

Jiachen (Tianhao) Wang

E-Quad, Princeton, NJ – 08540 – USA

🌐 Academic Website

🔍 Google Scholar

✉ tianhaowang@princeton.edu

RESEARCH INTERESTS

I am interested in developing principled **data-centric** methodologies to build **trustworthy AI at scale**. Poor data decisions cascade through the entire AI lifecycle, producing opaque and unreliable models, wasting compute on low-quality data, and exposing sensitive user information. To this end, I design *theoretically grounded* and *scalable* algorithms for **data attribution** (quantifying data influence), **data curation** (optimizing dataset quality), and **data privacy** (provably preventing data leakage).

EDUCATION

Princeton University

PhD Candidate in Electrical and Computer Engineering

Princeton, NJ

Sept 2021 - May 2026 (expected)

Advisor: **Prof. Prateek Mittal**

Harvard University

MEng in Computational Science and Engineering

Cambridge, MA

Sept 2019 - May 2021

Thesis: Concurrent Composition Theorem of Differential Privacy

Advisor: **Prof. Salil Vadhan**

University of Waterloo

BMath in Computer Science and Statistics

Waterloo, ON

Sept 2016 - May 2019

PUBLICATIONS

Summary: *Since 2021, I have led or contributed significantly to 20 research papers, including 15 as sole first author. My research has been recognized with ICLR 2025 Best Paper Honorable Mention and 9 Oral/Spotlight presentations at top-tier AI conferences. My work has achieved real-world impact, influencing data efforts at Google and being adopted by PISTIS (an EU data marketplace platform).*

Can Small Training Runs Reliably Guide Data Curation? Rethinking Proxy-Model Practice

Jiachen T. Wang, Tong Wu, Kaifeng Lyu, James Zou, Dawn Song, Ruoxi Jia, Prateek Mittal

ICLR 2026 {PDF}

Impacted Google's data efforts for frontier AI development.

Capturing the Temporal Dependence of Training Data Influence

Jiachen T. Wang, Dawn Song, James Zou, Prateek Mittal, Ruoxi Jia

ICLR 2025 (Oral, top 1.5% among submissions) {PDF}

Data Shapley in One Training Run

Jiachen T. Wang, Prateek Mittal, Dawn Song, Ruoxi Jia

ICLR 2025 (Oral, top 1.5% among submissions) {PDF}

🏆 Outstanding Paper Honorable Mention (6 out of 11,000+ submissions)

Impacted Google's data efforts for frontier AI development.

GREATS: Online Selection of High-Quality Data for LLM Training in Every Iteration

Jiachen T. Wang, Tong Wu, Dawn Song, Prateek Mittal, Ruoxi Jia

NeurIPS 2024 (Spotlight, top 3% among submissions) {PDF}

Rethinking Data Shapley for Data Selection Tasks: Misleads and Merits

Jiachen T. Wang, Tianji Yang, James Zou, Yongchan Kwon, Ruoxi Jia

ICML 2024 (Oral, top 1.5% among submissions) {PDF}

Efficient Data Valuation for Weighted Nearest Neighbor Algorithms

Jiachen T. Wang, Prateek Mittal, Ruoxi Jia

AISTATS 2024 (Oral, top 1.6% among submissions) {PDF}

DP-OPT: Make Large Language Model Your Privacy-Preserving Prompt Engineer

Junyuan Hong, Jiachen T. Wang, Chenhui Zhang, Zhangheng Li, Bo Li, Zhangyang Wang

ICLR 2024 (Spotlight, top 3% among submissions) {PDF}

Privacy-Preserving In-Context Learning for Large Language Models

Tong Wu*, Ashwinee Panda*, **Jiachen T. Wang***, Prateek Mittal
ICLR 2024 {PDF}

Threshold KNN-Shapley: A Linear-Time and Privacy-Friendly Approach to Data Valuation

Jiachen T. Wang, Yuqing Zhu, Yu-Xiang Wang, Ruoxi Jia, Prateek Mittal
NeurIPS 2023 (**Spotlight, top 3% among submissions**) {PDF}

A Randomized Approach for Tight Privacy Accounting

Jiachen T. Wang, Saeed Mahloujifar, Tong Wu, Ruoxi Jia, Prateek Mittal
NeurIPS 2023 {PDF}

Recognized with Princeton's Early Career Graduate Award.

