Kunlun Zhu

✉ kunlunz2@illinois.edu | ⚙ Github(Stars: Over 100k) | ✉ Scholar(Citations: 1800) | 🏠 Personal-Web

Education

University of Illinois Urbana Chanpaign (UIUC)	Aug 2024 - May 2026

Advisor: Prof. Jiaxuan You and Prof. Heng Ji

Master of Science in Computer Science (MSCS), GPA: 4.0/4.0

Sun Yat-Sen University (SYSU)	Aug 2015 - July 2020

B.E in Computer Science, WES Major GPA: 3.82/4.0

Experience

Stanford University| Prof. James Zou & Dr. Pan Lu | Visiting Student & Research Intern — June.2025 - Present
- Research on an AI scientist for scientific discovery in the Microbiome Biomedical Area.
- Lead on a collaboration project about the Agent's self-evolving, submitted to ICLR 2026

Argonne National Lab| Dr.Tanjin He | Remote Research Intern — May.2025 - Present
- AI scientist for material science discovery, with multimodal and memory use ability
- Help write the proposal that earned 8000$ OpenAI's credit for Agent for science research

UIUC | Prof. Jiaxuan You, Prof. Heng Ji | Master Student — Dec.2023 - Present
- Leading the OpenManus-RL team, earning 40k dollars in computing credits from AMD collaboration
- Leading the Multiagent Collaboration project, currently one paper accepted at ACL main, getting new designs inspired by Biology
- Leading the AI scientist projects: Results in one first author paper published in EMNLP main & demo, and ICML

ModelBest & Tsinghua NLP lab | Prof. Zhiyuan Liu | Research Assistant & NLP Engineer — Feb 2022 - August 2024
- Number 10th full-time employee at ModelBest startup company, now worth millions of dollars, and over 500 employees.
- Core contributor to the Tsinghua's 'QingXiaoDa': the first Tsinghua Undergraduate student AI companion assistant
- Main Participate in Open-source projects such as BMTools, XAgent, Agentverse, ChatDev, gaining over 30,000 stars in github in total
- First author or one of the main authors in 6 research papers about LLMs and their applications, having over 1000 citations in total.

Mila-Quebec AI Institute | Prof. Jian Tang | Research Assistant — July 2023 - December 2023
- Developed *DrugAgent*: A LLM-based RAG pipeline for drug discovery from scientific research papers and databases.
- One position paper about scientific agent safety get accepted at Nature Communications.

Carneigie Mellon University, Robotics Institute | Prof. Katia Sycara | Research Intern — July 2018 - July 2019
- Help improve the interpretability of the deep reinforcement agent with a object enhanced salience map.
- Improve the agent's interpretability with the influence function across training data.

Publications & Under Review

1. (*ICLR* 2026, *under review*) "Where LLM Agents Fail And How They can Learn From Failures." *K.Zhu*, Z. Liu, B. Li, M. Tian, Y. Yang, J. Zhang, P. Han, Q. Xie, F. Cui, W. Zhang, X. Ma, X. Yu, G. Ramesh, Y. Su, J. Wu, Z. Liu, P. Lu, J. Zou, J. You. *arXiv* : 2509.25370 2025.

2. (*EMNLP* 2025 *main*,*Recommend Oral*) "SafeScientist: Toward Risk-Aware Scientific Discoveries by LLM Agents." *K.Zhu*, J. Zhang, Z. Qi, N. Shang, Z. Liu, P. Han, Y. Su, H. Yu, J. You. *arXiv* : 2505.23559 2025.

3. (*ACL* 2025 *main*) "MultiAgentBench: Evaluating the Collaboration and Competition of LLM Agents."*K.Zhu*, H. Du, Z. Hong, X. Yang, S. Guo, Z. Wang, Z. Wang, C. Qian, X. Tang, H. Ji, J. You. *arXiv* : 2503.01935 2025.

4. (*ACL* 2025 *main*) "RAGEval: Scenario Specific RAG Evaluation Dataset Generation Framework."*K.Zhu*, Y. Luo, D. Xu, R. Wang, S. Yu, S. Wang, Y. Yan, Z. Liu, X. Han, Z. Liu, and M. Sun. *arXiv* : 2408.01262, 2024.

5. (*Nature Communications*) "Prioritizing Safeguarding Over Autonomy: Risks of LLM Agents for Science" X. Tang, Q. Jin, *K.Zhu*, T. Yuan, Y. Zhang*, W. Zhou, M. Qu, Y. Zhao, J. Tang, Z. Zhang, A. Cohan, Z. Lu, M. Gerstein *arXiv : 2402.04247*, 2024.

