

Education

The Chinese University of Hong Kong Ph.D., Information Engineering. Advisor: [Angela Yingjun Zhang](#). *Aug 2026~present*

Research Areas My research primarily studies lifecycle management of data in AI models, focusing on theoretical methods and practical problems that arise as data are generated, used, and deleted. My recent work aims to improve the reliability, interpretability, and controllability of AI models in heterogeneous, computation-constrained, and communication-constrained environments.

Zhejiang University M.Eng. in Artificial Intelligence. Advised by [Meng Zhang](#).

Zhejiang. Sep 2022 ~ Dec 2025

Shandong University B.Eng. in Communication Engineering.

Shandong. Sep 2018 ~ Jul 2022

Experience

Doctoral Research @ The Chinese University of Hong Kong

Hong Kong. Aug 2026 ~ present

Investigates data lifecycle management in distributed AI across generation, use, deletion, and federated learning.

- Proposed optimal-transport federated learning with barycentric multi-prototype classification. **Paper #5**

Research on Data-Centric ML Systems @ Zhejiang University

Zhejiang. Mar 2023 ~ Dec 2025

Developed data-deletion methods for continuous influence control, certified online removal, and dynamic tree ensembles.

- Extended machine unlearning beyond binary erasure through continuous example weights, enabling targeted fairness and robustness interventions. **Paper #2**
- Built Hessian-free online certified unlearning with recollected trajectory statistics, replacing explicit Hessian inversion for streaming deletion requests. **Paper #3**
- Built exact dynamic random-forest unlearning through affected-tree updates, avoiding full retraining. **Paper #4**

Research on Trustworthy LLM Systems @ NUSRI CQ

Chongqing. Jun 2025 ~ Dec 2025

Analyzed LLM reliability failures caused by synthetic-data feedback loops and prompt patterns that trigger spurious inference.

- Showed how low-resource verification amplifies sample-selection bias in recursive synthetic-data training, persistently pruning tail data and precipitating model collapse. **Paper #1**
- Linked illusory pattern perception to spurious LLM inference when perceived prompt patterns override evidence. **Paper #6**

Open-Source Contributions and Academic Service

- 🔗 **Research code releases**, public code for certified unlearning, soft-weighted unlearning, and model-collapse work.
- 🏠 **Xinbaopedia**, public academic homepage and wiki-style research archive with papers, figures, CV, and project notes.
- 📄 **Academic service, 2026**, reviewer for ICML, NeurIPS, and AAAI.
- 📄 **Academic service, 2025**, reviewer for NeurIPS, ICLR, AAAI, and IEEE TNNLS.

Selected Publications

Asterisks (*) denote co-first authorship; daggers (†) denote corresponding authors.

- When Sample Selection Bias Precipitates Model Collapse**  *ICML 2026.*
Xinbao Qiao[†], Xianglong Du, Wei Liu, Jingqi Zhang, Peihua Mai, Meng Zhang[†], Yan Pang[†]
Shows how low-resource verification turns selection bias into persistent tail pruning and model collapse.
- Beyond Binary Erasure: Soft-Weighted Unlearning for Fairness and Robustness**  *AAAI 2026.*
Xinbao Qiao, Ningning Ding, Yushi Cheng, Meng Zhang[†]
Generalizes binary deletion into continuous influence control for targeted fairness and robustness interventions.
- Hessian-Free Online Certified Unlearning**  *ICLR 2025.*
Xinbao Qiao, Meng Zhang[†], Ming Tang, Ermin Wei
Replaces explicit Hessian inversion with recollected trajectory statistics for scalable certified deletion.
- DynFrs: An Efficient Framework for Machine Unlearning in Random Forest**  *ICLR 2025.*
Shurong Wang, Zhuoyang Shen, Xinbao Qiao, Tongning Zhang, Meng Zhang[†]
Updates affected tree statistics to deliver exact random-forest unlearning in dynamic environments.
- Federated Learning as Optimal Transport: Barycentric Multi-Prototype Classification** *Under review.*
Xinbao Qiao, Wenjing Yan[†], Ying-Jun Angela Zhang
Frames federated learning as optimal transport with barycentric multi-prototype classification under communication constraints.
- Illusory Pattern Perception Drives Spurious Inference in Large Language Models** *Under review.*
Peihua Mai, Zhuoyan Shao, Xinbao Qiao, Meng Zhang, Xinyue Zhou[†], Yan Pang[†]
Links perceived prompt regularities to spurious LLM inference and interpretable failure behavior.