Data Banzhaf: A Robust Data Valuation Framework for Machine Learning

Jiachen T. Wang, Ruoxi Jia
AISTATS 2023 (**Oral, top 1.6% among submissions**) {PDF}

Featured in the course materials for Harvard CS236 (Topics in Computer Science and Economics).

LAVA: Data Valuation without Pre-Specified Learning Algorithms

Hoang Anh Just*, Feiyang Kang*, **Jiachen T. Wang**, Yi Zeng, Myeongseob Ko, Ming Jin, Ruoxi Jia
ICLR 2023 (**Spotlight, top 3% among submissions**) {PDF}

Incorporated into open-source ML tools by AppliedAI Institute.

Data Valuation in the Absence of a Reliable Validation Set

Himanshu Jahagirdar, **Jiachen T. Wang**, Ruoxi Jia
TMLR (**Featured Certification, top 2.5% among accepted papers**) {PDF}

One-Round Active Learning through Data Utility Learning and Proxy Models

Jiachen T. Wang, Si Chen, Ruoxi Jia
TMLR {PDF}

Rényi Differential Privacy of Propose-Test-Release and Applications to Private and Robust ML

Jiachen T. Wang, Saeed Mahloujifar, Shouda Wang, Ruoxi Jia, Prateek Mittal
NeurIPS 2022 {PDF}

Improving Cooperative Game Theory-based Data Valuation via Data Utility Learning

Tianhao Wang, Yu Yang, Ruoxi Jia
ICLR 2022 Workshop on Socially Responsible Machine Learning {PDF}
Incorporated into open-source ML tools by AppliedAI Institute.

Concurrent Composition of Differential Privacy

Salil Vadhan*, **Tianhao Wang***
TCC 2021 {PDF}

Invited to the Journal of Cryptology.

DPLUS: Boosting Utility of Differentially Private Deep Learning via Randomized Smoothing

Wenxiao Wang, **Tianhao Wang**, Lun Wang, Nanqing Luo, Pan Zhou, Dawn Song, Ruoxi Jia
PoPETS'21 {PDF}

Improving Robustness to Model Inversion Attacks via Mutual Information Regularization

Tianhao Wang, Yuheng Zhang, Ruoxi Jia
AAAI'21 {PDF}

A Principled Approach to Data Valuation for Federated Learning

Tianhao Wang, Johannes Rausch, Ce Zhang, Ruoxi Jia, Dawn Song
Book Chapter in Federated Learning: Privacy and Incentive (2020) {PDF}

I have also co-authored papers across diverse research fields in trustworthy artificial intelligence.

Does More Inference-Time Compute Really Help Robustness?

Tong Wu, Chong Xiang, **Jiachen T. Wang**, Weichen Yu, Chawin Sitawarin, Vikash Sehwal, Prateek Mittal
In submission, 2025 {PDF}

Effectively Controlling Reasoning Models through Thinking Intervention

Tong Wu, Chong Xiang, **Jiachen T. Wang**, G. Edward Suh, Prateek Mittal
In submission, 2025 {PDF}

A Sustainable Machine Learning Economy Needs Data Deals That Work for Generators

Ruoxi Jia, Luis Oala, Wenjie Xiong, Suqin Ge, **Jiachen T. Wang**, Feiyang Kang, Dawn Song
NeurIPS 2025 Position Paper Track

Boosting Alignment for Post-Unlearning Text-to-Image Generative Models

Myeongseob Ko, Henry Li, Zhun Wang, Jonathan Patsenker, **Jiachen T. Wang**, Qinbin Li, Ming Jin, Dawn Song, Ruoxi Jia
NeurIPS 2024 {PDF}

Language Models as Science Tutors

A. Chevalier, J. Geng, A. Wettig, H. Chen, S. Mizera, ..., **Jiachen T. Wang**, ..., Sanjeev Arora, Danqi Chen
ICML 2024 {PDF}

BaDExpert: Extracting Backdoor Functionality for Accurate Backdoor Input Detection

Tinghao Xie, Xiangyu Qi, Ping He, Yiming Li, **Jiachen T. Wang**, Prateek Mittal
ICLR 2024 {PDF}

Uncovering Adversarial Risks of Test-Time Adaptation

Tong Wu, Feiran Jia, Xiangyu Qi, **Jiachen T. Wang**, Vikash Sehwal, Saeed Mahloujifar, Prateek Mittal
ICML 2023 {PDF}

