Jiawei Xu

PhD 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 Agent, AI for Healthcare, 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
Accolade: Awarded direct admission to Master's program

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.* [Code] [PDF]
2. Xu, J., Yu, C., **Xu, J.**, Torvik, V. I., Kang, J., Sung, M., Song, M., Yi Bu, & 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. **Accepted and will appear** at *ASIS&T 2025 (88th Annual Meeting of the Association for Information Science and Technology).* [Demo] Acceptance rate: 28%.
4. 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]
5. 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. To appear in *Proceedings of IEEE ICHI 2025 (International Conference on Healthcare Informatics).*
6. **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).*
7. 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]
8. 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

9. **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]

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

Under Review & Preprints

11. 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. *arXiv preprint.* [arXiv]
12. **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]
13. 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]

PROJECTS

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

- Preparing 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

CM4AI Talent Knowledge Graph. Oct. 2024; Apr. 2025

Beyond Correlation: What Factors Influence Scientific Performance. Aug. 2023

Guest Lecture on LLM Prompt Engineering and Multi-Agent Systems.
UT Austin 2024–2025

TEACHING EXPERIENCE

Data Visualization, Teaching Assistant
Peking University 2021–2023

Complex Networks, Teaching Assistant
Peking University 2022

EMPLOYMENT

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

Applied Scientist Intern May 2025 – Aug. 2025
Amazon

SERVICE

Reviewer
Nature Humanities & Social Sciences Communications, Journal of the Association for Information Science and Technology, Online Information Review, Data Intelligence, Information & Culture, Journal of Clinical and Translational Science, Scientometrics

Student Volunteer 2024
NeurIPS 2024 GenAI4Health Workshop, Vancouver, Canada

Workshop Organizer 2025
GenAI4Health Workshop at NeurIPS 2025

PROFESSIONAL SKILLS

Coding	Python (PyTorch, Hugging Face, Streamlit), JavaScript, Node.js, Svelte
Design & Visualization	LaTeX, Photoshop, Premiere, PowerPoint *(Designed 7+ high-quality illustrative figures for ML publications)*
Languages	English (Professional), Mandarin (Native)