# Weixiang Sun

*Ph.D. Student, University of Notre Dame*

+1 574 850 3601
wsun4@nd.edu
weixiang-sun.github.io
Weixiang-Sun
Google Scholar

## Research Summary

I aim to build a science of **AI–AI interaction**. HCI and HAI study how people interact with computers and AI; as AI agents increasingly collaborate with, evaluate, and stand in for one another, I study them the way sociology studies people, from the individual **agent** to **micro**-level interaction to **macro**-level societies. A recurring lesson is that what holds for one agent does not hold for many: collaboration between agents is not free and does not always pay off [1], interaction can erode the diversity that makes a group valuable [3], and agents standing in for people flatten the variety of real populations [4]. I want to understand these dynamics well enough to make systems of many AI agents reliable and safe.

Agent How one agent plans, acts, and learns: computer-use agents in production (Li Auto), long-horizon planning [23], self-evolving agents [21], and agents that serve people in social services [13, 24].

Micro How agents shape each other: when collaboration pays off [1], how interaction shapes collective output [3], and multi-agent systems for creative and scientific work [8, 20]; as a human reference, how manipulation unfolds in conversation [2].

Macro How agent populations behave and stand in for people: how faithfully LLMs simulate individuals [4] and social roles [25], and simulating how policies play out across populations [16].

## Education

2025–present **Ph.D. in Computer Science and Engineering**, *University of Notre Dame*, South Bend, USA

2021–2025 **B.S. in Information and Computing Science**, *Northeastern University*, Shenyang, China

## Selected Publications

EMNLP'26 **Weixiang Sun**, Z. Wang, H. Huang, C. Nelson, Y. Ma, Y. Ye. *The Collaboration Tax: How Much LLM Multi-Agent Systems Pay to Coordinate.*
Shows that collaboration between LLM agents is not free: working together can cost performance relative to working alone, and the loss comes from how agents communicate rather than how well they reason.

COLM'26 **Weixiang Sun**, S. Ma, Y. Li, T. Ma, Z. Wang, C. Nelson, X. Xiao, Y. Ye. *PreScam: A Benchmark for Predicting Scam Progression from Early Conversations.*
A benchmark of real scam conversations for anticipating where an interaction is heading; current models pick up some cues but miss how manipulation escalates. Code and data.

Under review **Weixiang Sun**, D. Li, Y. Ye, J. G. Nelson. *Collaborating Without Converging: Quality and Diversity in Multi-Agent Creative Generation.*
Shows that multi-agent collaboration improves quality but pulls agents' ideas together, and explains this with opinion dynamics, pointing to communication designs that keep diversity.

Under review **Weixiang Sun**, Z. Wang, K. Shi, H. Huang, C. Nelson, Y. Li, Y. Ye. *The Average Patient: Do Simulated Psychiatric Patients Answer Like Real People?*
Shows that LLM-simulated patients answer like an "average patient", far less varied than real people, which limits what simulator-based training and evaluation can test.

## Experience

Jun–Oct 2026 **Research Intern**, *Adobe*, Seattle, USA
Led the study of quality and diversity in multi-agent creative generation [3], showing how communication topology and agent stubbornness trade one against the other.

Mar–Jun 2025 **Applied Scientist Intern**, *Li Auto Inc.*, Beijing, China
Architected the computer-use agent module of Li Auto's in-house LLM assistant, with hierarchical planning and tool-use orchestration; a zoom-in mechanism for high-resolution screens raised task success from 31% to 57%, and the module shipped to production.

2023–2025 **Research Intern**, *Lehigh University & MGH/Harvard Medical School*, Remote
Co-first author of Mora (multi-agent video generation, ∼1.6k GitHub stars), EfficientLLM, and Biomedical SAM-2; first author of Bora (text-guided biomedical video generation).

## Service & Mentorship

Reviewer NeurIPS 2025, COLM 2026, ICLR 2027

Mentorship Mentored 5 students from NUS, CUHK, Fudan, and Lanzhou University (2024–2025)

Membership IEEE Graduate Member

## Awards

2023 Tuoxin Innovation Scholarship (top 2%)

2023 Second-Class Scholarship (top 10%)

2022 Third-Class Scholarship (top 15%)

## Talks

Oct 2025 **Presenter**, *RISE AI Conference*, South Bend, USA
Presented the *Chemistry and Biology* chapters of LLMs4All on LLMs for drug discovery; panelist on AI-for-Science and Biomedicine.

## Full Publication List (first/co-first author first; * equal contribution, † corresponding)

[1] **Weixiang Sun**, Zehong Wang, Hong Huang, Colby Nelson, Yijun Ma, and Yanfang Ye. The Collaboration Tax: How Much LLM Multi-Agent Systems Pay to Coordinate. *EMNLP (Main)*, 2026.

