

个人简介

郑昊，中科院软件所 23 级硕士，预计于 26 年 6 月毕业。研究方向为 LLM Agents/Agentic RL，具备预训练、微调、RLHF 全栈能力和优秀的工程实现能力。独立开发的开源项目 PPTAgent 已获 **2,900+ Stars**，并以第一作者发表论文于 EMNLP 2025

中国科学院大学 · 软件工程硕士
软件研究所 · 中文信息处理实验室

2023 年 9 月 – 2026 年 6 月
导师：孙乐、韩先培研究员

浙江工业大学 · 软件工程学士
计算机科学学院

2019 年 9 月 – 2023 年 6 月

实习经验

个性化大模型，京东探索研究院

2025 年 6 月 – 2025 年 8 月

- 业务背景**：构建具有共情与口语表达能力的个性化大模型，服务用户个性化交互场景
- 数据构造**：基于 PersonaHub 种子数据设计多阶段增强 Pipeline，产出 2 万轮覆盖知识问答、情感陪伴、闲聊等场景的高质量对话数据
- 能力增强**：训练与推理阶段集成搜索工具调用，提升模型知识时效性与工具使用能力
- 偏好对齐**：基于微调模型多次采样、去重与人工标注，构建 5k 高质量偏序数据集
- 质量评估**：基于 Qwen3-14B 训练生成式奖励模型自动评估回复质量，实现了 87% 的准确率
- 成果**：模型回复合格率 86%，相比 Qwen3-235B-A22B-Instruct (55%) **提升 31pp**

项目经验

DeepPresenter：自主探索与环境驱动反思的幻灯片智能体

2025 年 8 月 – 至今

- 研究动机**：智能体对工作流的依赖限制了探索深度和推理能力 → 提出基于 ReAct 框架的环境驱动反思推理的双智能体
- 智能体框架**：设计 Researcher-Presenter 双智能体架构，通过共享文件系统协作解耦信息研究与视觉设计；配备沙盒环境支持代码执行等开放式探索
- 环境驱动反思**：引入感知工具获取真实环境反馈（如幻灯片渲染后的图像），支撑智能体的反思推理
- 轨迹数据构建**：构建多场景、约束增强的幻灯片任务集 UltraPresent；提出过程监督机制，在智能体调用感知工具时注入外部评判，引导有效反思合成高质量轨迹

PPTAgent：面向风格复用的演示文稿生成工作流

2024 年 6 月 – 2025 年 6 月

- 研究动机**：基于模板生成幻灯片是广泛需求，但人工标注模板成本高昂，而让模型从零设计又难以保证布局合理与设计美观 → 通过多模态幻灯片分析和 CodeAct 框架实现幻灯片的编辑生成
- 生成框架**：提出两阶段 Agent 方法——第一阶段通过多模态内容理解对参考幻灯片进行聚类与内容识别；第二阶段以 Code Action 为动作空间在原始幻灯片上进行内容编辑来生成新幻灯片，并通过代码执行反馈迭代纠错确保 95%+ 的成功率
- 模型训练**：基于 GPT-4o 构建 26k Agent 轨迹数据集，设计基于 ROUGE 相似度的 rule-based reward，通过 GRPO 算法将 Qwen2.5-3B 的单次执行成功率从 37.4% 提升至 99.3%
- 成果**：PPTEval 平均得分 3.67 (基线 3.00, +22.3pp)，生成幻灯片的多样性和样式显著提升

基座模型训练

2023 年 9 月 – 2024 年 5 月

- 数据工程**：开发细粒度 CommonCrawl 处理流水线，通过分布式域名聚合与 n-gram 匹配实现 TB 级行级去重；训练 BERT 分类器筛选高质量数据（准确率 91%+），GPT-2 困惑度降低 34%
- 模型对齐**：基于 Self-Instruct 构造 50k 中文推理数据，结合 260k 开源对话/偏好数据，完成指令对齐
- 成果**：C-Eval 中文理解能力 +21%，MT-Bench 多轮对话质量 +15%，模型的知识时效性显著提升

发表论文

[1] Hao Zheng, Xinyan Guan, et al. “PPTAgent: Generating and Evaluating Presentations Beyond Text-to-Slides” EMNLP 2025 Main Conference