Jiawei Xu

Ph.D. Student in Information Studies

School of Information, The University of Texas at Austin

✉ jiaweixu@utexas.edu 🌐 jiaweixu98.github.io

EDUCATION

Ph.D. Student in Information Studies   Aug. 2023–Present
The University of Texas at Austin
Focus: Language Model Agents, AI for Health, Science of Science

M.S. in Information Science   Sep. 2021–Jul. 2023
Peking University
Award: Excellent Graduate of Beijing

B.Sc. in Information Management & Systems   Sep. 2017–Jul. 2021
Peking University

PUBLICATIONS

Journal Articles

1. **Xu, J.**, Zheng, Z., Min, C., Huang, W., & Bu, Y*. (2025). Knowledge Integration and Diffusion Structures of Interdisciplinary Research: A Large-Scale Analysis Based on Propensity Score Matching. *Journal of the Association for Information Science and Technology (JASIST).* [Code] [PDF]

2. Xu, J., Yu, C., **Xu, J.**, Torvik, V. I., Kang, J., Sung, M., Song, M., Bu, Y., & Ding, Y*. (2025). PubMed Knowledge Graph 2.0: Connecting Papers, Patents, and Clinical Trials in Biomedical Science. *Scientific Data.*

Conference Papers

3. **Xu, J.**, Chen, J., Ye, Y., Sembay, Z., Thaker, S., Payne-Foster, P., Chen, J., & Ding, Y*. (2025). Interactive Graph Visualization and Teaming Recommendation in an Interdisciplinary Project's Talent Knowledge Graph. *ASIS&T 2025 (88th Annual Meeting of the Association for Information Science and Technology).* [Demo] *(Acceptance rate: 28%)*

4. **Xu, J.**, Xie, Q., Liu, M., Sembay, Z., Thaker, S., Payne-Foster, P., Chen, J. Y., & Ding, Y*. (2024). Decoding Patterns of Data Generation Teams for Clinical and Scientific Success: Insights from the Bridge2AI Talent Knowledge Graph. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL). (Acceptance rate: 30%)*

5. Pandit, S., **Xu, J.**, Hong, J., Wang, Z., Chen, T., Xu, K., & Ding, Y*. (2025). MedHallu: A Comprehensive Benchmark for Detecting Medical Hallucinations in Large Language Models. *EMNLP 2025 Main Conference (Empirical Methods in Natural Language Processing).* [Link] *(Acceptance rate: 22%)*

6. Cox, K., **Xu, J.**, Han, Y., Xu, R., Chen, T., Gerych, W., & Ding, Y*. (2025). Mapping from Meaning: Addressing the Miscalibration of Prompt-Sensitive Language Models. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI).* [Code] [Link] *(Acceptance rate: 23%)*

7. Cox, K.*, Qu, G.*, Hsu, C.-Y., **Xu, J.**, Zhou, Y., Tan, Z., Hu, M., Chen, T., Hu, Z., Zhao, Z.†, & Ding, Y.† (2025). Thought Graph: Balancing specificity and uncertainty in LLM-based gene set annotation. *Proceedings of IEEE ICHI 2025 (International Conference on Healthcare Informatics).*

8. Naeem, A., Li, T., Liao, H.-R., **Xu, J.**, Mathew, A. M., Zhu, Z., Tan, Z., Jaiswal, A. K., Salibian, R. A., Hu, Z., Chen, T., & Ding, Y*. (2024). Path-RAG: Knowledge-Guided Key Region Retrieval for Open-ended Pathology Visual Question Answering. In *Proceedings of the Machine Learning for Health (ML4H).* [Code] [Link]

9. Hsu, C.-Y., Cox, K., **Xu, J.**, Tan, Z., Zhai, T., Hu, M., Pratt, D., Chen, T., Hu, Z., & Ding, Y*. (2024). Thought Graph: Generating Thought Process for Biological Reasoning. In *Companion Proceedings of the ACM Web Conference 2024 (WWW'24 Companion).* [Link]

