PURDUE UNIVERSITY | Elmore Family School of Electrical and Computer Engineering

# CompE Seminar Series


## Associate Professor Xiaojing Liao

Department of Computer Science
University of Illinois at Urbana-Champaign


Tuesday, March 31, 2026
Presentation: 12:00 P.M. - 1:30 P.M.
MSEE 112

## Towards Trustworthy Agent Development Frameworks


**Abstract:** The proliferation of large language model-integrated systems has introduced a new era of technological advancements and transformations across numerous aspects of our daily lives. As agents become more autonomous and deeply embedded in real workflows, the bar rises from "it works" to "it can be trusted". Building trustworthy and resilient agents requires end-to-end safeguards that preserve correctness, robustness, and accountability, especially when agents handle sensitive context and make consequential decisions.

In this talk, I will present our group's recent work toward a trustworthy framework for agent development, with a focus on building privacy-accountable LLM agents by design and by development. We outline principles and practical mechanisms for proactive privacy enhancement, runtime policy enforcement, and auditable accountability throughout the agent lifecycle, thereby setting the stage for deeper discussion on threat modeling and proactive defense strategies.


**Bio:** Xiaojing Liao is an Associate Professor in the Siebel School of Computing and Data Science at the University of Illinois Urbana-Champaign. She received her Ph.D. from the Georgia Institute of Technology in 2017. Her research interests span trustworthy machine learning systems, cyber threat intelligence, privacy compliance, and enforcement. She has published papers on leading system security venues such as S&P, USENIX Security, CCS, and NDSS. She co-authored the children's STEM book "Lorie in Cybersecurity Wonderland: The Fun of Camping & Mobile Security". She is the recipient of the NSF CAREER Award, the Amazon Research Award, the Meta Privacy-Enhancing Technology Research Award, and several Distinguished Paper and Distinguished Artifact Awards.