Rongzhe Wei

Tel: (765)-746-9720 Email: rongzhe.wei@gatech.edu
Address: 12th Floor, Coda Building, 756 W Peachtree St NW, Atlanta, GA 30308

EDUCATION BACKGROUND

Georgia Institute of Technology, Atlanta, Georgia Currently, Anticipated Graduation Date: 05/2026
Ph.D. in Machine Learning
School of Electrical and Computer Engineering

Xi'an Jiaotong University (XJTU), Xi'an, China 09/2017 - 07/2021
Bachelor of Science in Mathematics (Overall GPA: **3.89**/4.0)
Honors Math Program by Ministry of Education, P.R. China

RESEARCH INTERESTS

- **Trustworthy Machine Learning (Machine Unlearning; Privacy), Graphs Analysis (Graph Diffusion, Graph Neural Networks), Large Language Models, etc.**

SELECTED PUBLICATIONS (Sorted by Year)

Under Review:

- **Rongzhe Wei***, Peizhi Niu*, Hans Hao-Hsun Hsu, Ruihan Wu, Haoteng Yin, Mohsen Ghassemi, Yifan Li, Vamsi K. Potluru, Eli Chien, Kamalika Chaudhuri, Olgica Milenkovic, Pan Li. Do LLMs Really Forget? Evaluating Unlearning with Knowledge Correlation and Confidence Awareness. Under-review (NeurIPS).
  Summary: *In this project, we propose a novel knowledge unlearning definition with corresponding evaluation framework. Our framework accurately captures the implicit structure of real-world knowledge by representing relevant factual contexts as knowledge graphs with associated confidence scores.*

- Haoteng Yin, **Rongzhe Wei**, Eli Chien, Pan Li. Privately Learning from Graphs with Applications in Large Language Model Finetuning. Under-review (COLM). In *SFLLM-NeurIPS'24 Workshop*.
  Summary: *In this project, we propose a privacy-preserving relational learning pipeline that ensures differential privacy in fine-tuning large language models on sensitive graph data using a tailored DP-SGD*

- Shuaiqi Wang, **Rongzhe Wei**, Mohsen Ghassemi, Eleonora Kreacic, Vamsi K. Potluru. Guarding Multiple Secrets: Enhanced Summary Statistic Privacy for Data Sharing. In *ICLR'24 PML Workshop.*
  Summary: *In this project, we propose a framework to define, analyze, and protect multi-secret summary statistics privacy by designing tailored privacy metrics and release mechanisms, balancing privacy and data distortion, and evaluating their effectiveness on real-world data.*

- Yinan Huang, Haoteng Yin, Eli Chien, **Rongzhe Wei**, Pan Li. Node-level Differential Private Relational Learning on Graphs. Under-review (NeurIPS).
  Summary: *In this project, we consider the problem of node-level privacy-preserving relational learning on graphs with novel privacy amplification analysis.*

- Tianchun Li, Tianci Liu, Xingchen Wang, **Rongzhe Wei,** Pan Li, Lu Su, Jing Gao. Towards Universal Debiasing for Language Models-based Tabular Data Generation. Under-review (EMNLP).
  Summary: *In this project, we developed a universal debiasing framework for large language model-based tabular data generation that mitigates fairness issues by minimizing group-level dependencies, combining mutual information estimation with DPO and targeted debiasing techniques.*

Published:

- **Rongzhe Wei**, Mufei Li, Mohsen Ghassemi, Eleonora Kreacic, Yifan Li, Xiang Yue, Bo Li, Vamsi K. Potluru, Pan Li, Eli Chien. Underestimated Privacy Risks for Minority Populations in Large Language Model Unlearning. In *International Conference on Machine Learning* **(ICML'25)**

Summary: *In this project, we identify a critical flaw that the privacy risks faced by minority groups within the training data are often significantly underestimated in large language model unlearning.*

