ANDY (ZEHAO) ZHAO
609-937-3646 | zehao.andy.zhao@gmail.com | zehao.org/academic_homepage | No Sponsorship Needs

RESEARCH KEYWORDS

RSI, Distillation, Multi-agent RL and Game Theory, Post-training alignment, Test-time Training, Proactive Agent

EDUCATION

UNIVERSITY OF SOUTHERN CALIFORNIA Los Angeles & Bay Area, CA
PhD in Electrical and Computer Engineering, Center for Autonomy and Artificial Intelligence, GPA:4.0/4.0 *2023 – 2027*
Master's degrees earned en route to the PhD: Computer Science and Pure Mathematics

MASSACHUSETTS INSTITUTE OF TECHNOLOGY Cambridge, MA
Master in Business Analytics and Operation Research, Operation Research Center (Admit as PhD) GPA:5.0/5.0 *2023*

UNIVERSITY OF CALIFORNIA, BERKELEY *(4 Majors, First in Berkeley School History)* Berkeley, CA
Bachelor in Computer Science (Honors), Statistics, Math, and Data Science (Honors), GPA:3.9/4.0 *2022*

EXPERIENCE

Research Scientist Intern Menlo Park, CA
Meta Superintelligence Lab (FAIR) *Jul 2026 -*
- Research on test-time training, model plasticity and efficient KV-cache architectures.
- Research on Proactive Agent with Situation Awareness and under multi-agent environment changes.
- Research on Recursive self-improvement and on-Policy Distillation and reasoning with black-box teachers.

Research Scientist Intern Redmond, WA
Microsoft Research, Reinforcement Learning and System Group *April 2026 – Jul 2026*
- Developed learning frameworks for populations of self-evolving AI agents using game theory and systems optimization.
- Studied data-center design and heterogeneous GPU, CPU, memory, and storage allocation for large-scale AI workloads.

Research Engineer Intern Mountain View, CA
Samsung Research America, Advanced ML Lab *May 2025 – Dec 2025*
- Built agentic systems to reduce hallucinations and improve the reasoning capabilities of internal language models.
- Researched preference-based post-training, model evaluation, and multi-agent reasoning.

Research Assistant Los Angeles, CA
USC Center for Autonomy and Artificial Intelligence *Sep 2023 – Dec 2026*
- Advised by Professor Rahul Jain; developed algorithms for multi-agent reinforcement learning and post-training alignment
- Applied game-theoretic methods to Language-model post-training, Human-AI strategic play, Censorship Games, etc.
- Developed multi-agent and multi-objective direct preference optimization methods for language-model alignment.

Founding Research Engineer Bay Area, CA
Multiple Startups including OmniGPT, FlowGPT, Ditto.AI, work in Training and Agentic AI *Aug 2023 – May 2025*
- Built personalized AI systems for a social platform designed to support friendship formation and dating experiences.
- Post-trained language models using interactions from 100+ users and 10k + documents over long context and time horizon.

Data Scientist Boston, MA
Ameriprise Financial *Feb 2023 – Aug 2023*
- Researched on data-specific training and evaluation metrics for deep learning to enhance model performance for investment

Research Assistant Cambridge, MA
MIT Computer Science and Artificial Intelligence Laboratory and Operation Research Center *Sep 2021 – Dec 2023*
- Advised by Professor Dimitris Bertsimas; developed optimization algorithms for multi-agent and counterfactual learning.
- Developed polynomial-time approximation algorithms for NP-hard optimization and online learning to improve heuristics.

Research Assistant Berkeley, CA
Berkeley Artificial Intelligence Research (BAIR) *Sep 2019 – May 2021*
- Advised by Professor James Demmel, researched model compression and routing algorithms for federated learning