Towards a Proactive ML Approach for Detecting Backdoor Poison Samples

Xiangyu Qi, Tinghao Xie, **Jiachen T. Wang**, Tong Wu, Saeed Mahloujifar, Prateek Mittal
USENIX 2023 {PDF}

Just Rotate it: Deploying Backdoor Attacks via Rotation Transformation

Tong Wu, **Jiachen T. Wang**, Vikash Sehwal, Saeed Mahloujifar, Prateek Mittal
CCS Workshop on Artificial Intelligence and Security (AISec'22) {PDF}

RIGA: a Covert and Robust White-Box Watermarking of Deep Neural Networks

Tianhao Wang, Florian Kerschbaum
WWW'21 {PDF}

TECHNICAL WRITINGS

An Economic Solution to Copyright Challenges of Generative AI

Jiachen T. Wang, Zhun Deng, Hiroaki Chiba-Okabe, Boaz Barak, Weijie Su
Technical Note, April 2024 {PDF}

A Note on “Efficient Task-Specific Data Valuation for Nearest Neighbor Algorithms”

Jiachen T. Wang, Ruoxi Jia
Technical Note, April 2023 {PDF}

[Adopted by PISTIS \(a data marketplace platform by the EU\) for data valuation and recommendation.](#)

A Note on “Towards Efficient Data Valuation Based on the Shapley Value”

Jiachen T. Wang, Ruoxi Jia
Technical Note, Feb 2023 {PDF}

SCHOLARSHIPS & AWARDS

ICLR 2025 Best Paper Honorable Mention

April 2025

One of 6 papers recognized among 11,000+ submissions (for “Data Shapley in One Training Run”).

Apple Scholars in AIML PhD Fellowship

March 2025

One of 21 recipients worldwide; full academic support from 2025-2027.

Yan Huo *94 Graduate Fellowship @ Princeton

Jan 2025

One of 3 recipients department-wide; full academic support in Spring 2025.

Rising Stars in Data Science

Nov 2024

One of 30 early-career researchers worldwide selected by UCSD, UChicago, and Stanford data science institutes.

Pramod Subramanian Early Career Graduate Award @ Princeton

May 2023

One of 2 second-year Ph.D. recipients department-wide for outstanding research and general exam performance.

Gordon Y. S. Wu Fellowship @ Princeton

Sept 2021

For the highest honor that Princeton University can bestow on an incoming graduate student.

Winston and Diana Cherry Scholarship @ UWaterloo

May 2019

For the highest major average at Spring Convocation with an Honors Statistics Major designation.

J.A. Brzozowski Scholarship @ UWaterloo

Dec 2018

For outstanding achievement in courses focusing on **theoretical computer science**.

INVITED TALKS & TUTORIALS

Princeton ECE574 (Security and Privacy) Guest Lecture Introduction to Differential Privacy	<i>Oct 2025</i>
Apple Machine Learning Research Seminar Can Small Training Runs Reliably Guide Data Curation? Rethinking Proxy-Model Practice	<i>Aug 2025</i>
Google NYC Algorithms and Optimization Seminar Data Shapley in One Training Run	<i>July 2025</i>
Tsinghua University ML Foundations Seminar Capturing the Temporal Dependence of Training Data Influence	<i>Jan 2025</i>
NeurIPS 2024 Tutorial (2 hours, with Ruoxi Jia and Ludwig Schmidt) Advancing Data Selection for Foundation Models: From Heuristics to Principled Methods	<i>Dec 2024</i>
IDEAL Institute Workshop on Harmonious Human-AI Ecosystems Capturing the Temporal Dependence of Training Data Influence	<i>Nov 2024</i>
Virginia Tech ECE6514 (Trustworthy Machine Learning) Guest Lecture Privacy-preserving Language Model Inference	<i>Nov 2024</i>
Brandeis University Michtom School of Computer Science Principled Data Attribution at Scale	<i>Sept 2024</i>
Microsoft Research NYC Machine Learning Reading Group An Influence-Preserving Approach to Continual Pretraining of LLMs	<i>Aug 2024</i>
Google Privacy Seminar TKNN-Shapley: A Linear-Time and Privacy-Friendly Approach to Data Valuation	<i>Jan 2024</i>
Norwegian Computing Center Explaining AI Seminar TKNN-Shapley: A Linear-Time and Privacy-Friendly Approach to Data Valuation	<i>Jan 2024</i>
Google Privacy Seminar A Randomized Approach to Tight Privacy Accounting	<i>May 2023</i>
Boston-Area Data Privacy Seminar Concurrent Composition of Differential Privacy	<i>Aug 2021</i>

