# Weizhao Jin

weizhaoj@usc.edu | https://weizhaojin.netlify.app/ | www.linkedin.com/in/wzjin

## Education

**University of Southern California** Aug. 2020 – May 2025 (Expected)
*Ph.D. in Computer Science*

**University of Virginia** Aug. 2018 – May 2020
*Master of Engineering in Computer Engineering*

**Zhejiang University** Sep. 2014 – June 2018
*Bachelor of Engineering in Electrical Engineering and Automation*

## Experience

**Applied Scientist Intern** May 2024 – Aug. 2024
*Amazon* *AWS Privacy Security Automation*

- **GenAI Alignment & Privacy Analysis**
  Built a mutation-based LLM alignment and privacy analysis tool for AWS Bedrock LLM models including Claude 3.5 and Llama 3.1; the library tool is vended via CDK-based AWS CodeArtifact pipeline for easy access on SageMaker, used in Amazon Bedrock Guardrails pen-testing process

**Applied Scientist Intern** May 2023 – April 2024 (Extended)
*Amazon* *AWS Privacy Engineering*

- **AWS Differential Privacy Library**
  Built the key components including concurrent privacy budget accountant via DynamoDB and verified discrete sampler adapter for the first (easy-to-use, hard-to-misuse) differential privacy library at Amazon to raise the privacy bar of Amazon services including private machine learning training (integrated as part of AWS Clean Rooms)

**ML Privacy Researcher** Oct. 2022 – Mar. 2023
*TensorOpera AI* *Privacy-Preserving Federated Learning*

- **FedML with Homomorphic Encryption**
  Integrated homomorphic encryption to FedML's open-source library and MLOps services to facilitate the adoption of privacy enhancement for users; improve the efficiency of encrypted computation by 20x via proposed sensitive parameter selection

**Graduate Research Assistant** Aug. 2020 – Present
*University of Southern California* *Advisor: Srivatsan Ravi*

- **Efficient Two-Stage Privacy-Preserving Deep Entity Resolution**
  Built an efficient privacy-preserving deep entity resolution using encrypted rich entity embedding
- **Privacy-Preserving Path Validation for 5G Network Slicing**
  Designed a decentralized path validation protocol for 5G network slicing using NIZK; implemented the protocol on a testbed orchestrated using Ansible
- **Secure Publish-Process-Subscribe IoT System**
  Built a secure Publish-Process-Subscribe IoT system supporting functions like private set intersection and federated learning with Yao's Garbled Circuits, homomorphic encryption and proxy re-encryption atop MQTT protocol

**Applied Scientist Intern** May 2022 – Aug. 2022
*Amazon* *Buyer Risk Prevention*

- **Privacy-Preserving Federated Learning Using Fully Homomorphic Encryption**
  Built a privacy-preserving federated learning framework using homomorphic encryption for solving cross-region fraud detection data restriction as well as facilitating cross-team collaboration on sensitive data; worked with the engineering team to design and implement an AWS-based FL system; integrated our framework with tabular neural networks like TabNet and Tab1DCNN

**Student Research Assistant** Oct. 2018 – May 2020
*Security Lab, University of Virginia* *Advisor: Yuan Tian*

- **Vulnerabilities of Autonomous Vehicle Sensor Fusion Algorithm**
  Composed adversarial examples against perception module; tested attacks on the sensor fusion algorithm of the autonomous vehicle platform Baidu Apollo
- **Vulnerabilities of Dedicated Short-Range Communication**
  Analyzed existing vulnerabilities in the current version of DSRC protocol for connected vehicles; designed several attacks on DSRC protocol on connected vehicle modules

## Selected Publications (Full list)

- **Mutation-Based LLM Safety Prompt Testing Automated at Scale**
  Weizhao Jin, Aws Albarghouthi, Tancrède Lepoint, Matthew Schwartz, Nur Gucu, Chandu Chinthala
  Amazon Machine Learning Conference, 2024
- **Efficient Two-Stage Privacy-Preserving Deep Entity Resolution**
  Yixiang Yao*, Weizhao Jin*, Shanshan Han, Yuhang Yao, Carlee Joe-Wong, Srivatsan Ravi
  preprint, 2024
- **FedSecurity: A Benchmark for Attacks and Defenses in Federated Learning and Federated LLMs**
  Shanshan Han, Baturalp Buyukates, Zijian Hu, Han Jin, Weizhao Jin, Lichao Sun, Xiaoyang Wang, Chulin Xie, Yuhang Yao, Kai Zhang, Qifan Zhang, Yuhui Zhang, Chaoyang He, Salman Avestimehr
  ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2024
- **FedML-HE: An Efficient Homomorphic-Encryption-Based Privacy-Preserving Federated Learning System**
  Weizhao Jin*, Yuhang Yao*, Shanshan Han, Carlee Joe-Wong, Srivatsan Ravi, Salman Avestimehr, Chaoyang He
  preprint (short version: NeurIPS 2023 Federated Learning Workshop), 2023
- **Cross-Region Privacy-Preserving Federated Tabular Learning for Fraud Detection**
  Weizhao Jin, Shahin Navardi, Gaoyuan Du, Daniel Cociorva, Hakan Brunzell, Xiaoyang Liu
  Amazon Machine Learning Conference (Oral), 2023
- **FedGCN: Convergence-Communication Tradeoffs in Federated Training of Graph Convolutional Networks**
  Yuhang Yao, Weizhao Jin, Srivatsan Ravi, Carlee Joe-Wong
  Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS), 2023
- **P3V: Privacy-Preserving Path Validation System for Multi-Authority Sliced Networks**
  Weizhao Jin, Erik Kline, TK Satish Kumar, Lincoln Thurlow, Srivatsan Ravi
  preprint, 2023
- **Secure Publish-Process-Subscribe System for Dispersed Computing**
  Weizhao Jin, Bhaskar Krishnamachari, Muhammad Naveed, Srivatsan Ravi, Kwame-Lante Wright
  41st International Symposium on Reliable Distributed Systems (SRDS), 2022
- **Decentralized Privacy-Preserving Path Validation for Multi-Slicing-Authority 5G Networks**
  Weizhao Jin, Srivatsan Ravi, Erik Kline
  IEEE Wireless Communications and Networking Conference (WCNC), 2022
- **SMS Goes Nuclear: Fortifying SMS-Based MFA in Online Account Ecosystem**
  Weizhao Jin*, Xiaoyu Ji*, Ruiwen He, Zhou Zhuang, Wenyuan Xu, Yuan Tian
  Workshop on Data-Centric Dependability and Security (co-located with the IEEE/IFIP International Conference on Dependable Systems and Networks), 2021

## Skills

Python, Java/Kotlin, C/C++, JavaScript, Rust, AWS (DynamoDB, SageMaker, S3, IAM, CodeArtifact, CloudFormation), gRPC, PyTorch, Docker, MySQL, MATLAB

## Miscellaneous

- Reviewer: Amazon Research Awards, NeurIPS 2024, ICLR 2025, IEEE Transactions on Information Forensics and Security, IEEE IoTJ, IEEE Transactions on Network Science and Engineering
- USC Graduate School Research Award (2022)
- Amazon BRP Trustworthy and Privacy ML Working Group Talk: Homomorphic Encryption and Privacy-Preserving Federated Learning (2022)
- USC ISI NCD Talk: Decentralized Privacy-Preserving Path Validation for Multi-Authority 5G Networks (2022)
- Graduate Teaching Assistant: USC CSCI 103 Introduction to Programming (C++), UVA APMA 3100 Probability