[2] **Weixiang Sun**, Shang Ma, Yiyang Li, Tianyi Ma, Zehong Wang, Colby Nelson, Xusheng Xiao, and Yanfang Ye. PreScam: A Benchmark for Predicting Scam Progression from Early Conversations. In *COLM*, 2026.

[3] **Weixiang Sun**, Dingzeyu Li, Yanfang Ye, and John G Nelson. Collaborating Without Converging: Quality and Diversity in Multi-Agent Creative Generation. *Under review at ACL Rolling Review*, 2026.

[4] **Weixiang Sun**, Zehong Wang, Kaiwen Shi, Hong Huang, Colby Nelson, Yiyang Li, and Yanfang Ye. The Average Patient: Do Simulated Psychiatric Patients Answer Like Real People? *Under review at ACL Rolling Review*, 2026.

[5] Rong Zhou*, Zhengqing Yuan*, Zhiling Yan*, **Weixiang Sun***, Kai Zhang, Yiwei Li, Yanfang Ye, Xiang Li, Lifang He, and Lichao Sun. TTT-Unet: Enhancing U-Net with Test-Time Training Layers for Biomedical Image Segmentation. In *MIDL (Oral)*, 2026.

[6] Zhengqing Yuan*, **Weixiang Sun***, Yixin Liu*, Huichi Zhou*, Rong Zhou, Yiyang Li, Zheyuan Zhang, Wei Song, Yue Huang, Haolong Jia, et al. EfficientLLM: Efficiency in Large Language Models. *arXiv preprint arXiv:2505.13840*, 2025.

[7] Yanfang Ye*, Zheyuan Zhang*, Tianyi Ma*, Zehong Wang*, Yiyang Li*, Shifu Hou*, **Weixiang Sun***, Kaiwen Shi*, Yijun Ma, Wei Song, Ahmed Abbasi, et al. LLMs4All: A Review on Large Language Models for Research and Applications in Academic Disciplines. *arXiv preprint arXiv:2509.19580*, 2025.

[8] Zhengqing Yuan*, Yixin Liu*, Yihan Cao*, **Weixiang Sun***, Haolong Jia, Ruoxi Chen, Zhaoxu Li, Bin Lin, Li Yuan, Lifang He, et al. Mora: Enabling Generalist Video Generation via a Multi-Agent Framework. *arXiv preprint arXiv:2403.13248*, 2024.

[9] **Weixiang Sun**, Xiaocao You, Ruizhe Zheng, Zhengqing Yuan, Xiang Li, Lifang He, Quanzheng Li, and Lichao Sun. Bora: Biomedical Generalist Video Generation Model. *arXiv preprint arXiv:2407.08944*, 2024.

[10] Zhiling Yan*, **Weixiang Sun***, Rong Zhou*, Zhengqing Yuan*, Kai Zhang, Yiwei Li, Tianming Liu, Quanzheng Li, Xiang Li, Lifang He, and Lichao Sun. Biomedical SAM 2: Segment Anything in Biomedical Images and Videos. *arXiv preprint arXiv:2408.03286*, 2024.

[11] **Weixiang Sun**, Yixin Liu, Zhiling Yan, Kaidi Xu, and Lichao Sun. Medical Unlearnable Examples: Securing Medical Data from Unauthorized Training via Sparsity-Aware Local Masking. *ICML NextGenAISafety Workshop*, 2024.

[12] **Weixiang Sun**. Research and Application of Improved Neural Network Optimization Algorithm. In *Third International Symposium on Computer Engineering and Intelligent Communications (ISCEIC 2022)*, volume 12462, pages 602–606. SPIE, 2023.

[13] Yiyang Li, **Weixiang Sun**, Tianyi Ma, Kaiwen Shi, Zheyuan Zhang, and Yanfang Ye. HoosierHelp: Benchmarking LLM Agents for Social Service Navigation. *arXiv preprint arXiv:2608.09946*, 2026.

[14] Kaiwen Shi, **Weixiang Sun**, Zheyuan Zhang, Lichao Sun, Nitesh V Chawla, and Yanfang Ye. CiteAudit: You Cited It, but Did You Read It? A Benchmark for Verifying Scientific References in the LLM Era. *arXiv preprint arXiv:2602.23452*, 2026.

[15] Hong Huang, **Weixiang Sun**, Zhijian Wu, Donghuan Lu, Xian Wu, and Yefeng Zheng. IQE-CLIP: Instance-aware Query Embedding for Zero-/Few-shot Anomaly Detection in Medical Domain. *IEEE Transactions on Image Processing*, 2026.

