

Shihao Weng

✉ shweng@smail.nju.edu.cn

🌐 <https://wengshihao.github.io/>

📍 Singapore

🔗 wengshihao



Education

- 2024.09 – Present 📖 **Ph.D. in Software Engineering**, Nanjing University, China. Advisor: Prof. Yang Feng.
- 2022.09 – 2024.06 📖 **M.S. in Software Engineering**, Nanjing University, China.
- 2018.09 – 2022.06 📖 **B.E. in Computer Science and Technology**, Jiangnan University, China. Advisor: Prof. Hengyang Lu.

Research Experience

- Present 📖 **Visiting Ph.D. Student**, Singapore Management University, Singapore. Host: Prof. Xiaofei Xie.

Research Interests

Trustworthy AI agents; self-evolving agents; agent defense; software testing and quality assurance for AI systems.

Research Publications

- 1 **Shihao Weng**, Yang Feng, Xiaofei Xie, and Jiongchi Yu, *AgentAntibody: An adaptive immune system for defending LLM agents against prompt injection*, Preprint, 2026. arXiv: 2608.04053.
- 2 **Shihao Weng**, Yang Feng, and Xiaofei Xie, *Beyond accuracy: Policy invariance as a reliability test for LLM safety judges*, Preprint, 2026. arXiv: 2605.06161.
- 3 **Shihao Weng**, Yang Feng, Jinrui Zhang, Xiaofei Xie, Jiongchi Yu, and Jia Liu, *ARGUS: Defending LLM agents against context-aware prompt injection*, Preprint, 2026. arXiv: 2605.03378.
- 4 **Shihao Weng**, Yang Feng, Jincheng Li, Yining Yin, Zhenlun Zhang, Lyuxi Liu, and Jia Liu, “TACO: Trust assessment of large language models in coding assistance tasks,” in *Proceedings of the 48th IEEE/ACM International Conference on Software Engineering (ICSE)*, **CCF-A**, 2026.
- 5 Zhenlun Zhang, Yang Feng, **Shihao Weng**, Yining Yin, Jincheng Li, and Jia Liu, “VADA: A multicultural benchmark for value-aware data generation and alignment evaluation in LLMs,” in *Proceedings of the 48th IEEE/ACM International Conference on Software Engineering (ICSE)*, **CCF-A**, 2026.
- 6 **Shihao Weng**, Yang Feng, Jincheng Li, Yining Yin, Xiaofei Xie, and Jia Liu, “AtPatch: Debugging transformers via hot-fixing over-attention,” in *Proceedings of the 48th IEEE/ACM International Conference on Software Engineering (ICSE)*, **CCF-A**, 2026.
- 7 **Shihao Weng**, Yang Feng, Yining Yin, Zhenlun Zhang, and Baowen Xu, “Data preparation and quality for code-centric generative software engineering tasks: A systematic literature review,” *Frontiers of Computer Science*, 2025, **CCF-B**.
- 8 Yining Yin, Yang Feng, **Shihao Weng**, Xinyu Gao, Jia Liu, and Zhihong Zhao, “Lightweight probabilistic coverage metrics for efficient testing of deep neural networks,” in *Proceedings of the 16th Asia-Pacific Symposium on Internetware*, **CCF-C**, 2025.

- 9 **Shihao Weng**, Yang Feng, Yining Yin, Yuxuan Dai, Jia Liu, and Zhihong Zhao, “Seeing the invisible: Test prioritization for object detection system,” *Empirical Software Engineering (EMSE)*, vol. 29, no. 6, p. 154, 2024, **CCF-B**.
- 10 Yining Yin, Yang Feng, **Shihao Weng**, Yuan Yao, Jia Liu, and Zhihong Zhao, “Dataactive: Data fault localization for object detection systems,” in *Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA)*, **CCF-A**, 2024, pp. 895–907.
- 11 Yining Yin, Yang Feng, **Shihao Weng**, Zixi Liu, Yuan Yao, Yichi Zhang, Zhihong Zhao, and Zhenyu Chen, “Dynamic data fault localization for deep neural networks,” in *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (FSE)*, **CCF-A**, 2023, pp. 1345–1357.
- 12 **Shihao Weng**, Yang Feng, Yining Yin, and Jia Liu, “Prioritizing testing instances to enhance the robustness of object detection systems,” in *Proceedings of the 14th Asia-Pacific Symposium on Internetwork*, **CCF-C**; Distinguished Paper Award, 2023, pp. 194–204.

Awards and Achievements

Awards and Honors

- | | |
|-----------|--|
| 2023–2025 |  Outstanding Graduate Student of Nanjing University (three consecutive years). |
| 2023 |  National Scholarship for Graduate Students , awarded by the Ministry of Education of China. |
| |  Distinguished Paper Award , Internetwork 2023. |
| 2021 |  CCF Outstanding Undergraduate Student , awarded by the China Computer Federation. |
| 2020–2021 |  National Scholarship for Undergraduate Students (two consecutive years), awarded by the Ministry of Education of China. |

Competition Achievements

- | | |
|------|--|
| 2021 |  Bronze Medal , 45th ACM International Collegiate Programming Contest (ICPC), Asia Regional Contest. |
|------|--|

Academic Service

- | | |
|------|--|
| 2026 |  Reviewer , NeurIPS 2026 Workshop (TAE). |
|------|--|