

楼仁泽

📞 电话: 158-6970-7203 📧 微信: Reza0247atTU ✉ 邮箱: marionojump0722@gmail.com 🌐 个人主页 📖 谷歌学术

🎓 教育背景

- 🎓 宾夕法尼亚州立大学, 美国 State College 2023.01 – 2026.06
计算机科学与工程 | 博士研究生在读 | 硕博连读 (已获硕士学位)
- 🎓 浙江大学城市学院, 中国杭州 2018.09 – 2022.06
计算机软件工程 | 全日制本科

🔬 研究方向

大语言模型 | 指令微调 | 后训练 | RLHF | GUI 网页智能体 | 工具调用智能体 | 代码智能体 | AI4Science

📁 实习经历

算法研究员 @ 微软研究院 🇺🇸 (美国 Redmond) 2025.06 - 2026.01

从事基于工具调用 (Tool-use / MCP) 的网页 GUI 智能体研究, 面向复杂网页环境提升自动化执行效率与鲁棒性。围绕网页智能体的工具构造、设计与复用展开两方面工作:

- 1) 提出一套面向网页场景的智能体系统, 使智能体能够高效探索网页并自动构造可复用工具 API; 通过构建网页元素哈希映射 (Element Hash) 与 UI 向量库, 在工具复用阶段显著增强了系统对动态网页环境和 UI 更新的适应能力。
- 2) 基于现有 Tool-use 网页智能体, 收集并整理多源异构工具, 在大规模基础模型与多类网站场景上开展系统性实证分析, 研究工具引入对网页智能体性能、效率与稳定性的影响, 并总结形成一套面向未来网页智能体的工具设计与构造经验准则。

业绩: 以第一作者身份完成论文一篇, 顶会投稿中。

算法研究员 @ Salesforce 研究院 🇺🇸 (美国 Palo Alto) 2024.05 - 2024.08

面向大规模代码仓库场景, 研究编码智能体在软件工程任务中的代码理解、检索与修复能力, 提升其在 SWE-bench 基准上的问题求解效果, 包括 1) 多智能体编码框架: 提出面向软件缺陷修复的多智能体 (Multi-Agent) 框架, 将 SDE Bug Fixing 任务拆解为“缺陷定位”和“缺陷修复”两个阶段。2) 缺陷定位: 结合 AST、Tree-sitter 等代码解析工具, 将代码仓库组织为紧凑的树状结构; 随后由 Search Agent 联合理解仓库代码与缺陷描述, 检索与问题相关的候选代码片段作为上下文输入。3) 缺陷修复: 在候选代码片段基础上进一步补充相关代码上下文, 并结合测试样例构造完善功能依赖信息, 使智能体能够在充分理解上下文的前提下, 对目标代码进行更精确的诊断与修复, 同时减少对原有功能的破坏。

业绩: 在 SWE-bench 上相较于当时基线智能体有明显性能提升。部分核心算法已应用于 Salesforce Coding Copilot 代码智能体产品, 为 Copilot 增加了一项重要功能。

🔧 科研成果 (引用次数: 1,853 | h-index: 12)

一作论文 7 篇, 其中 CCF-A 类 2 篇, CCF-B 类 3 篇; 另有多篇论文在审。
研究成果发表于 ICLR, ICML, ACL, EMNLP, AAAI, Computational Linguistics 等国际顶级会议与期刊。

2025 & 2026:

- **Renze Lou**, Baolin Peng, Wenlin Yao, Qianhui Wu, Hao Cheng, Suman Nath, Wenpeng Yin, and Jianfeng Gao. The Tool Illusion: Rethinking Tool Use in Web Agents. *Under review at COLM 2026*.
- Zhihan Zhang, **Renze Lou**, Fangkai Jiao, Wenpeng Yin, and Meng Jiang. Advancing Language Models through Instruction Tuning: Recent Progress and Challenges. *Tutorial @ EMNLP 2025*.
- Xiaojie Gu, Ziyang Huang, Weicong Hong, Jian Xie, **Renze Lou**, and Kai Zhang. Surface Compliance or Genuine Memory Change? Rethinking Evaluation and Conflict in Knowledge Editing. *arXiv 2025*.
- Yusen Zhang, Wenliang Zheng, ..., **Renze Lou**, Wenpeng Yin, Rui Zhang. HRScene: How Far Are VLMs from Effective High-Resolution Image Understanding? *ICCV 2025*.

