995-2024；Derivation 规则符合率 59.1%→88.2%，超过人工标注基线；模型 Precision 74.0%、Recall 90.0%。

## 实习经历

- 北京大学区域与国别研究院-战略研究团队分析师 2022年01月 - 至今
- 主导设计并落地 AI 战略研究工作流，将 LLM 深度嵌入信息检索、多源材料整合、初步分析、系统润色全链路，研究产出效率提升 60%+；独立完成 33 篇研究报告报送（1 篇获中宣部等部委采纳），产出学术论文 3 篇、中英文媒体政策文章 9 篇
  - 公派赴印尼大学担任为期 6 个月的访问学者，实地调研中印尼产业合作、主权AI情况，接受《中国贸易报》专题采访。
- 广西壮族自治区外事办公室-综合协调处实习生 2025年08月 - 2025年09月
- 聚焦中国-东盟国家AI领域合作动态与政策走向，独立起草外事内参、领导人谈话参考及调研报告等决策信息产品7份，为高层外事决策提供关键参考依据；深度参与第22届中国-东盟博览会统筹工作，主导将区域国别研究成果系统转化为面向外办青年干部的专题培训内容，主讲“区域国别研究论坛”，提升内容产品化能力与公众演讲影响力。

## 相关技能

- AI 产品能力: 0-1 AI 产品全链路落地、需求洞察与产品指标定义、Agent / RAG / 架构设计、评测驱动迭代；定性深度访谈。
- Vibe Coding 能力: 熟练使用Codex/Deepseel /HarnessWorkbuddy等平台，擅长以自然语言驱动完成从原型到线上部署。
- 外语能力: 英语能够作为工作语言（CET-6 543 / IELTS 6.5）够直接与海外客户、合作方进行进行商务谈判与交流对接。