



🎓 教育经历

东南大学 (985) - 交通运输工程 硕士 专业排名前10% (16/282) 2024.09 - 2027.06
2025年度研究生国家奖学金、2021年度本科生国家奖学金
东南大学 (985) - 交通工程 本科 学院排名前10% (9/93) 2020.09 - 2024.06
2020年全国大学生英语竞赛国家级二等奖、2022年全国大学生数学建模大赛江苏省二等奖

💼 实习经历

华为技术有限公司 终端云服务部 AI应用开发 2026.05 - 2026.08

工作背景: 参与面向华为开发者技术支持场景知识生产与管理平台的功能开发, 实现工单到知识自动化转换与全生命周期管理

主要职责:

- **重构工单转知识为Agent驱动:** 针对原先python硬编排单阶段调用, token消耗大且效率低的问题, 将流程重构为LLM两阶段加python中间处理的ReAct Agent架构, 并将判定与生成分离, 降低约60%的token消耗, 每日工单处理耗时缩减约50%
- **优化知识质量评估流水线:** 针对工具检测、语义评估、代码规则、代码审查与编译验证等异构检查串行耗时长, 单阶段失败影响整体结果问题, 设计基于协程并行的流水线将确定性检查与LLM评估并行化, 每阶段独立降级策略隔离故障传播, 单条知识评估耗时从18-45秒缩短至10-25秒
- **设计并实现Skill自演化系统:** 针对人工优化Skill的低效率难闭环问题, 设计基于审核意见驱动的Map-Reduce流程, 涉及LLM自动定位缺陷并生成结构化修改意见、Skill版本管理、数据看板等功能, 单月内提高知识内容审核通过率15%
- **实现LLM多实例池化机制:** 针对本地部署下, 单模型调用易触发限流问题, 实现多模型轮询分摊流量与单模型撞限额后自动切换下一模型的能力, 约8%的调用触发降级切换且最终成功调用

📁 项目经历

EventGuard-电商订单事件溯源与智能异常治理平台 2026.03 - 2026.08

项目描述: 面向电商订单全生命周期的追溯与异常治理场景, 区别于仅保存最终状态的传统订单平台, 将每次状态变更沉淀为可回放、可审计的事件, 并结合实时异常检测、告警分析、人工补偿和运营查询, 实现从订单追溯到异常处置的闭环

- **重构订单写入为事件溯源:** 针对状态覆盖、并发冲突和重复入账, 以不可变事件记录状态, 采用唯一约束、版本校验和重试保证安全; 通过commandId和command_log去重, 每100个事件生成快照。重复提交不新增事件, 重建仅回放增量数据
- **完善读写分离异步读模型:** 针对投影重复、乱序和数据滞后, 引入消费去重表、版本守卫、投影通知、50毫秒兜底轮询和2秒超时机制; 重复消息不重复更新, 乱序不覆盖新状态
- **搭建数据库变更捕获链路:** 针对数据库与消息发布双写不原子, 通过Debezium读取PostgreSQL WAL并写入Kafka, 消费端采用手动提交、3次重试和死信队列; 数据库宕机或Kafka暂停后均在2秒内恢复, 已提交事件零丢失
- **设计可降级的智能查询与处置闭环:** 针对LLM生成查询语句的注入风险、超时和告警无处置, 采用意图分类、参数抽取和模板接口, 8秒超时降级摘要; 8个查询场景结果均正确, 支付重试超限与库存失败两类补偿均成功, 高风险操作进入人工审批

Synapse-可观测、可扩展、可评测的Code Agent Harness 2026.03 - 2026.08

项目描述: 面向代码仓库长链路任务构建本地Code Agent Harness, 作为模型与真实开发环境之间的运行时控制层, 统一治理上下文、记忆、工具执行和结果验证, 解决多轮执行中的信息膨胀、状态丢失、越界操作及结果不可验证问题

- **设计可扩展Agent Harness运行时:** 为解耦模型、规划器与工具并降低扩展、排障成本, 划分可替换组件边界, 统一管理Agent、Session与事件生命周期; 接入6家Provider、4类Planner、2种MCP传输, 扩展无需改动Agent主循环
- **完善长任务上下文治理机制:** 为避免历史消息和无关文件挤占预算、仓库内容重复读取, 采用Git-aware检索、符号提取、相关性排序、四区预算与摘要式compaction; 按配置的Token预算注入内容, 检索故障默认最多阻塞10s
- **构建分层记忆与任务恢复体系:** 为解决跨轮事实丢失与中断重跑, 采用原子持久化、/resume续接与Git快照回滚; 状态拆为Session、Project、User、Semantic四类, 会话跨进程恢复, 项目规则与用户偏好支持跨任务复用
- **建立工具调用安全与执行治理机制:** 因静态Tool风险声明无法识别参数副作用, 采用action-time鉴权, 按工作区、shlex命令链、重定向与SSRF判险; 覆盖5级风险决策、14类危险命令模式和12类敏感路径, 高风险操作执行前拦截或确认
- **建立冻结数据与外部grader评测链:** 针对模型自报成功不可信, 以completion gate、隔离workspace与重复统计; SWE-bench冻结20题gold子集三模型×3评测, success@3达40%-55%、误报率58%-77%

🔧 专业技能

- **Java与后端开发:** 熟悉Java语言与面向对象编程, 理解常用集合的底层实现及源码、多线程与JVM核心机制; 理解IoC、AOP设计思想, 能够使用Spring Boot、MyBatis开发后端服务
- **数据库与中间件:** 熟悉MySQL表结构与索引设计, 了解存储引擎、事务机制及SQL优化; 熟悉Redis数据结构与缓存应用, 了解分布式锁、主从复制等机制; 能够使用Kafka实现异步处理、系统解耦与消息传递
- **AI应用开发:** 理解LLM核心概念, 熟悉Prompt Engineering、RAG、向量数据库及检索优化; 能够使用LangGraph等框架进行Workflow编排与Agent开发, 理解Agent核心机制及主流Harness设计; 熟练使用Coding Agent辅助研发
- **工程化与部署:** 熟悉Git、Linux常用命令与Docker部署实践; 了解Prometheus、Grafana监控体系及ELK、Loki日志采集