

王鸿

求职意向：大模型应用 (Agent)

电话：(+86) 18627061195 | 邮箱：wh_5233@163.com | 微信：whang5233

个人主页：wh5233.me | GitHub：github.com/wanghong5233



教育背景

2024.09-2027.06

苏州大学(211)

人工智能硕士

2019.09-2024.06

武汉科技大学

软件工程本科

实习经历

上海启湾智驰科技有限公司——AI 全栈工程师

2026.05-至今

工作职责：在 AI 漫剧生产中，按需求从零到一负责剧本解析微服务的后端开发与迭代，覆盖小说选材推荐、长文本结构化、带证据报告、Agent 改写与分镜脚本生成，并完成与业务主站的报告展示及 Agent 联调。

主要贡献：

- 剧本解析从零到一：**独立完成可运行全栈原型，演进为独立部署与存储的剧本解析微服务，并在业务主站打通报告展示、任务轮询、证据定位与 Agent 可审阅写回；结构、情绪、悬疑、投流等分析通道并行，单通道失败降级，不拖死整次任务。
- 小说选材推荐与分镜脚本生成：**建设小说原材料库与离线画像批处理管线（幂等、可续跑、单本失败隔离），对百万字级长篇分块浅读（Map-Reduce 归并）形成推荐画像，完成题材、走势与改编意图三路召回的选材推荐，以及分镜脚本生成。
- 集场切分与人物解析：**针对半结构剧本切分不稳、人物易漏抽/易污染，以确定性格式识别与自声明对账为主、LLM 兜底为辅，并保留切分诊断；人物与大纲改全文单次抽取并共享实体 ID，关系摘要由关系边派生；按任务标定输入覆盖与上下文/输出预算，避免固定阈值静默少喂，局部失败显式暴露、禁止空值伪装成功。
- 高光识别与曲线解析稳定性：**以目标达成、观众悬疑、观众情绪等曲线刻画剧情起伏，按波峰波谷与突变识别高光候选，并支持人工勾选写入校准账本，服务投流选段；针对同剧重跑曲线漂移、结果不可比，验证降温度与多模型投票无法消除定义漂移后，改用固定模板受控填槽与确定性归约，并以稳定性/正确性双轴探针与 gold 锚点回归；当前可支撑报告消费，客观性与稳定性仍按真实剧本迭代。
- 剧本专家 Agent：**建设垂直域剧本助手（领域人格、工具分工、领域知识外置），打通报告结论、爆款视频到剧本的结构迁移与 LiteXML Diff 写回；超长剧本先给场次索引、再按需读正文，同会话靠对话历史连续；工具写回不信任模型自述，修改前后对比文档状态，无变化则显式失败；共享编辑器上多 Agent 并行会争抢节点 id 并丢失 Diff，改为单 Agent 批量修改。
- 长任务异步化与可恢复执行：**针对分钟级剧本分析在同步 HTTP 下超时、丢状态，改为 PostgreSQL 持久化任务与 Serverless Worker，上传即返回任务标识；以原子 claim、租约、有限重试、repair 与幂等写入处理超时、重复触发与卡死，进度与失败可查可恢复；入库与分析状态分离，避免分析失败覆盖已入库剧本。

工作成果：上述能力已上线，供短剧业务方真实使用并快速反馈迭代，再逐步扩展至其他使用团队。

项目经历

RivalLens——Multi-Agent 竞品分析与追踪系统

2026.05.23-2026.06.23

项目概述：字节跳动 AI 全栈项目挑战赛作品。构建 Multi-Agent 竞品分析与追踪系统，将一句自然语言需求转为可执行调研任务，覆盖竞品发现、证据采集、跨竞品分析与带引用报告生成。担任队长，主导架构设计与核心实现，约 60 支队伍中入围 30 强答辩。

