

Xuanming Zhang

Researcher in Humanoid Intelligence

☎ +1 (608) 259-0539 ✉ zhangxm@stanford.edu 🌐 web.stanford.edu/people/zhangxm 📧 XMZhangAI

Summary

Visiting B.S. researcher at Stanford University with SSR Fellowship, focusing on Humanoid Intelligence: Cognitive Learning, World Interaction, and Policy Simulation. Senior undergraduate at UW-Madison (ISci, GPA: 3.9/4.0) with research experience at Amazon, Tsinghua, and Peking University. Published 10+ papers at top venues (NeurIPS, ACL, ICML) with 300+ citations and h-index of 7.

Education

Stanford University

Visiting Researcher, Computer Science (SSR Fellowship)

May 2025 – Present

Stanford, CA

University of Wisconsin-Madison

B.S. in Information Science (Transfer, Honors Program)

Dec 2024 – May 2026

Madison, WI

GPA: 3.9/4.0 • CDIS, ISci#11, CS#13

Peking University

Artificial Intelligence (EECS/EIE, Withdrawn)

Sep 2022 – Dec 2024

Beijing, China

National Elite Institute of Engineering Program

Research Experience

Amazon - Trustworthy AI Research

May 2025 – Present

Research Scholar

Remote

- Working with Dr. Rahul Gupta and Dr. Jwala Dhamala on responsible AI and world interaction
- Developing evaluation frameworks for LLM deceptive behavior in workplace scenarios

UW-Madison - Deep Learning Lab

Feb 2025 – Present

Research Assistant (Advisor: Prof. Sharon Li)

Madison, WI

- Leading 3 research projects on LLM reasoning vulnerability and cognitive architectures
- Developing novel approaches for reliable AI through cognitive learning frameworks (NeurIPS 2025)

ByteDance - Seed

Sep 2024 – Feb 2025

AI Architect Intern (Advisors: Xingnan Wang, Xuebin Lin, Muyang Li)

Beijing, China

- Research solutions and deployment of Applet
- Co-Design Fourier Analysis Networks for LLM periodic pattern

Tsinghua University - CoAI Lab & Zhipu-AI

Apr 2024 – Feb 2025

Research Scientist Intern (Advisors: Prof. Minlie Huang, Prof. Hongning Wang)

Beijing, China

- Led Seeker project approaching robustness code generation for LLMs (TOSEM Under Review)
- Co-Led SocialEval project evaluating social intelligence of LLMs (ACL 2025)

Peking University - Key Lab of High Confidence Software Technologies

Jun 2023 – May 2024

Research Assistant (Advisors: Prof. Zhi Jin, Prof. Ge Li)

Beijing, China

- Co-Led EvoCodeBench project evaluating code generation for LLMs in real scenario (NeurIPS 2024)
- Co-Led MGI project evaluating data contamination for LLMs (ACL 2024)

Zen-AI

Jan 2023 – Mar 2023

SDE Intern (Advisor: Libing Yan)

Beijing, China

Selected Publications

10. MetaMind: Modeling Human Social Thoughts with Metacognitive Multi-Agent Systems X. Zhang, Y. Chen, M. Yeh, S. Li	NeurIPS 2025
9. Theoretical Proof that Generated Text in the Corpus Leads to the Collapse of Auto-regressive Language Models L. Wang, X. Shi, ..., X. Zhang, ..., H. Mei	Sci China Inf Sci (Review)
8. Human Decision-making is Susceptible to AI-driven Manipulation S. Sabour, J. Liu, ..., X. Zhang, ..., M. Huang	Nature Comm. (Review)
7. SocialEval: Evaluating Social Intelligence of Large Language Models J. Zhou*, Y. Chen*, Y. Shi*, X. Zhang*, ..., M. Huang (*Equal contribution)	ACL 2025
6. Towards Exception Safety Code Generation with Intermediate Representation Agents Framework X. Zhang, Y. Chen, Y. Yuan, M. Huang	TOSEM (Review)
5. EvoCodeBench: An Evolving Code Generation Benchmark with Domain-Specific Evaluations J. Li, G. Li, X. Zhang, Y. Dong, Z. Jin	NeurIPS 2024
4. Generalization or Memorization: Data Contamination and Trustworthy Evaluation for Large Language Models Y. Dong, X. Jiang, X. Zhang, ..., M. Yang, G. Li	ACL 2024; TPAMI (Review)
3. DevEval: A Manually-Annotated Code Generation Benchmark Aligned with Real-World Code Repositories J. Li, G. Li, ..., X. Zhang, ..., F. Huang, Y. Li	ACL 2024
2. Multicollinearity Resolution Based on Machine Learning: A Case Study of Carbon Emissions X. Zhang	ICML Workshop
1. The average human lifespan is only 18 years X. Zhang	Beijing Evening News

Key Projects

RAG4Code Achieved SOTA on CrossCodeEval with dependency-chain-retrieval for cross-file code completion	PKU 2024
Common Exception Enumeration (CEE) Curated knowledge base systematizing exception handling patterns across programming languages	THU 2024
Applet - Mobile UI Agent Lightweight intent-driven framework using Swift/CoreML for real-time multimodal interaction	ByteDance 2024
H3O - Startup (CTO) Received 500K RMB investment from XbotPark; led architecture and prototype development	2022

Technical Skills

Languages: Python, C++, Swift, Java, Shell, MATLAB

ML/AI: PyTorch, Transformers, PEFT, vLLM, CoreML

Tools: Git, Docker, LaTeX, AWS, Distributed Systems

Selected Awards

National Innovation & Entrepreneurship Excellence Award

Chinese Mathematics Competition First Prize

Mathematical Contest in Modeling Second Prize

Outstanding Student President Scholarship (Top 1)

Service & Leadership

Academic: Program Committee Member for ICLR, NeurIPS, OOPSLA, AAAI • Student Member at CCF, IEEE

Leadership: Student Union President (BNDS 2020, EIE 2022)

Certifications: National Level II Psychological Counselor (PCIA)