Haizhong Zheng

Email: haizhonz@andrew.cmu.edu Homepage: zhenghaizhong.com

INTERESTS

Large Language Model; Efficient Machine Learning; Data Efficiency Algorithm; Machine Learning System.

My research focuses on building models, algorithms, and systems for scalable and efficient ML. The goal is to bridge the gap between the rapid scaling of models and the slower scaling of hardware and high-quality data. In particular, my work has been along two lines: 1) Designing and training hardware-aware and system-aware models for fast model inference. 2) Designing algorithms for efficient data selection, augmentation, and human feedback.

WORK EXPERIENCE

Carnegie Mellon University (CMU) Jan. 2025 - Current
Postdoctoral Research Associate
Advisor: Prof. Beidi Chen

Lawrence Livermore National Laboratory (LLNL) May. 2023 - Aug. 2023
Research Intern
Mentor: Dr. Bhavya Kailkhura

Amazon Web Service (AWS) May. 2021 - Aug. 2021
Applied Scientist Intern
Mentor: Dr. Wei Zhang and Dr. Qian Cui

EDUCATION

University of Michigan, Ann Arbor 2024
Ph.D., Computer Science
Advisor: Prof. Atul Prakash

Shanghai Jiao Tong University (SJTU) 2015, 2018
B.S. & M.S., Computer Science and Engineering
Advisor: Prof. Haojin Zhu

SELECTED PUBLICATIONS

- Prosperity before Collapse: How Far Can Off-Policy RL Reach with Stale Data on LLMs?
  Haizhong Zheng, Jiawei Zhao, Beidi Chen
  Preprint
- When "Correct" Is Not Safe: Can We Trust Functionally Correct Patches Generated by Code Agents?
  Yibo Peng, James Song, Lei Li, Xinyu Yang, Mihai Christodorescu, Ravi Mangal, Corina Pasareanu, Haizhong Zheng, Beidi Chen
  Preprint
- RLBoost: Harvesting Preemptible Resources for Cost-Efficient Reinforcement Learning on LLMs
  Yongji Wu*, Xueshen Liu*, Haizhong Zheng, Juncheng Gu, Beidi Chen, Z. Morley Mao, Arvind Krishnamurthy, Ion Stoica
  NSDI 2026
- WAVE: Leveraging Architecture Observation for Privacy-Preserving Model Oversight
  Haoxuan Xu*, Chen Gong*, Beijie Liu*, Haizhong Zheng, Beidi Chen, Mengyuan Li

ASPLOS 2026

- Act Only When It Pays: Efficient Reinforcement Learning for LLM Reasoning via Selective Rollouts

  Haizhong Zheng, Yang Zhou, Brian R. Bartoldson, Bhavya Kailkhura, Fan Lai, Jiawei Zhao, Beidi Chen

  NeurIPS 2025

- Kinetics: Rethinking Test-Time Scaling Laws

  Ranajoy Sadhukhan*, Zhuoming Chen*, Haizhong Zheng, Yang Zhou, Emma Strubell, Beidi Chen

  NeurIPS 2025

- Plato: Plan to Efficiently Decode for Large Language Model Inference

  Shuowei Jin*, Xueshen Liu*, Yongji Wu*, Haizhong Zheng, Qingzhao Zhang, Atul Prakash, Matthew Lentz, Danyang Zhuo, Feng Qian, Z. Morley Mao

  COLM 2025

- Harmful Terms and Where to Find Them: Measuring and Modeling Unfavorable Financial Terms and Conditions in Shopping Websites at Scale

  Elisa Tsai, Neal Mangaokar, Boyuan Zheng, Haizhong Zheng, Atul Prakash

  WWW 2025, *Oral*

- Label-Free Coreset Selection with Proxy Training Dynamics

  Haizhong Zheng*, Elisa Tsai*, Yifu Lu, Jiachen Sun, Brian R. Bartoldson, Bhavya Kailkhura, Atul Prakash

  ICLR 2025

- Learn To be Efficient: Build Structured Sparsity in Large Language Models

  Haizhong Zheng, Xiaoyan Bai, Xueshen Liu, Z. Morley Mao, Beidi Chen, Fan Lai, Atul Prakash

  NeurIPS 2024, *Spotlight*

- Leveraging Hierarchical Feature Sharing for Efficient Dataset Condensation

  Haizhong Zheng, Jiachen Sun, Shutong Wu, Bhavya Kailkhura, Zhuoqing Mao, Chaowei Xiao, Atul Prakash