核心技术：

- Multi-Agent 动态编排：**基于 LangGraph 实现 Intake / Planner / Supervisor / Researcher / Analyst / Writer / QA 闭环；Supervisor 以 LLM tool calling 在运行时决定竞品发现、并行度、补调研与终止，而非固定流水线；需求澄清与计划确认采用 interrupt + PostgreSQL checkpoint，支持 HITL 暂停与跨进程恢复。
- 结构化协议与证据质检：**Agent 间以 Pydantic 强 Schema 传递产品画像、功能、定价、人群与结论矩阵，降低自然语言接力漂移；结论强制绑定可核验来源，QA 以规则 DSL + LLM 语义双层审查无证据 / 证据不支撑 claim 的输出，按 reject_to 精准打回上游，设最大打回次数防止死循环。
- 可观测与调用工程化：**SSE Live + Trace DAG 回放决策与 token；LLM 按任务槽位分档路由，配合并发信号量、TPM 令牌桶、Retry-After/退避与 wall-clock 重试预算，保障 Multi-Agent 并发稳定执行。

项目概述: 面向服饰电商尺码犹豫导致的弃单与错码退货，独立开发并上架 Shopify App Store。在商品页 Size 旁构建 Fit Finder，将商家确认尺码表、顾客画像与买家评论转为可直接加购的尺码推荐；规则引擎负责可回放、可评测的选码决策，LLM 负责评论信号抽取、证据解释与缺失信息追问，并通过推荐加购与合身反馈实现商品级校准闭环。

Pulse——通用个人智能体助手

2026.03–2026.05

项目概述: 独立设计并实现通用个人 Agent 平台，以三层 Agent Runtime (Agent OS / Task Runtime / Memory Runtime) 为底座，统一任务调度、断点恢复、安全门控与飞书/微信多渠道交互。在此底座上落地多平台求职自动投递与技术雷达，真实求职场景持续运行，单平台日均自动触达 50+ 位 HR。

ScholarMind——面向学术研究的 AI 助手

2025.07–2026.03

项目概述: 面向学术场景中大模型易编造虚假文献、引用难溯源的问题，独立研发 AI 研究助手。以混合 RAG 检索为事实底座，集成 DeepResearch (Multi-Agent 调研) 与 Doc Studio (ReAct 文档编辑)；支持本地知识库与学术网络双源检索、精确引用定位，以及 LaTeX 长文档语义改写、编译校验与审批回滚。

科研经历

- **H. Wang**, X. Fan, Z. Cheng, et al., "Safe Multi-Agent Deep Reinforcement Learning for Privacy-Aware Edge-Device Collaborative DNN Inference," IEEE Transactions on Vehicular Technology, Early Access, 2026, doi: 10.1109/TVT.2026.3670497. (**中科院二区 Top/IF 7.1**)
- Z. Cheng, **H. Wang**, M. Liwang, et al., "Graph-Enhanced Deep Reinforcement Learning for Share-Aware Multi-Task DNN Inference in Low-Altitude Intelligent Service Networks," IEEE Transactions on Services Computing, Under Review, 2026. (**CCF-A/IF 5.8**)
- Z. Cheng, X. Xia, **H. Wang**, et al., "Privacy-Aware Joint DNN Model Deployment and Partitioning Optimization for Collaborative Edge Inference Services," IEEE Transactions on Services Computing, vol. 18, no. 5, pp. 3079-3092, 2025. (**CCF-A/IF 5.8**)

技术栈

- **编程语言与工程:** Python、TypeScript、FastAPI、asyncio、PostgreSQL、Redis、Milvus、Docker、Linux、Pydantic、WebSocket、SSE。
- **Agent / LLM 系统:** LangGraph、Agent Runtime、Multi-Agent、ReAct、Tool Calling / Function Calling、Prompt Engineering、Context Engineering、Structured Output、MCP、HITL、长期记忆、Checkpoint / Resume、Hook / Gate、Tracing、Workflow Orchestration。
- **RAG / 检索增强生成:** RAG、Embedding、Hybrid Search (BM25 + Vector)、Rerank / Cross-Encoder、RRF、MMR、Multi-Query、HyDE。
- **大模型后训练与评估:** HuggingFace Transformers、LLaMA-Factory、QLoRA、SFT、DPO、LLM-as-a-Judge、Qwen。
- **AI 辅助开发:** 熟练使用 Cursor、Codex、Claude Code 等 AI 编程工具进行方案设计、代码生成、重构与复杂 Debug。

证书

软件设计师 (中级) ; CET-6