

张舒浩

邮箱:M202421730@xs.ustb.edu.cn

电话:(+86)132-3323-6973



教育背景

北京科技大学	通信工程·硕士	2024.09-至今
GPA3.6(专业前20%)	一等奖学金	
中北大学	人工智能·本科	2019.09-2023.06
GPA3.4(专业前20%)	一等奖学金 校人工智能比赛二等奖	

实习经历

北京字节跳动科技有限公司	AI 技术产品 (豆包)	2026.03-至今
--------------	--------------	------------

【产运】豆包研发风险策略平台

- 针对**豆包侧**研发流程中合规风险识别分散、治理进度不可追踪的问题，参与从0到1设计**豆包 SDLC 合规检测框架**，覆盖 PRD、技术文档、代码、MR、APK 等研发产物，形成“资源接入→工具检测→风险判定→治理跟进→指标监控”闭环。
- 梳理司内可接入风险识别工具、设计治理指标体系、协同产品、研发、安全 BP 等角色推动数据与检测能力接入，实现合规总覆盖率 **81.29%**、成熟数据接入率 **96%**；
- 针对规则匹配与复杂风险识别需求，协同算法团队设计结构化结果，基于 LangGraph，将 RAG、LLM 能力封装为可复用检测 **Agent**，单条处理耗时约 19 秒，映射准确率 **96.4%**，使所有检测结果可规范化映射，从而支持策略快速配置与迭代。累计覆盖 14 个检测工具和 271 个风险点。

【产运】DOLA 安全策略评估与优化

- 针对 DOLA 隐晦年龄表达等识别未成年难的问题，设计“数据取送—Agent 判定—case 回归—策略迭代”**自动化闭环链路**，实现周期数据的自动化分析。
- 基于误伤 Case 开展归因，提炼优化策略，持续协同 TnS 侧推动修改线上链路并优化策略落地，初步将端上判定准确率由 **88%** 提升至 **92%**。

【产运】AI 标识与元数据治理

- 响应生成式 AI 标识监管要求，针对豆包测试场景分散、执行标准及统计口径不统一等问题，主导建设多端、多模式合规测试体系；基于端级去重口径沉淀千余个**标准化测试用例**，并制定统一执行规范，搭建风险 Tracker 与量化指标，形成“测试—定位—整改—验收”闭环。
- 推动 AI 标识、元数据等系统性问题整改，整体不合规率由约 36% 降至 8% 以内，核心基础模式实现 100% 合规，并为自动化检测平台接入提供统一基线。

【技术】业务自进化客服 Agent (Agent Harness)

- 面向合规咨询、通告与工单分散的业务场景，基于开源 AI Manus 二次开发智能客服 Agent，串联知识检索、上下文记忆、任务规划与工具执行，支持政策咨询、风险查询、工单跟进和通告追踪 4 类任务。
- 设计**轻量化 Agent Harness**，基于 Plan-and-Execute 与原生 Tool Calling 编排 Browser、File、Search、MCP 等工具；以 OpenClaw 作为消息入口，通过任务级隔离沙箱执行工具，并完善会话状态管理、失败重试、权限校验、任务中断与人工接管机制。
- 构建 240 条业务评测集，围绕失败案例开展知识库、Prompt 与工具路由优化；经 3 轮迭代，端到端任务完成率由 74.2% 提升至 **86.7%**，平均处理时长由约 8 分钟缩短至 1.6 分钟，按每周 150 次请求测算节省约 16 人时/周。

【技术】业务流程自动化工具 (Skill / Bot)

- 隐私政策 Skill：针对多语言隐私政策回填效率低、格式易错的问题，设计“文档解析—自动写入—格式校验”链路并封装为可复用 **Agent Skill**，支持 20 种语言高保真迁移，将单次处理时间缩短至约 3 分钟，效率提升约 20 倍。
- 信息抓取 Bot：针对合规通知分散、人工收集与跟进易遗漏的问题，基于 OpenClaw 搭建合规信息自动化 Bot，实现通知抓取、结构化沉淀和进展跟踪，降低人工收集与跟进遗漏风险。

北京快手科技有限公司	安全技术工程师	2025.10-2026.02
------------	---------	-----------------

【产运】相册 SDK 统一接入与策略优化

- 针对原相册 SDK 合规能力不足及业务调用分散的问题，设计 API 准入、权限告知、频控、防劣化与场景化配置机制，推动相册能力统一收口；完成 2 个 SDK 整改验收，覆盖 11 个部门、247 处代码调用及 106 个业务场景，降低监管风险

【产运】App 隐私问题治理与整改跟进

- 熟练掌握隐私合规要求，结合自动化报告、日志分析、人工测试及 Frida 动态检测开展 App 隐私合规治理，发现 76 项问题并推动闭环 57 项，问题闭环率约 **75%**；并将敏感 API 监测、规则识别和用例生成能力沉淀为可复用方案。

【技术】隐私评估 Agent 与合规通用 Prompt

- 针对隐私评估工单依赖人工、复杂场景判断口径不一的问题，设计“需求识别—要素抽取—完整性校验—场景分流—风险聚合”**Agent 决策链路**；固定场景匹配规则库，复杂场景按“字段×行为”循环评估，自动输出风险结论与整改建议。
- 将隐私合规要求整理成可复用的 AI 提示词，根据不同应用端和 6 类业务场景，按统一模板生成合规检查用例，并通过规则约束减少 AI 编造内容的情况。

【技术】权限调用信息采集

- 复用现有 Frida 检测工具及采集代码，根据隐私核验需求确定重点关注的权限调用与采集范围，制定**定向抓取策略**，提取后台报告中的目标调用信息，为业务场景核验和隐私问题定位提供依据。

北京经纬恒润科技有限公司	AI 应用开发	2025.07-2025.09
--------------	---------	-----------------

【技术】企业问答系统(RAG 平台)

- 针对企业文档分散、信息检索困难及数据需本地可控的问题，基于 Python、LangChain、Chroma 和本地大模型搭建问答系统，实现文档处理、混合检索、答案生成与来源追溯，并通过 Streamlit 实现问答与知识库管理，完成本地部署。
- 构建 50 条**问答评测集**并人工标注正确证据，分析检索与回答中的错误，优化文档分块、混合检索及提示词；使 Top-5 正确证据命中数由 36 条提升至 44 条，命中率由 72% 提升至 **88%**

科研经历

基于 OAM 的水下可见光通信	AI&通信	2024.09-至今
-----------------	-------	------------

- 搭建融合 AI 的端到端水下涡旋光通信系统：仿真端提出相位恢复方法并构建高质量光束数据集，建立可控海洋湍流信道，设计融合注意力机制与 Transformer 的解调网络，实现受扰信号的高精度恢复；系统端搭建基于 SLM 与相机的收发平台，实现光束调制、同步采集与智能解调，通过采采数据微调及水下传输实验，完成算法从仿真环境到真实场景的适配与验证。

专业技能

产品运营	需求分析、流程设计、指标拆解、策略效果评估、跨团队协作
技术与工具	Python、MySQL、LangChain / LangGraph、RAG、Agent、OpenClaw
语言能力	CET-6

自我评价

聚焦 AI 与业务治理交叉方向，擅长从产品化视角拆解业务问题，可将业务规则、数据指标和研发流程抽象为可落地策略，擅长 Agent/Skill/Bot 等自动化提效链路搭建；做事认真负责、抗压能力强，能快速适应新平台与新业务并独立推进落地。