6. (*TMLR : 2024*) "Exploring format consistency for instruction tuning." S. Liang*, R. Tian*, *K.Zhu**, Y. Qin, H. Wang, X. Cong, Z. Liu, X. Liu, M. Sun. Transactions on Machine Learning Research, 2023.

7. (*ICLR : 2024 spotlight*) "Toolllm: Facilitating large language models to master 16000+ real-world apis." Y. Qin, S. Liang, Y. Ye, *K.Zhu*, L. Yan, Y. Lu, Y. Lin, X. Cong, et al.*arXiv : 2307.16789*, 2023. 900+ citations

8. (*ICLR 2026, under review*) "Which LLM Multi-Agent Protocol to Choose?" H. Du, J. Su, J. Li, L. Ding, Y. Yang, P. Han, X. Tang, *K.Zhu*†, J. You†. 2025. † corresponding authors

9. (*EMNLP Demo 2025*) "TinyScientist: A Lightweight Framework for Building Research Agents." H. Yu, K. Xuan, F. Li, *K.Zhu*, Z. Lei, J. Zhang, Z. Qi, J. You. Code 2025.

10. "Advances and Challenges in Foundation Agents: From Brain-Inspired Intelligence to Evolutionary, Collaborative, and Safe Systems." B. Liu, X. Li, J. Zhang, J. Wang, T. He, S. Hong, H. Liu, S. Zhang, K. Song, *K.Zhu*, Y. Cheng, et al. *arXiv : 2504.01990* 2025.

11. (*ICML 2025*) "ResearchTown: Simulator of Human Research Community.' H. Yu*, Z. Hong*, Z. Cheng*, *K.Zhu**, K. Xuan, J. Yao, T. Feng, J. You. *arXiv : 2412.17767*, 2024.

12. (*ACL 2023 main*) "WebCPM: Interactive Web Search for Chinese Long-form Question Answering." Y. Qin, Z. Cai, D. Jin, L. Yan, S. Liang, *K.Zhu*, Y. Lin, X. Han, N. Ding, H. Wang, et al. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, vol. 1, pp. 8968-8988, 2023. 100+ citations

13. (*ICLR 2025*) "Scaling Large-Language-Model-based Multi-Agent Collaboration", C. Qian, Z. Xie, Y. Wang, W. Liu, *K.Zhu**, Y. Dang, Z. Du, W. Chen, C. Yang, Zhiyuan Liu, Maosong Sun.*arXiv : 2406.07155*, 2024.

14. (*TMLR Survey*) "How Far Are We From AGI", T. Feng*, *K.Zhu**, C. Jin*, J. Liu*, H. Tu, Z. Cheng, G. Lin, J. You.*arXiv : 2405.10313*, 2024.

15. (*ACM Computing Surveys*) "Tool learning with foundation models." Y. Qin, S. Hu, Y. Lin, W. Chen, N. Ding, G. Cui, Z. Zeng, Y. Huang, C. Xiao, C. Han, et al.*arXiv : 2304.08354*, 2023. 200+ citations

OpenSource Projects

- *ToolBench* Benchmark for 16,000+ tool using, github 5,000+ stars
- *XAgent* State-of-the-art autonomous agent, github 8,000+ stars
- *AgentVerse* multi-agent simulation and task solving framework, github 3,300+ stars
- *ChatDEV* (*MACNET*) A multi-agent collaboration framework for software development, github 26,400+ stars
- *ResearchTown* A multi-agent platform for simulating and studying automatic AI research
- *MARBLE* Code for MultiAgentBench : Evaluating the Collaboration and Competition of LLM agents
- *OpenManus*&*OpenManus − RL* A live stream development of RL tunning for LLM agents, github 50,000+ stars

Awards & Honors

- Silver Medal in Kaggle 'Feedback Prize' NLP Competition over 1557 teams (top 2.5%)
- 2019 Top 5 teams in Amazon Alexa Hackathon workshop at Carnegie Mellon University
- 2017&2018 Out-standing Student Member Award and scholarships of CS Department, Sun Yat-Sen University

Service & Talk

- Invited Speaker at AMD Advancing AI 2025, "Developing AI Agents with AMD GPUs" conference
- Invited Speaker at Alibab Yunxi Conference Agent workshop about 'XAgent', 2024
- ICML 2025, ICLR 2025, ICLR 2024, NeurIPS 2024, 2025, ACL (ARR) 2024, etc Reviewer
- Teaching Assistant at CS107 Data Science Discovery UIUC, Fall 2024 & Spring 2025, and Fall 2025
- Serve as one of Founding Organizers of 'Foundation Agents' Organization and member of OpenBMB.