- Haoyu Wang, Shikun Liu, **Rongzhe Wei**, Pan Li. Generalization Principles for Inference over Text-Attributed Graphs with Large Language Models. In *International Conference on Machine Learning* **(ICML'25)**
  Summary: *In this project, we address the challenges of applying large language models to text-attributed graph learning by proposing the LLM-BP framework, which integrates task-adaptive embeddings and a generalizable graph information aggregation mechanism.*

- **Rongzhe Wei**, Eli Chien, Pan Li. Differentially Private Graph Diffusion with Applications in Personalized PageRanks. In *Advances in Neural Information Processing Systems* (**NeurIPS'24**)
  Summary*: In this project, we propose a novel graph diffusion framework with a Wasserstein Distance tracking method that extends beyond traditional Privacy Amplification by Iteration analysis, eliminating the diameter assumption and achieving state-of-the-art privacy-utility trade-offs in personalized PageRank applications.*

- **Rongzhe Wei**, Eleonora Kreačić, Haoyu Wang, Haoteng Yin, Eli Chien, Vamsi K Potluru, Pan Li. On the Inherent Privacy Properties of Discrete Denoising Diffusion Models. In **TMLR'24**, selected to **ICLR'25**.
  Summary: *In this project, we analyze and demonstrate the weak inherent privacy guarantees of discrete denoising diffusion models over discrete data and outliers suffers from higher privacy leakages.*

- Yizhou Wang, Can Qin, **Rongzhe Wei**, Yi Xu, Yue Bai, & Yun Fu. SLA^{2}P: Self-supervised Anomaly Detection with Adversarial Perturbation. In **TKDE'24**.

- Tianyi Zhang*, Haoteng Yin*, **Rongzhe Wei**, Pan Li, Anshumali Shrivastava. Learning Scalable Structural Link Representations with Bloom Signatures. Proceedings of the *ACM Web Conference*. (**WWW'24**)

- **Rongzhe Wei**, Haoteng Yin, Junteng Jia, Austin R. Benson, Pan Li. Understanding Non-linearity in Graph Neural Networks from the Bayesian-Inference Perspective. In *Advances in Neural Information Processing Systems* (**NeurIPS'22**)

PROFESSIONAL SERVICES

- **Conference Reviewer:** NeurIPS'22-25, ICML'24-25, ICLR'25, AISTATS'23-24, ISIT'25, AAAI'24, LoG'22-24
- **Journal Reviewer:** Computers and Mathematics with Applications, TMLR
- **Graduate Teaching Assistant –** ECE 6720 Convex Optimization (Graduate Level) / ECE 3077 Introduction to Probability and Statistics for ECEs / ECE 8003 Conversational AI

INDUSTRIAL EXPERIENCES

- **Amazon** **Seattle, WA**
  AI Research Intern May 2025 – Now
  Project: LLM Agent for Decision-making

- **JP Morgan Chase & Co.** **Manhattan, NYC, NY**
  AI Research Intern June 2023 – August 2023
  Project: A General Framework for Graph Data Generation Control via Margin Relaxed Schrodinger Bridges
  Mentors: Eleonora Kreačić and Vamsi K Potluru.

SELECTED HONORS & SCHOLARSHIP

- Travel Award for NeurIPS 2022 2022

- Student Award in 2019 IEEE International Conference on BigData 2019
- The First Prize of "*Zhufeng*" Scholarship, established for "*Pilot Scheme of Top-notch Talent Cultivation in Basic Disciplines*", Ministry of Education for three times 2017 – 2018, 2018 – 2019, 2019 - 2020
- Outstanding Student Award for three times, XJTU 2017 – 2018, 2018 – 2019, 2019 - 2020
- First-class Scholarship for three times, XJTU 2017 – 2018, 2018 – 2019, 2019 - 2020

TECHNICAL SKILLS

- **Languages:** Chinese (native), English
- **Programming Languages:** Python, C#, Matlab, C, SQL, HTML
- **Piano:** Band 9 out of 9 (Central Conservatory of Music) achieved at the age of 12, and won the 5th China Outstanding Talents Art Festival National Competition (Gold Award)