

Wenyu Li

✉ wenyuli2355@outlook.edu

☎ +1(412)418-2355

🌐 LinkedIn

🌐 <https://wenyuli23.github.io/>

EDUCATION

- 2023 – present 📌 **Ph.D. Computer Science** in Washington University in Saint Louis. St. Louis, MO
GPA: 3.7/4.0, Advisors: Yixin Chen, Fuhai Li
- 2021 – 2023 📌 **M.S. Computer Science** in Washington University in Saint Louis. St. Louis, MO
GPA: 3.9/4.0, thesis title: *Decipher macrophage-fibroblast-cardiomyocyte signaling interactions associated with heart failure using deep graph neural network models and single-cell RNA-seq data.*
- 2016 – 2021 📌 **B.S. Computer Science; Mathematics** University of Pittsburgh. Pittsburgh, PA
GPA: 3.6/4.0, graduated with *Magna Cum Laude* and Program Honor in Computer Science

EXPERIENCE

- May. 2026 – present 📌 **AI/ML Intern.** Apple, Inc. Seattle, WA
- Jan. 2026 – May 2026 📌 **Student Consultant.** Microsoft Corporation. St. Louis, MO
- Built an end-to-end pipeline integrating topic modeling, information extraction, and knowledge graph construction to extract insights from customer support tickets and enable automated response generation.
- Collaborated with Microsoft stakeholders by giving business strategies advice based on our proposed prototype method
- Jun. 2019 – Sept. 2019 📌 **Web Development Intern.** Bentley System, Inc. Pittsburgh, PA
Worked on Project ALIM (Asset Lifecycle Information Management); used the Web Service Gateway to communicate between interfaces and plugins

SELECTED PROJECTS

OmniCellAgent: Towards AI Co-Scientists for Scientific Discovery in Precision Medicine

- Implemented the Omic-agent for an agentic AI system enabling non-experts (e.g., wet-lab scientists) to run end-to-end scRNA-seq investigations, including data loading, QC, differential expression, pathway analysis, and therapy hypothesis generation with explainable step-by-step traces.

Retrieval-Augmented Generation (RAG) for Mealworm Plastic Degradation

- Benchmarked three RAG systems, including KG-aware methods (GraphRAG, LightRAG), on a curated corpus of mealworm biodegradation papers with domain-expert-designed Q&A and quantitative evaluation tasks.
- Implemented a vector-based RAG pipeline.
- Contributed to study conceptualization and supervision.



Graph-Augmented Language Model for Explainable Reinforcement-Guided Subgraph Reasoning

- Co-designed GALAX, a graph-augmented language model integrating multi-omic data, biomedical knowledge graphs, and LLM reasoning with a Graph Process Reward Model for reinforcement-guided, explainable subgraph reasoning in target discovery.
- Conducted Target-QA validation and ablation studies on models including RoG, SubgraphRAG, and GNN-RAG.


PUBLICATIONS

Journal Articles





- 1 F. Li, H. Zhang, D. Huang, H. Li, **W. Li**, T. Xu, Y. Chen, M. Province, and P. R. O. Payne, “AI for scientific discovery in omics data-driven precision medicine,” *en, Mo. Med.*, vol. 123, no. 1, pp. 55–66, Jan. 2026.
- 2 **W. Li**, R. Zhao, P.-T. Sun, J. Lee, O. Lu, S. Yuan, P. Sureshkumar, Y. Chen, Z. Xiao, and Y. J. Tang, “Large language models for peer review in biotechnology,” *New Biotechnology*, vol. 93, pp. 218–225, 2026, ISSN: 1871-6784. 🌐 DOI: <https://doi.org/10.1016/j.nbt.2026.03.007>.
- 3 R. Zhao, **W. Li**, J. Gonzalez-Aguirre, Y. Chen, J. S. Yuan, H. B. Pakrasi, C. Fei, S. Park, G. Roell, Y. Chen, and Y. J. Tang, “Development of multimodal ai for photobiorefineries via knowledge syntheses, transfer learning, and techno-economic analysis,” *Bioresource Technology*, vol. 442, p. 133 646, 2026, ISSN: 0960-8524. 🌐 DOI: <https://doi.org/10.1016/j.biortech.2025.133646>.