PUBLICATIONS

1) **Andy (Zehao) Zhao**, Rui Liu, Mark Ibrahim, Hamed Firooz, Kavosh Asadi, Yuchen Lu, Chongyang Bai, Hang Cui, "Proactive Agent with Situation Awareness", *In Progress*, ICML 2027.
2) **Andy (Zehao) Zhao**, Rui Liu, Mark Ibrahim, Hamed Firooz, Kavosh Asadi, Yuchen Lu, Chongyang Bai, Hang Cui, "Black-Box Teachers: On-Policy Distillation with Forward KL", *In Progress*, ICML 2027.
3) **Andy (Zehao) Zhao**, Rahul Jain, "Forecast, Act, Correct: Prediction Residual-Stability with Fallible Opponent Models", *In Submission*, ICML 2027.
4) **Andy (Zehao) Zhao**, Yanyang Yang, Xiaole Zhang, Xiyu Deng, "ART-MPG: Autoregressive R/T Manifold Path Guiding", *arXiv preprint*, arXiv:2610.29049, 2026.
5) Shangjian Yin*, **Andy (Zehao) Zhao***, Kavosh Asadi, Rui Liu, Yuchen Lu, Shike Mei, Hang Cui, Luke Simon, Zhouxing Shi, Hamed Firooz, "Recursive Self-Improvement via On-Policy Distillation for Reasoning", *arXiv preprint*, arXiv:2609.30652, 2026.
6) **Andy (Zehao) Zhao**, Rui Liu, Mark Ibrahim, Hamed Firooz, Kavosh Asadi, Yuchen Lu, Chongyang Bai, Hang Cui, "Beyond update counts: A control map for evaluating test-time model updates", *arXiv preprint*, arXiv:2610.36109, 2026.
7) **Andy (Zehao) Zhao**, Rahul Jain, Vijay Subramanian, "Correlated Quantal Response Equilibrium: A New Equilibrium Notion for Human-AI Strategic Interactions", *arXiv preprint*, arXiv:2610.29020, AISTATS 2026.
8) Xiaole Zhang*, Ziyi Zhang*, **Andy (Zehao) Zhao***, Stephen Tu, Guannan Qu, Yorie Nakahira, Paul Bogdan, "Learning Fractional-Order Dynamics from a Single Trajectory", *arXiv preprint*, arXiv:2609.18127, NeurIPS 2026.
9) **Andy (Zehao) Zhao**, Apurv Shukla, Rahul Jain, Vijay Subramanian, "Efficient Decentralized Learning of Generalized Quantal Response Equilibrium", *arXiv preprint*, arXiv:2507.09928, UAI 2025.
10) Chang-Wei Yueh*, **Andy (Zehao) Zhao***, Ashutosh Nayyar, Rahul Jain, "Bayesian Learning in Episodic Zero-Sum Games", *arXiv preprint*, arXiv:2603.20604, IEEE TAC 2026.
11) Md Sakaouth Hossain, Muhammad Yasir, Ping Wang, Saleem Ullah, Maftuha Jahan, Sheng Hui**, Andy (Zehao) Zhao**, "Automatic Shoreline Extraction and Change Detection: A Study on the Southeast Coast of Bangladesh", *Marine Geology*, 2021.
12) **Andy (Zehao) Zhao**, Xiongfeng Ma, Hongyi Zhou, "Performance Optimization on Practical Quantum Random Number Generators: Modification on Min-entropy Evaluation and Acceleration on Post Processing", *arXiv preprint*, arXiv:2011.04130, 2020.
13) **Andy (Zehao) Zhao**, Hui Sheng, Xianfeng Xu, "Intelligent Detection and Removing of Straight Periodic Streaks on Digital Images", *Journal of Physics: Conference Series*, 887(1), 012078, 2017

AWARDS AND HONORS

- Y Combinator (YC'26) Startup School Award 2026
- MHI Scholar Award 2026
- Top 10 Google DeepMind Hackathon 2026
- OpenAI Research Credits 2026
- MATS Research Award 2025
- Early Career Innovator Award 2025
- USC Conference Travel Grant 2024
- Ming Hsieh Fellowship 2024
- Viterbi Graduate Fellowship 2024
- Startup accepted by MIT Accelerator 2023
- Top 10 in MIT Hackathon out of 1000+ attendees 2023
- 2nd place in MIT Analytics Lab Contest 2022
- Berkeley EECS Honors Program 2022
- Berkeley Data Science Honors Program (Department Citation) 2022
- Outstanding Graduate Student Instructor Awards (UC Berkeley) 2021
- 1st place in California actuarial league competition 2021
- Finalist for Mathematics Contest in Modeling 2018
- Putnam Competition Top 500 2018
- Research Science Institute (RSI'17) 2017
- USAMO Qualifier 2016

TALKS, PRESENTATIONS, AND PROFESSIONAL SERVICE.

- **Reviewer**: NeurIPS (2025,2026), AISTATS (2025, 2026), ICML (2025, 2026), UAI (2026)
- Invited research talks at YCML (Y Combinator) on "Human-AI Agents Co-existence: Game Theory in AI Era", 2026
- Invited research talks at Cornell University, MIT, Caltech, USC, UC Irvine on "Strategic Learning in Games: Online Learning, Quantal Response, and LLM Reasoning", 2026
- Invited talk at 47th SoCal Control Workshop on "Bayesian Learning in Episodic Zero-Sum Games", April 2026
- MATS 9.0 (Winter 2026) under Neel Nanda from Google DeepMind, "Interpretable Feature Circuits in LLM Chain-of-Thought Reasoning". Jan 2026
- **Andy (Zehao) Zhao,** Feiyu Zhu, Xiaole Zhang, Interpretable Feature Circuits in LLM Reasoning on Math Problems, NeurIPS Workshop, Dec 2025
- **Andy (Zehao) Zhao**, Apurv Shukla, Rahul Jain, Vijay Subramanian, Towards "Nash" Equilibrium for Mixed Agents, 2025 MURI/TTIC Summer Workshop on Algorithmic Game Theory
- Invited talk at Samsung Advanced ML Lab on "Evaluation and Benchmark of Agentic System and Multi-Agent LLM Post-Training". July 2025
- Presented "End-to-End Learning Framework for Solving Non-Markovian Optimal Control," ICML, July 2025; acknowledged contributor.