

# Wenxin Ding

Email: [wenxind@uchicago.edu](mailto:wenxind@uchicago.edu)

Website: [wenxind.github.io](https://wenxind.github.io)

## RESEARCH INTEREST

---

My research interest lies in machine learning security and privacy. Specifically, I focus on bridging the gap between theoretical understanding and empirical practice. My research studies the safety behavior of machine learning models under strategically optimized training data. Recently, I have been working on problems regarding vulnerabilities of text-to-image diffusion models and developing tools for content creators against copyright infringement.

## EDUCATION

---

### University of Chicago

Ph.D. in Computer Science

Advisors: Prof. Heather (Haitao) Zheng and Prof. Ben Y. Zhao

Chicago, IL

June, 2026

### Carnegie Mellon University

M.S. in Computer Science – Research Thesis

Advisors: Prof. Nihar B. Shah and Prof. Weina Wang

Pittsburgh, PA

August, 2021

B.S. in Computer Science and B.S. in Mathematical Sciences

Minor in Computational Finance

May, 2020

## Peer-Reviewed Publications

---

### Conferences

- **Wenxin Ding**, Cathy Li, Shawn Shan, Ben Y. Zhao, Haitao Zheng. “Understanding Implosion in Text-to-Image Generative Models.” in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2024.
- Shawn Shan, **Wenxin Ding**, Josephine Passananti, Haitao Zheng, Ben Y. Zhao. “Prompt-Specific Poisoning Attacks on Text-to-Image Generative Models.” in *Proceedings of IEEE Symposium on Security and Privacy (S&P)*, 2024.
- **Wenxin Ding**, Arjun Nitin Bhagoji, Ben Y. Zhao, and Haitao Zheng. “Towards Scalable and Robust Model Versioning.” in *Proceedings of IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, 2024.
- Sihui Dai\*, **Wenxin Ding**\*, Arjun Nitin Bhagoji, Daniel Cullina, Ben Y. Zhao, Haitao Zheng, and Prateek Mittal. “Characterizing the Optimal 0-1 Loss for Multi-class Classification with a Test-time Attacker.” in *Proceedings of Advances in Neural Information Processing Systems (NeurIPS)*, 2023. **Spotlight paper.** (\* for equal contribution)
- Shawn Shan, **Wenxin Ding**, Emily Wenger, Haitao Zheng, and Ben Y. Zhao. “Post-breach recovery: Protection against white-box adversarial examples for leaked DNN models.” in *Proceedings of ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2022.
- **Wenxin Ding**, Gautam Kamath, Weina Wang, and Nihar B. Shah. “Calibration with privacy in peer review.” in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, 2022.

## Workshops

- Wenxin Ding, Nihar B. Shah, and Weina Wang. “On the privacy-utility tradeoff in peer-review data analysis.” *AAAI Privacy-Preserving Artificial Intelligence (PPAI) workshop, 2021*. **Spotlight paper**.

## TEACHING EXPERIENCE

---

### Teaching Assistant

University of Chicago

- CMSC 25800 Adversarial Machine Learning
- CMSC 25300/35300 Mathematical Foundations of Machine Learning

Carnegie Mellon University

- 15110 Principles of Computing (Head Teaching Assistant)
- 15213 Introduction to Computer Systems
- 15440 Distributed Systems

### Mentor

- Strong Women Strong Girls, Pittsburgh, PA

## SERVICE

---

### Technical Program Committee

- 2025 ACM Conference on Computer and Communications Security (CCS)
- 2025 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)
- 2024 ACM Workshop on Artificial Intelligence and Security (AISec)

### Reviewer

- Nature
- 2025 ACM SIGGRAPH Asia
- 2024, 2025 The Conference on Uncertainty in Artificial Intelligence (UAI)
- SIAM Journal on Mathematics of Data Science (SIMODS)

## AWARDS

---

- 2025 University of Chicago William Rainey Harper Dissertation Fellowship
- 2024 University of Chicago UU Fellowship
- 2021 University of Chicago Eckhardt Scholar
- 2020 Carnegie Mellon University Senior Leadership Recognition
- 2019 Mark Stehlik SCS Alumni Undergraduate Impact Scholarship
- 2017 William Lowell Putnam Mathematical Competition (Rank: 255 / 4638)