- 4 W. Li, Z. Mao, Z. Xiao, X. Liao, M. Koffas, Y. Chen, H. Ma, and Y. J. Tang, "Large language model for knowledge synthesis and AI-enhanced biomanufacturing," en, *Trends Biotechnol.*, vol. 43, no. 8, pp. 1864–1875, Aug. 2025.
- 5 B. Liu, W. Li, R. Zhao, R. R. Klauer, A. Hansen, N. Miller, Y. Chen, M. A. Blenner, and Y. J. Tang, "Leveraging retrieval-augmented generation to accelerate discoveries on mealworm larvae and plastic degradation," *Environmental Science & Technology*, vol. 59, no. 50, pp. 27 437–27 448, 2025, PMID: 41363662.  DOI: 10.1021/acs.est.5c14258. eprint: <https://doi.org/10.1021/acs.est.5c14258>.
- 6 Z. Xiao, W. Li, H. Moon, G. W. Roell, Y. Chen, and Y. J. Tang, "Generative artificial intelligence gpt-4 accelerates knowledge mining and machine learning for synthetic biology," *ACS Synthetic Biology*, vol. 12, no. 10, pp. 2973–2982, 2023, PMID: 37682043.  DOI: 10.1021/acssynbio.3c00310. eprint: <https://doi.org/10.1021/acssynbio.3c00310>.





Conference Proceedings

- 1 H. Zhang, D. Huang, W. Li, M. Province, Y. Chen, P. Payne, and F. Li, "Galax: Graph-augmented language model for explainable reinforcement-guided subgraph reasoning in precision medicine," in *International Conference on Learning Representations*, C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust, Eds., vol. 2026, 2026.  URL: <https://iclr.cc/virtual/2026/poster/10011059>.





Preprints

- 1 D. Huang, H. Li, W. Li, H. Zhang, P. Dickson, M. Zhan, J. P. Miller, C. Cruchaga, M. Province, Y. Chen, P. Payne, and F. Li, *Omnicellagent: Towards ai co-scientists for scientific discovery in precision medicine*, 2025.  DOI: 10.1101/2025.07.31.667797. eprint: <https://www.biorxiv.org/content/early/2025/08/04/2025.07.31.667797.full.pdf>.
- 2 H. Zhang, W. Li, D. Huang, Y. Tang, Y. Chen, P. Payne, and F. Li, *Kogner: A novel framework for knowledge graph distillation on biomedical named entity recognition*, 2025. arXiv: 2503.15737 [cs.CL].  URL: <https://arxiv.org/abs/2503.15737>.
- 3 H. Zhang, S. Liang, T. Xu, W. Li, D. Huang, Y. Dong, G. Li, J. P. Miller, S. P. Goedegebuure, M. Sardiello, J. Cooper, W. Buchser, P. Dickson, R. C. Fields, C. Cruchaga, Y. Chen, M. Province, P. Payne, and F. Li, *Biomedgraphica: An all-in-one platform for biomedical prior knowledge and omic signaling graph generation*, 2024.  DOI: 10.1101/2024.12.05.627020. eprint: <https://www.biorxiv.org/content/early/2024/12/09/2024.12.05.627020.full.pdf>.
- 4 W. Li, J. Feng, P. Payne, Y. Chen, and F. Li, *Decipher macrophage-fibroblast-cardiomyocyte signaling interactions associated with heart failure using deep graph neural network models and single-cell rna-seq data*, 2022.  DOI: 10.1101/2022.11.02.514928. eprint: <https://www.biorxiv.org/content/early/2022/11/04/2022.11.02.514928.full.pdf>.

SKILLS

Languages		Strong reading, writing and speaking competencies for English, Mandarin, intermediate for German.
Coding		Python, R, Java, JavaScript, SQL, XML/XSL, ...
AI/ML		Hugging Face, PyTorch, scikit-learn, AutoGen, LangChain/LangGraph, ...
Misc		Notion, Office 365, Slack, \LaTeX , ...

TEACHING

Aug. 2025 – Dec. 2025		CSE 543T Algorithms of Non-linear Optimization , WashU Taught course materials, supervised and graded student presentations and projects	(8h/week)
Aug. 2022 – Dec. 2022		CSE 514A Data Mining , WashU graded assignments and exams	(2h/week)
Jan. 2022 – Dec. 2022		CSE 412A Intro to AI , WashU graded assignments and exams	(9h/week)
Sep. 2020 – Dec. 2020		CS 449 Intro to Systems Software , University of Pittsburgh Supported course operations including office hours, proctoring, and teaching staff coordination.	(5h/week)