Workshop Papers

10. **Xu, J.** (co-first), Raza, Z. (co-first), Lim, T., Boddy, L., Mery, C., Well, A., & Ding, Y*. (2025). LLM-TA: An LLM-Enhanced Thematic Analysis Pipeline for Transcripts from Parents of Children with Congenital Heart Disease. *GenAI4Health Workshop at AAAI.* [Code] [Link]

11. **Xu, J.**, Ding, Y., & Bu, Y*. (2024). Position: Open and Closed Large Language Models in Healthcare. *GenAI4Health Workshop at NeurIPS 2024.* [Link]

12. **Xu, J.**, Sembay, Z., Thaker, S., Payne-Foster, P., Chen, J. Y., & Ding, Y*. (2024). Demo: Interactive Visualization of Semantic Relationships in a Biomedical Project's Talent Knowledge Graph. *GenAI4Health Workshop at NeurIPS 2024.* [Link] [Interactive Demo Link]

Under Review & Preprints

13. **Xu, J.**, Lee, Y., Youssef, A. E., Yun, E., Huang, T., Guo, T., Saber, H., Ying, R., & Ding, Y*. (2025). Beyond Feature Importance: Feature Interactions in Predicting Post-Stroke Rigidity with Graph Explainable AI. *arXiv preprint.* [arXiv]

14. Xu, H., Yi, S., Lim, T., **Xu, J.**, Well, A., Mery, C., Zhang, A., Zhang, Y., Ji, H., Pingali, K., Leng, Y., & Ding, Y*. (2025). TAMA: A Human-AI Collaborative Thematic Analysis Framework Using Multi-Agent LLMs for Clinical Interviews. *arXiv preprint.* [arXiv]

EMPLOYMENT

Graduate Research Assistant Aug. 2023–Present
School of Information, The University of Texas at Austin

Applied Scientist Intern May 2025–Nov. 2025
Amazon Search

PROJECTS

MATRIX: Multi-Agent Teaming Recommendation via Interactive Expertise Gap Identification
Prof. Ying Ding & Prof. Jiliang Tang Apr. 2024–Present

- A paper: ***MATRIX: Multi-Agent Teaming Recommendation through Interactive Expertise Gap Identification***
- Developing an LLM Agent-based RAG system for scientific teaming recommendations and question answering. *(Available at: https://cm4aiteaming.streamlit.app/)*
- Curating a benchmark dataset for teaming evaluation using large-scale scientific collaboration data

Cell Maps for AI (CM4AI) Data Generation Project
Prof. Ying Ding Aug. 2023–Present

- Built a knowledge graph of CM4AI project contributors.
- Visualized the knowledge graph using Svelte and Pixi.js. *(Available at: https://cm4aikg.vercel.app/)*

INVITED TALKS AND TEACHING EXPERIENCE

CM4AI Talent Knowledge Graph
Virtual Talk Oct. 2024; Apr. 2025
Beyond Correlation: What Factors Influence Scientific Performance
Virtual Talk Aug. 2023
Guest Lecture: LLM Prompt Engineering and Multi-Agent Systems
The University of Texas at Austin 2024–2025
Teaching Assistant: Data Visualization
Peking University 2021–2023
Teaching Assistant: Complex Networks
Peking University 2022

SERVICE

Reviewer
Journals

- ACM Computing Surveys
- Journal of the Association for Information Science and Technology
- Nature Humanities & Social Sciences Communications
- Online Information Review
- Data Intelligence

- Information & Culture
- Journal of Clinical and Translational Science
- Scientometrics

Conferences

- AAAI 2026 Main Conference
- AAAI 2026 AI Alignment Track

Workshop Organizer 2025
GenAI4Health Workshop at NeurIPS 2025

Student Volunteer 2024
NeurIPS 2024 GenAI4Health Workshop, Vancouver, Canada

PROFESSIONAL SKILLS

Coding	Python (PyTorch, Hugging Face, Streamlit), JavaScript, Node.js, Svelte
Design & Visualization	LaTeX, Photoshop, Premiere, PowerPoint
Languages	English (Professional), Mandarin (Native)