EXPERIENCES

Google Research <i>Student Researcher</i>	Remote <i>March 2025 - May 2026</i>
o Mentored by Lin Chen , Mohammad Hossein Bateni , and Vahab Mirrokni on <i>Gemini Data</i> .	
Microsoft Research <i>Research Intern</i>	New York City, NY <i>June 2024 - Aug 2024</i>
o Mentored by Huseyin A. Inan , Janardhan Kulkarni , and Ida Momennejad on <i>continual pretraining of LLMs</i> .	
Virginia Tech <i>Research Assistant</i>	Blacksburg, VA <i>June 2021 - Aug 2021</i>
o Mentored by Prof. Ruoxi Jia on <i>privacy-preserving machine learning</i> and <i>data-centric machine learning</i> .	
Harvard University <i>Research Assistant</i>	Cambridge, MA <i>Aug 2020 - Aug 2021</i>
o Mentored by Prof. Salil Vadhan on the <i>theoretical foundation of differential privacy</i> .	
University of California, Berkeley <i>Visiting Student Scholar</i>	Remote <i>May 2020 - Aug 2020</i>
o Mentored by Prof. Dawn Song and Dr. Ruoxi Jia on <i>data valuation for machine learning</i> .	
University of Waterloo <i>Undergraduate Research Assistant</i>	Waterloo, ON <i>June 2018 - May 2019</i>

- o Mentored by **Prof. Florian Kerschbaum** on *machine learning security*.

TEACHING

Information Security (COS432) @ Princeton Graduate Teaching Assistant for Prof. Maria Apostolaki	<i>Spring 2024</i>
Fairness and Privacy: Perspectives of Law and Probability (CS126) @ Harvard Teaching Fellow for Prof. Cynthia Dwork	<i>Fall 2020</i>
Statistics (STAT231) @ UWaterloo Teaching Assistant	<i>Fall 2018</i>
Algebra (MATH135) @ UWaterloo Teaching Assistant	<i>Fall 2017</i>

MENTORING

Mahavir Dabas (Junior PhD) o Resulted in his first-author conference paper submission.	<i>Since 2025</i>
Weida Li (Junior PhD)	<i>Since 2024</i>
Luxi He (Junior PhD)	<i>Since 2023</i>
Tianji Yang (Undergraduate) o Resulted in his second-author ICML 2024 paper (Oral Presentation).	<i>2023</i>
Tinghao Xie (Junior PhD) o Resulted in his first-author ICLR 2024 paper.	<i>Since 2023</i>
Himanshu Jahagirdar (Master) o Resulted in his first-author TMLR paper (Featured Certification).	<i>2023</i>

SERVICE

Lead Organizer for the 3rd Workshop on Data Problems for Foundation Models With Ruoxi Jia, Zheng Xu, Martin Jaggi, Mónica Ribero, Pratyush Maini, Yuzheng Hu, Luxi He.	<i>ICLR 2026</i>
Lead Organizer for the 2nd Workshop on Data Problems for Foundation Models With Ruoxi Jia, Pang Wei Koh, Dawn Song, Jerone Andrews, Hoang Anh Just, Feiyang Kang.	<i>ICLR 2025</i>
Conference Paper Reviewer	
o International Conference on Artificial Intelligence and Statistics (AISTATS)	2023, 2024, 2025
o International Conference on Learning Representations (ICLR)	2023, 2024, 2025
o Conference on Neural Information Processing Systems (NeurIPS)	2023, 2024, 2025
o International Conference on Machine Learning (ICML)	2023, 2024, 2025
o Privacy Enhancing Technologies Symposium (PETS)	2021, 2022
Workshop Program Committee	
o Data-centric Machine Learning	ICML 2023, ICLR 2024, ICML 2024
o Privacy Regulation and Protection in Machine Learning	ICLR 2024
Research Mentor for AI4ALL Ignite Program	<i>2024</i>
o Mentored \approx 100 students from underrepresented groups (Black, Indigenous, women, and non-binary) in developing AI/ML portfolio projects, providing both technical guidance and career advice.	