[16] Yijun Ma, Zehong Wang, **Weixiang Sun**, Zheyuan Zhang, Kaiwen Shi, Nitesh Chawla, and Yanfang Ye. Policy4OOD: A Knowledge-Guided World Model for Policy Intervention Simulation against the Opioid Overdose Crisis. *arXiv preprint arXiv:2602.12373*, 2026.

[17] Yijun Ma, Zehong Wang, **Weixiang Sun**, and Yanfang Ye. Temporal Graph Pattern Machine. *arXiv preprint arXiv:2601.22454*, 2026.

[18] Yunhong He, Zhengqing Yuan, **Weixiang Sun**, Yiyang Li, Yixin Liu, Yanfang Ye, and Lichao Sun. Vision-MoR: Scaling Vision Transformer via Patch-Level Mixture-of-Recursions. In *AAAI*, pages 4699–4707, 2026.

[19] Tianyi Ma, Yijun Ma, Zehong Wang, **Weixiang Sun**, Ziming Li, Connor R Schmidt, Chuxu Zhang, Matthew J Webber, and Yanfang Ye. SupraBench: A Benchmark for Supramolecular Chemistry. *arXiv preprint arXiv:2606.13477*, 2026.

[20] Zehong Wang, Yijun Ma, Connor R Schmidt, Tianyi Ma, **Weixiang Sun**, Ziming Li, Xiaoguang Guo, Chuxu Zhang, Matthew J Webber, and Yanfang Ye. MDForge: Agentic Molecular Dynamics Pipeline Design under Sparse Simulator Feedback. *EMNLP (Main, Oral)*, 2026.

[21] Yijun Ma, Zehong Wang, Yiyang Li, Ziming Li, Xiaoguang Guo, **Weixiang Sun**, Chuxu Zhang, and Yanfang Ye. ProPlay: Procedural World Models for Self-Evolving LLM Agents. *arXiv preprint arXiv:2606.12780*, 2026.

[22] Zehong Wang, Xiaolong Han, Qi Yang, Xiangru Tang, Fang Wu, Xiaoguang Guo, **Weixiang Sun**, Tianyi Ma, Pietro Lio, Le Cong, et al. Molecular Representations in Implicit Functional Space via Hyper-Networks. *arXiv preprint arXiv:2601.22327*, 2026.

[23] Zehong Wang, Fang Wu, Hongru Wang, Xiangru Tang, Bolian Li, Zhenfei Yin, Yijun Ma, Yiyang Li, **Weixiang Sun**, Xiusi Chen, et al. Why Reasoning Fails to Plan: A Planning-Centric Analysis of Long-Horizon Decision Making in LLM Agents. *arXiv preprint arXiv:2601.22311*, 2026.

[24] Yiyang Li, **Weixiang Sun**, Tianyi Ma, Kaiwen Shi, Zheyuan Zhang, and Yanfang Ye. Food4All: An Agentic Framework and Benchmark for Food Resource Navigation with Adaptive User Understanding. *arXiv preprint arXiv:2510.18289*, 2025.

[25] Yue Huang, Zhengqing Yuan, Yujun Zhou, Kehan Guo, Xiangqi Wang, Haomin Zhuang, **Weixiang Sun**, Lichao Sun, Jindong Wang, Yanfang Ye, et al. Social Science Meets LLMs: How Reliable Are Large Language Models in Social Simulations? *COLM*, 2025.

[26] Zhengqing Yuan, Xiaoyu Ma, Yunhong He, Zixuan Weng, Rong Zhou, **Weixiang Sun**, Lifang He, Lichao Sun, and Yanfang Ye. Can Text-to-Video Models Generate Responsible Realistic Human Motion? *NeurIPS ResponsibleFM Workshop*, 2025.

[27] Zhiling Yan, Sifan Song, Dingjie Song, Yiwei Li, Rong Zhou, **Weixiang Sun**, Zhennong Chen, Sekeun Kim, Hui Ren, Tianming Liu, et al. SAMed-2: Selective Memory Enhanced Medical Segment Anything Model. In *MICCAI*, pages 540–550, 2025.

[28] Hang Wu and **Weixiang Sun†**. FMamba: A Highly Scalable Financial Time-Series Prediction Model. In *Proceedings of the International Conference on Digital Economy, Blockchain and Artificial Intelligence*, pages 45–49, 2024.

[29] Boming Miao, Chunxiao Li, Yao Zhu, **Weixiang Sun**, Zizhe Wang, Xiaoyi Wang, and Chuanlong Xie. AdvLogo: Adversarial Patch Attack against Object Detectors Based on Diffusion Models. *arXiv preprint arXiv:2409.07002*, 2024.