2024:

- **Renze Lou**, Hanzi Xu, Sijia Wang, Jiangshu Du, Ryo Kamoi, and et al. AAAR-1.0: Assessing AI's Potential to Assist Research. *ICML 2025*. (Best Paper @ AAAI 2025 AI4Research Workshop)
- Jiangshu Du, Yibo Wang, Wenting Zhao, Zhongfen Deng, Shuaiqi Liu, **Renze Lou**, and et al. LLMs Assist NLP Researchers: Critique Paper (Meta-)Reviewing. *EMNLP 2024*.
- Hanzi Xu, **Renze Lou**, Jiangshu Du, Vahid Mahzoon, and et al. LLMs' Classification Performance is Overclaimed. *arXiv Preprint 2024*.
- Kexun Zhang, Weiran Yao, Zuxin Liu, Yihao Feng, Zhiwei Liu, Tian Lan, Lei Li, **Renze Lou**, and et al. Diversity Empowers Intelligence: Integrating Expertise of Software Engineering Agents. *ICLR 2025*.
- Xi Li, Yusen Zhang, **Renze Lou**, and et al. Chain-of-Scrutiny: Detecting Backdoor Attacks for Large Language Models. *ACL-Findings 2025*.
- Ryo Kamoi, Sarkar Snigdha Sarathi Das, **Renze Lou**, and et al. Evaluating LLMs at Detecting Errors in LLM Responses. *COLM 2024*.
- Jian Xie*, Kai Zhang*, Jiangjie Chen, Tinghui Zhu, **Renze Lou**, Yuandong Tian, Yanghua Xiao, Yu Su. TravelBench: Toward Real-world Planning for Language Agents. *ICML 2024*. (Spotlight, 3.5%)

2023:

- **Renze Lou**, Kai Zhang, Jian Xie, Yuxuan Sun, Janice Ahn, Hanzi Xu, Yu Su, Wenpeng Yin. MUFFIN: Curating Multi-Faceted Instructions for Improving Instruction-Following. *ICLR 2024*.
- Jian Xie*, Kai Zhang*, Jiangjie Chen, **Renze Lou**, Yu Su. Adaptive Chameleon or Stubborn Sloth: Unraveling the Behavior of Large Language Models in Knowledge Conflicts. *ICLR 2024*. (Spotlight, 5%)
- Lin Sun*, Kai Zhang*, Qingyuan Li*, **Renze Lou**. UMIE: Unified Multimodal Information Extraction with Instruction Tuning. *AAAI 2024*. (Oral)
- **Renze Lou**, Kai Zhang, Wenpeng Yin. Large Language Model Instruction Following: A Survey of Progresses and Challenges. *Computational Linguistics 2024*.
- **Renze Lou** and Wenpeng Yin. Toward Zero-Shot Instruction Following. *SRW@EACL 2024*.

- Janice Ahn, RISHU VERMA, **Renze Lou**, Di Liu, Rui Zhang, Wenpeng Yin. Large Language Models for Mathematical Reasoning: Progresses and Challenges. *SRW@EACL 2024*.

2021 & 2022:

- Xiaojie Gu, **Renze Lou**, Lin Sun, Shangxin Li. PAGE: A Position-Aware Graph-Based Model for Emotion Cause Entailment in Conversation. *ICASSP 2023*.
- Yuanhe Tian, **Renze Lou**, Xiangyu Pang, Lianxi Wang, Shengyi JIANG and Yan Song. Improving English-Arabic Transliteration with Phonemic Memories. *EMNLP-Findings 2022*.
- Weicheng Ma, **Renze Lou**, Kai Zhang, Lili Wang and Soroush Vosoughi. GradTS: A Gradient-Based Automatic Auxiliary Task Selection Method Based on Transformer Networks. *EMNLP 2021*.
- Weicheng Ma*, Kai Zhang*, **Renze Lou**, Lili Wang and Soroush Vosoughi. Contributions of Transformer Attention Heads in Multi- and Cross-lingual Tasks. *ACL 2021*. (Oral)
- **Renze Lou***, Yutong Wang*, Kai Zhang*, Maoyan Chen and Yujiu Yang. MORE: A Metric Learning Based Framework for Open-Domain Relation Extraction. *ICASSP 2021*.
- **Renze Lou***, Fan Zhang*, Xiaowei Zhou, Yutong Wang, Minghui Wu, Lin Sun. A Unified Representation Learning Strategy for Open Relation Extraction with Ranked List Loss. *CCL 2021*.

获奖情况

宾夕法尼亚州立大学 最佳科研助理奖 (全系前 1%)	2025
英特尔 OpenVINO 卓越技术贡献奖	2022
浙江省优秀毕业生 (前 1%)	2022
浙江省政府奖学金 (前 1%)	2021
校科研一等奖学金 (前 5%)	2021
校竞赛一等奖学金 (前 5%)	2021
校学业与创新二等奖学金 (前 5%)	2020,2021
校学业与创新一等奖学金 (前 1%)	2019

学术社区服务

宣讲会/研讨会组织者: SciCap Challenge @ IJCAI 2024; Instruction Following Tutorial @ EMNLP 2025。

领域主席 (Area Chair) : EMNLP 2025; AI4Research @ AAI 2025; SEA @ NeurIPS 2025。

审稿人 (Reviewer) : ICCV (自 2025); COLM (自 2025); ICLR (自 2024); ACL Rolling Review (自 2022); Instruction Workshop @ NeurIPS 2023; UDM @ KDD 2023; ToM @ ICML 2023; StarSEM (自 2023); AKBC 2022; NLPCC 2023; CCL 2023; IJCAI 2021; JISF。