

Yili Shen

Ph.D. Student | Generative Modeling, Model Steering & Reliability, AI for Science

yshen7@nd.edu | (+1) 646-584-5431 | [Website](#) | [LinkedIn](#) | [GitHub](#) | [Google Scholar](#)

Education

University of Notre Dame

Ph.D. in Computer Science and Engineering; Advisor: Prof. Xiangliang Zhang

Notre Dame, IN
Aug. 2023 – May 2028 (*Expected*)

Tongji University

B.Eng. in Software Engineering (AI Concentration); GPA: 91.77/100

Tongji University, Department of Chemistry

First-Year Chemistry Major

Shanghai, China
2019 – 2023
Shanghai, China
2018 – 2019

Research Experience

University of Notre Dame, MINE Lab

Graduate Research Assistant

Notre Dame, IN
Aug. 2023 – Present

- Developed [LatentRxnFlow](#) and UniRxnAxis, direction-conditioned generative frameworks for reaction prediction that expose latent trajectories, endpoint-orthogonal steering, failure modes, and uncertainty beyond endpoint accuracy.

Caltech, Nelson Lab

Visiting Researcher - AI for Crystallography

Pasadena, CA
Feb. 2025 – May 2025

- Built an end-to-end learning workflow for microcrystal electron diffraction, from data curation and embedding to diffraction classification, improving robustness to out-of-distribution molecular crystals.

University of Massachusetts Amherst, Lin and Guan Groups

Undergraduate Research Assistant

Amherst, MA
Jul. 2021 – May 2023

- Developed graph-learning methods for many-body potential energy surfaces and subgraph representation learning.

Industry Experience

Zillow

Applied Scientist Intern

Seattle, WA
May 2026 – Present

- Identified and mitigated *never-stop*, an LLM termination failure caused by activation steering.

Baidu, Bioinformatics / NLP Department

Research Intern

Shanghai, China
Mar. 2022 – Jul. 2022

- Built a transformer for drug-response prediction from molecular graphs and multi-omics data, and developed gene-expression pipelines using protein interaction networks and knowledge graphs.

SAP, Data Intelligence & Innovation

Data Engineer Intern

Shanghai, China
Oct. 2021 – Mar. 2022

- Built automated ETL and predictive analytics pipelines for SAP HANA.

Selected Publications & Manuscripts

- Persistent Activation Steering Disrupts Language-Model Termination.** *Under review.*
- Emergent Steering Beyond Endpoint Alignment in Chemical Reaction Models.** Y Shen, K Guo, H Zhuang, X Zhang. *Under review.*
- Driving Reaction Trajectories via Latent Flow Matching.** Y Shen, X Zhang. *KDD 2026 (Oral).*
- AgentTrap: Measuring Runtime Trust Failures in Third-Party Agent Skills.** H Zhuang, H Xing, Y Zhou, Y Ma, Y Huang, Y Shen, Y Han, X Zhang. *arXiv, 2026.*
- SenseMath: Do LLMs Have Number Sense.** H Zhuang, X Wang, Y Shen, Y Cheng, X Zhang. *arXiv, 2026.*
- Towards Few-shot Chemical Reaction Outcome Prediction.** Y Shen, Y Tian, C-W Ju, O Wiest, X Zhang. *CIKM 2025.*
- Transferable FB-GNN-MBE for Potential Energy Surfaces.** S Chen, Z Wang, Y Shen, X Deng, X Cheng, C-W Ju, J Yi, G Ling, D Alhmoud, H Guan, Z Lin. *The Journal of Chemical Physics, 2026.*
- Improving Subgraph Representation Learning via Multi-View Augmentation.** Y Shen, J Yan, C-W Ju, J Yi, Z Lin, H Guan. *ICML 2022 AI4Science.*

Honors & Awards

National Scholarship (Top 0.2% Nationwide), Ministry of Education of China

2021