

## Faculty Candidate Seminar

# Sheng Liu

Stanford University



## Reliable and Actionable AI Copilots for Medicine

Date: January 29, 2026

Time: 3:00-4:15 p.m.

Location: Shaffer 3

**Abstract:** While modern AI models achieve impressive performance on standard benchmarks, they remain brittle in high-stakes domains such as medicine, where uncertainty is pervasive and errors are costly. The central challenge is not model capacity, but bridging the gap between predictive accuracy in controlled settings and reliable behavior in real-world clinical workflows. Existing methods often struggle with noisy, multimodal data and lack principled mechanisms for incorporating expert knowledge during deployment.

In this talk, I present a set of machine learning methods designed to address these challenges. I will first describe approaches that leverage training dynamics and implicit signals to learn robust models from uncurated, imperfect data without amplifying noise. I will then introduce an inference-time framework that treats natural language not merely as input, but as a control interface, allowing domain experts to steer model behavior and agentic systems through structured linguistic feedback. By integrating robust learning and language-based model steering within agent-based framework, I demonstrate how AI agents can operate in closed-loop clinical settings with transparent reasoning, adaptive behavior, and human oversight.

**Bio:** Sheng Liu is a Postdoctoral Researcher at Stanford University, working with James Zou and Lei Xing. He received his PhD in Data Science from New York University. His research develops theory and algorithms for trustworthy and steerable AI, with the goal of bridging the gap between AI developed in the lab and actionable AI systems in the real world. His work is motivated by healthcare, with applications to dementia prediction and cancer treatment planning. Sheng's research has been published in leading venues, including Nature and other Nature Research journals, and presented in oral and spotlight sessions at top machine learning conferences such as CVPR, NeurIPS, ICML, and ICLR. His work has been featured by media outlets including The New York Times, YouTube, and Stanford News.