Xiyuan Yang

Computer Science Ph.D. Student @ University of Illinois Urbana-Champaign

xiyuany4@illinois.edu · (+1) 217 766 5798 · Google Scholar (Citations 500+) · GitHub

Research & Selected Experiences

- My research focuses on agentic foundation LLMs/VLMs, specifically on scalability and efficiency.
- Have implemented large scale agentic-RL pipelines with **fully async** rollout / training loops, scaling multi-turn, long-horizon **(64k+ tokens traj.)** training across **16+ nodes** on ∼**10B dense and** ∼**30B MoE** models.

Education

Ph.D. in Computer Science 2025-2030
University of Illinois Urbana-Champaign Champaign, USA
Advisor: Prof. Jingrui He

B.Eng. in Computer Science and Technology 2021-2025
Wuhan University Wuhan, China
Advisor: Prof. Mang Ye

Internship Experience

Amazon, Prime Video Group 2026
Mentor(s): Zhenyu Liao, Zhan Shi, Sheikh Sarvar Sunnyvale, USA
Project : World Models for Scalable AutoResearch Agents

Microsoft Research, Social Computing Group 2024-2025
Mentor: Fangzhao Wu Beijing, China
Project 1: Detecting Prompt Injections via Representation Engineering (in EMNLP 2025)
Project 2: Identifying Cross-lingual Inequalities of Foundation Models (in PNAS 2025)
Project 3: Defending Jailbreak Attacks by Vision Language Models (under review of Nature Communications)
Project 4: Measuring Human Contributions in AI-Generated Content (in ACL 2026)
Competition: Foundation Model based News Recommendation (Ranked 1/157 groups)

University of Chicago, Department of Computer Science 2024
Mentor: Prof. Tian Li Chicago, USA
Project: Differentially Private Distributed Optimization (under review)

Invited Presentations

Jump Trading AI Research Symposium (selected poster, ∼ 25 worldwide) 2026
Optimizer Aware Kernels for Foundation Model Training Dynamics New York City, USA

Publications

- **Code as Agent Harness** [Huggingface Daily Paper #1]
  Xuying Ning et al. (43 authors, incl. **Xiyuan Yang**)
  arXiv 2026

- **Heterogeneous Scientific Foundation Model Collaboration** [Huggingface Daily Paper #1]
  Zihao Li, Jiaru Zou, Feihao Fang, Xuying Ning, Mengting Ai, Tianxin Wei, Sirui Chen, **Xiyuan Yang**, Jingrui He
  arXiv 2026

- **Preconditioning Neural Tangent Kernel for Adaptive Optimization**
  Xiyuan Yang, Wenxuan Bao, Katherine Tieu, Jingrui He
  ICML 2026

- **Latent Collaboration in Multi-Agent Systems** [Huggingface Daily Paper #1]
  Jiaru Zou, Ruizhong Qiu, Gaotang Li, **Xiyuan Yang**, Katherine Tieu, Pan Lu, Ke Shen, Hanghang Tong, Yejin Choi, Jingrui He, James Zou, Mengdi Wang, Ling Yang
  ICML 2026

- **Measuring Human Contribution in AI-Assisted Content Generation**
  Yueqi Xie, Tao Qi, Jingwei Yi, **Xiyuan Yang**, Ryan Whalen, Junming Huang, Qian Ding, Yu Xie, Xing Xie, Fangzhao Wu
  ACL 2026

- **Defending LLMs Against Jailbreak Attacks Utilizing Cross-Modality Generalization Gap**
  Xiyuan Yang*, Chenglong Wang*, Haoyu Tang*, Yueqi Xie, Bin Zhu, Lingjuan Lyu, Mang Ye, Fangzhao Wu
  Under Review 2025

- **Uncovering Inequalities in New Knowledge Learning by Large Language Models Across Different Languages**
  Chenglong Wang, Haoyu Tang, **Xiyuan Yang**, Yueqi Xie, Jina Suh, Sunayana Sitaram, Junming Huang, Yu Xie, Pengjun Zhao, Zhaoya Gong, Xing Xie, Fangzhao Wu
  PNAS 2025

- **Defending Against Indirect Prompt Injection by Instruction Detection**
  Tongyu Wen*, Chenglong Wang*, **Xiyuan Yang***, Haoyu Tang, Yueqi Xie, Lingjuan Lyu, Zhicheng Dou, Fangzhao Wu
  EMNLP 2025

- **Differentially Private Federated Clustering with Random Rebalancing**
  Xiyuan Yang, Shengyuan Hu, Soyeon Kim, Tian Li
  Under Review 2025

- **FedAS: Bridging Inconsistency in Personalized Federated Learning**
  Xiyuan Yang, Wenke Huang, Mang Ye
  CVPR 2024

- **Dynamic Personalized Federated Learning with Adaptive Differential Privacy**
  Xiyuan Yang*, Wenke Huang*, Mang Ye
  NeurIPS 2023

- **Robust Heterogeneous Federated Learning under Data Corruption**
  Xiuwen Fang, Mang Ye, **Xiyuan Yang**
  ICCV 2023

Community Contributions

- **PFLlib**: Personalized and Distributed Optimization Library — Github Stars **2000+**
  Core Contributor

- **LatentMAS**: Latent Communication of Multi-agent System Github Stars **600+**
  Core Contributor and Co-first Author
- **Conference Review**: NeurIPS, ICML, ICLR, AAAI, AISTATS, CVPR, ICCV
- **Journal Review**: IEEE TMC, IEEE TNNLS, IEEE TIFS, IEEE TKDE, IEEE TDSC, Info. Fusion

Awards & Scholarships

- PhD Student Fellowship (first-year), University of Illinois Urbana-Champaign 2025
- Star of Tomorrow Award, Microsoft Research 2025
- Overseas Exchange and Study Award, Wuhan University 2024
- Leijun Computer Science Research Award, Wuhan University 2024

Skills

- **Machine Learning**: PyTorch, CUDA, HuggingFace Transformers, vLLM, NumPy
- **Systems & Tools**: Linux, Distributed Systems, Docker, Git, LaTeX