Yue Liu

E-mail: charlie.liu.offduty@outlook.com

Mobile: +61 0405786122

Research interest: Responsible AI, AI Engineering, Agent Architecture, Software Engineering for AI, AI for Health, Safety, and Environment

Teaching interest: Foundational computing courses (data structures, algorithms), Programming courses, Software engineering courses, Professional research courses

EDUCATIONAL BACKGROUND

University of New South Wales, Australia (09/2020 – 09/2024)

Ph.D. in Computer Science and Engineering

China University of Petroleum (East China), China (09/2013 – 06/2020)

Master by Research in Software Engineering

Bachelor in Computer Science and Technology; English (double major)

AWARDS AND DISTINCTIONS

1. "**Early Career in Science Award**", Data61 Software & Computational Systems (2024)
2. "**University International Postgraduate Award**", UNSW (2020-2024)

WORK EXPERIENCE

1. Research Fellow (03/2025 to present, **Australian National University**)

   - Comparative analysis of duties, prohibitions, and compliance structures of Australian WHS regulations via LLM.

   - Surveying AI acceptance in HSE, contributing to the broader understanding of societal trust in AI.

   - Organising interdisciplinary workshops engaging legal, industry, and research stakeholders to disseminate findings and foster responsible AI adoption.

2. Research Engineer (09/2023 to 03/2025, **CSIRO's Data61**)

   - Propose an architectural pattern catalogue for foundation model-based agents, incorporating considerations of transparency, bias mitigation, and alignment.

   - Design and evaluate the architecture for an LLM-based tax copilot, applying responsible AI design trade-offs in collaboration with a startup (**Empathetic AI**).

   - Lead RAI pattern catalogue mapping to the EU AI Act, ISO 42001 standard, and Australia Voluntary AI Safety standard draft with **Australia National AI Centre**.

   - Co-contribution to the ESG-AI framework development collaborating with **Alphinity**.

- Technical lead for the Data Bill of Materials project collaborated with **CSIRO IM&T department**.

3. Research Technician (07/2022 to 09/2023, **CSIRO's Data61**)

   - Conduct a multivolcal literature review of risk assessment for Responsible AI, highlighting technical and organisational risk factors.

   - Develop AI risk assessment metrics, supporting explainability, fairness, and robustness evaluation strategies.

TEACHING AND SUPERVISION

1. Teaching Assistant (05/2021 to 08/2021, **University of New South Wales**)

   - Lead the tutorial of COMP6452 Blockchain Software Architecture course with ~100 students at undergraduate/postgraduate levels.

   - Provide project mentorship to 20 teams, fostering skills in developing secure and accountable distributed systems.

2. Student Supervision (research/project)

   - Pengqi Li (Master at University of New South Wales)

   - Jack Yuan (Undergraduate at Australian National University)

   - Jiahui Fu (Master at Lingnan University)

   - Huanchen Zhang (Master at Lingnan University)

   - Suhai Luo (Master at Lingnan University)

PUBLICATION   (No. of paper: 37, citations: 1801, h-index: 18)

1. **Liu, Y**., Lo, S. K., Lu, Q., Zhu, L., Zhao, D., Xu, X., ... & Whittle, J. (2025). Agent design pattern catalogue: A collection of architectural patterns for foundation model based agents. *Journal of Systems and Software*, *220*, 112278.

2. **Liu, Y**., Lu, Q., Zhu, L., & Paik, H. Y. (2024). Decentralised governance for foundation model based AI systems: exploring the role of blockchain in responsible AI. *IEEE Software*.

3. Lo, S. K., **Liu, Y**., Lu, Q., Wang, C., Xu, X., Paik, H. Y., & Zhu, L. (2022). Toward trustworthy AI: Blockchain-based architecture design for accountability and fairness of federated learning systems. *IEEE Internet of Things Journal*, *10*(4), 3276-3284.

4. Lu, Q., Zhao, D., **Liu, Y**., Zhang, H., Zhu, L., Xu, X., ... & Tan, T. (2024). Evaluating The Architecture of Large Language Model-Based Agents. Submitted to *Journal of Systems and Software*.

5. Lee, S. U., Perera, H., **Liu, Y**., Xia, B., Lu, Q., & Zhu, L. (2024). Integrating ESG and AI: A Comprehensive Responsible AI Assessment Framework. *arXiv preprint arXiv:2408.00965*.

6. Lee, S. U., Perera, H., **Liu, Y**., Xia, B., Lu, Q., Zhu, L., ... & Whittle, J. (2024). Responsible AI Question Bank: A Comprehensive Tool for AI Risk Assessment. *arXiv preprint arXiv:2408.11820.*

7. Zhang, D., Xia, B., **Liu, Y**., Xu, X., Hoang, T., Xing, Z., ... & Zhu, L. (2024, April). Privacy and copyright protection in generative AI: A lifecycle perspective. In *Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI* (pp. 92-97).

8. Lu, Q., Zhu, L., Xu, X., **Liu, Y**., Xing, Z., & Whittle, J. (2024). A taxonomy of foundation model based systems through the lens of software architecture. In *Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI* (pp. 1-6).

9. Lee, S. U., Perera, H., Xia, B., **Liu, Y**., Lu, Q., Zhu, L., ... & Whittle, J. (2024). QB4AIRA: A Question Bank for Responsible AI Risk Assessment. *IEEE Software.*

10. Xia, B., Lu, Q., Zhu, L., Lee, S. U., **Liu, Y**., & Xing, Z. (2024, April). Towards a responsible ai metrics catalogue: A collection of metrics for ai accountability. In *Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI* (pp. 100-111).

11. Xia, B., Lu, Q., Perera, H., Zhu, L., Xing, Z., **Liu, Y**., & Whittle, J. (2023). Towards Concrete and Connected AI Risk Assessment (C2AIRA): A Systematic Mapping Study. In *2023 IEEE/ACM 2nd International Conference on AI Engineering–Software Engineering for AI* (CAIN).

For the full publication list please see: https://scholar.google.com.au/citations?user=cG34KO4AAAAJ

PROFESSIONAL SERVICES

1. ICSE 2025, 3rd International Workshop on Responsible AI Engineering, PC member

2. ICSA 2025, 4th International Workshop on Software Architecture and Machine Learning, PC member

3. IEEE S&P 2025, Human-Machine Intelligence for Security Analytics, PC member

4. IEEE ISPA 2024, Main track, PC member

5. IEEE Blockchain 2024, Main track, PC member

6. ICSE 2024, 2rd International Workshop on Responsible AI Engineering, PC member

7. IEEE Blockchain 2023, Main track, PC member

8. ACM Computing Surveys, Invited reviewer

9. Journal of Systems & Software, Invited reviewer

10. IEEE Transactions on Software Engineering, Invited reviewer

11. IEEE Transactions on Computational Social Systems, Invited reviewer

12. IEEE Transactions on Intelligent Vehicles, Invited reviewer

For the full professional service list please see: https://yue-charlie-liu.github.io//news/