Xinyu(Jessica) Wang

xwang2775@wisc.edu | Personal Website | Linkedin | Google Scholar

Madison, Wisconsin, USA

Research Interests

My research centers on **Agentic AI** and **Human–Computer Interaction (HCI)**, with the goal of building **AI systems that can act, reason, and collaborate in the real world**. I study how **agents use tools**, **interact with humans**, and operate within **complex socio-technical environments**. Beyond improving **model capability**, I also focus on designing **human-centered interaction paradigms** that make AI systems **reliable, understandable, and aligned with human intent**.
More broadly, I aim to develop **AI agents that augment human decision-making**, **coordinate with people and other agents**, and ultimately **reshape workflows** by **reducing cognitive and operational burden** while **preserving human oversight and control**.

Education

- **University of Wisconsin–Madison** *2025 – Present*
  Ph.D. Student in Computer Science Madison, USA
  ◦ Research focus: Agentic AI, LLM fine-tuning and prompting, Decision Process Optimization (DPO), Retrieval-Augmented Generation (RAG), and LLM reasoning & verification.
- **University of Wisconsin-Madison** *2023 - 2025*
  B.S. in Computer Science Madison, USA
  ◦ Graduated with Honors
- **Beijing Institute of Technology** *2021 – 2023*
  B.E. in Electronic and Computer Engineering Beijing, China
  ◦ Ranked Top 10% in ECE cohort

Research Highlights

- 2 publications including **CHI**; 3 additional manuscripts under review at **ICML**, **DIS**, and **TOCHI journal**.
- Fine-tuned and prompted small LLMs (up to 7B parameters); applied **DPO** for agent reasoning.
- Hands-on experience with **WebArena agents** and **embodied agents in IsaacSim environment**.
- Developed and extended **modular verification frameworks** for LTL-based reasoning and step-by-step output evaluation.
- Designed and conducted **user studies** for adaptive learning and verification-based planning systems, informing iterative design improvements.

Publications & Manuscripts

* = *equal contribution*

- **Xinyu Jessica Wang***, Haoyue Bai*, Yiyou Sun, Haorui Wang, Shuibai Zhang, Wenjie Hu, Mya Schroder, Bilge Mutlu, Dawn Song, Robert D. Nowak, "Position: The Long-Horizon Task Mirage? Diagnosing Where and Why Agentic Systems Break"
  Under Review — ICML 2026 (Position)
- **Xinyu Jessica Wang**, Christine P Lee, Bilge Mutlu, "LearnMate^2: Design and Evaluation of an LLM-powered Personalized and Adaptive Learning Support System for Online Learning"
  Under Review — DIS 2026
- Christine P Lee, David Porfirio, **Xinyu Jessica Wang**, Aws Albarghouthi, Bilge Mutlu, "U-Define: Integrating User-Defined Constraint Types into LLM-Based Planning"
  Under Review — TOCHI journal
- **Xinyu Jessica Wang**, Christine P Lee, Bilge Mutlu, "LearnMate: Enhancing Online Education with LLM-Powered Personalized Learning Plans and Support"
  CHI 2025 (Late-Breaking Work)
- Christine P Lee, David Porfirio, **Xinyu Jessica Wang**, Kevin Chenkai Zhao, Bilge Mutlu, "VeriPlan: Integrating Formal Verification and LLMs into End-User Planning"
  CHI 2025

Research Experience

- **Honda Research Institute USA** — *Nov 2025 – Present*
  Research Assistant — Madison, USA
- **People and Robots Lab** — *May 2024 – Present*
  Research Assistant — Madison, USA
- **Sprocket Lab** — *Mar 2025 – May 2025*
  Research Assistant — Madison, USA
- **NVIDIA Applied Deep Learning Research** — *Mar 2024 – May 2024*
  Research Assistant — Madison, USA

Teaching Experience

- **COMP SCI 540: Introduction to Artificial Intelligence** — *Fall 2025*
  Teaching Assistant — University of Wisconsin–Madison

Technical Skills

- **Language Models & Agents:** Autoregressive LLMs, RAG, DPO, small model fine-tuning (up to 7B), web and embodied agents.
- **Machine Learning:** Deep Learning, model benchmarking, logical reasoning evaluation, experiment design for human-AI collaboration.
- **Formal Methods:** Model checking (PRISM/Stormpy), LTL formalization, step-by-step verification frameworks for reasoning consistency.
- **Tools & Infrastructure:** Python, PyTorch, NumPy, Git, Docker, Linux.