

Jiawei Xu

PhD student in Information Studies, School of Information, UT Austin
jiaweixu@utexas.edu — Homepage: jiaweixu98.github.io

EDUCATION

- Ph.D. student in Information Studies, UT Austin** Aug. 2023 – Present
Focus: Language Agent, AI for Healthcare, Science of Science
- M.S. in Information Science, Peking University** Sep. 2021 – Jul. 2023
Award: Excellent Graduate of Beijing
- B.Sc. in Information Management & Systems, Peking University** Sep. 2017 – Jul. 2021
Accolade: Awarded direct admission to Master’s program

PUBLICATIONS

Published / Accepted

- 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]
- 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. **Accepted and will appear** in *Journal of the Association for Information Science and Technology*.
- 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]
- Xu, J.**, Ding, Y., & Bu, Y*. (2024). Position: Open and Closed Large Language Models in Healthcare. *GenAI4Health Workshop at NeurIPS 2024*. [Link]
- 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]
- 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)*.
- 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)*. [Code] [Link]
- 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)*.
- 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]
- Xu, J.**, Min, C., Huang, W., & Bu, Y*. (2021). Interdisciplinarity vs. Unidisciplinarity: A Structural Comparison of Multi-Generation Citations and References. In *Proceedings of the 18th International Conference on Scientometrics and Informetrics (ISSI 2021)*. [Link]
- Min, C., **Xu, J.**, Han, T., & Bu, Y*. (2021). References of References: How Far is the Knowledge Ancestry? In *2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 262–265. [Link]
- Bu, Y., **Xu, J.**, & Huang, W*. (2021). Citation-Based Quantitative Evaluations on Scientific Publications: A Literature Review on Citation-Based Impact Indicators. *Documentation, Information & Knowledge*, 38(06), 47–59+46. [Link]

Preprints / Under Review

- 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:2502.14302*

14. Xu, J., Yu, C., **Xu, J.**, Torvik, V. I., Kang, J., Sung, M., Song, M., & Ding, Y*. (2024). PubMed Knowledge Graph 2.0: Connecting Papers, Patents, and Clinical Trials in Biomedical Science. *arXiv preprint*. [Link]
15. **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. **Under review** at *ASIS&T 2025 (88th Annual Meeting of the Association for Information Science and Technology)*.
16. **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. **Under review** at *AMIA 2025 (Annual Symposium of the American Medical Informatics Association)*.
17. 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. **Under review** at *AMIA 2025*.

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://cm4aig.vercel.app/>*)

Understanding, Tracking, and Predicting Social, Emotional, and Behavioral Dynamics During Public Health Crises
Prof. Ying Ding Aug. 2023–Nov. 2023

Prof. Ying Ding Aug. 2023–Nov. 2023

- Collected and analyzed COVID-19-related Twitter data from Austin, Seattle, and nine U.S. states.
- Leveraged LIWC for emotion analysis and correlated findings with policy timelines and COVID-19 outcomes (e.g., hospitalizations, deaths).
- Applied difference-in-differences (DiD) and time-series forecasting models to predict hospitalization trends in Austin and Seattle.

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
Complex Networks, Teaching Assistant, Peking University

2021–2023
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, NeurIPS 2024 GenAI4Health Workshop, Vancouver, Canada 2024

PROFESSIONAL SKILLS

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