  ECCV 2024

- CALICO: Self-Supervised Camera-LiDAR Contrastive Pre-training for BEV Perception

  Jiachen Sun, Haizhong Zheng, Qingzhao Zhang, Atul Prakash, Z. Morley Mao, Chaowei Xiao

  ICLR 2024

- Coverage-centric Coreset Selection for High Pruning Rates

  Haizhong Zheng, Rui Liu, Fan Lai, Atul Prakash

  ICLR 2023

- Efficient Adversarial Training with Transferable Adversarial Examples

  Haizhong Zheng, Ziqi Zhang, Juncheng Gu, Honglak Lee, Atul Prakash

  CVPR 2020

- Analyzing the Interpretability Robustness of Self-Explaining Models

  Haizhong Zheng, Earlence Fernandes, Atul Prakash

  ICML 2019 Workshop

- Smoke Screener or Straight Shooter: Detecting Elite Sybil Attacks in User-Review Social Networks

  Haizhong Zheng, Minhui Xue, Hao Lu, Shuang Hao, Haojin Zhu, Xiaohui Liang, Keith Ross

  NDSS 2018

(* indicates equal contribution)

GRANT PROPOSALS

I actively contributed to the proposal design, proposal writing, and presentation for the following grants:

Assessing and Improving Safety and Alignment of LLMs — NAIRR Pilot, 2024
GPU budgets — PI: Atul Prakash

Data Efficiency of LLMs Fine-tuning with RLHF — Cisco, 2023
$150*K* — PI: Atul Prakash

Intelligent Assistants for Detecting Social Engineering Scams — OpenAI, 2023
$100*K* — PI: Atul Prakash

TEACHING

Co-Lead Instructor — *Winter 2023*
Secure and Trustworthy Machine Learning — UMich EECS 598

- My responsibilities include designing the course, teaching the lectures, leading discussions, advising course projects, and office hours.
- EECS 598 covers research topics in machine learning security and privacy.
- Course rating: 4.9 out of 5; Instructor rating: 5 out of 5.

Graduate Student Instructor — *Fall 2021*
Data Structures and Algorithms — UMich EECS 281

- My responsibilities include teaching lectures, leading discussions on a weekly lab section, grading, and holding 6hrs of office hours each week.

STUDENT MENTORED

I am fortunate to mentor and co-advise the research of the following students:

- **Ph.D. Students**: Elisa Tsai (Umich CSE), Shuowei Jin (Umich CSE), Ruiyang Zhu (Umich CSE), Xueshen Liu (Umich CSE)
- **Masters**: Tianyu Zhang (→ ByteDance), Tiejin Chen (→ ASU Ph.D.), Shutong Wu (→ UWsic Ph.D.), Dongfang Ling (CMU CS), Yibo Peng (CMU CS)
- **Undergraduates**: Ziqi Zhang (→ Google), Xiaoyan Bai (→ UChicago Ph.D.), Yifu Lu (→ Princeton Ph.D.), Chengshuo Jiang (Umich CSE), James Song (Umich CSE), Lei Li (Peking University), Qilong Feng (NUS)

SERVICE

Conference/Journal Paper Reviewer: ECCV 2020, TPAMI 2020, ICLR 2022-2024, NeurIPS 2022-2025, ICML 2023-2025, AAAI 2024, ICML 2024 ES-FoMo-II Workshop

AWARDS & HONORS

- Rackham Travel Grant: ICML'19 — 2019
- Academic Excellence Scholarship of SJTU — 2015-2017
- Best Presentation Award in A3 Foresight Program 2015 Annual Workshop — 2015
- Academic Excellence Scholarship of SJTU — 2012-2014
- National Olympiad in Informatics in Provinces (NOIP), First Prize — 2010

REFERENCES

Prof. Atul Prakash
Division Chair, Professor
Computer Science & Engineering
University of Michigan
aprakash@umich.edu

Prof. Beidi Chen
Assitant Professor
Electrical and Computer Engineering
Carnegie Mellon University
beidic@andrew.cmu.edu

Prof. Chaowei Xiao
Assitant Professor
Information School
University of Wisconsin, Madison
cxiao34@wisc.edu

Dr. Bhavya Kailkhura
Principal Research Scientist
Center for Applied Scientific Computing
Lawrence Livermore National Laboratory
kailkhura1@llnl.gov