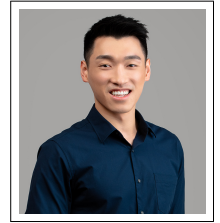


# Xinyi XU

✉ [michaelxysg@gmail.com](mailto:michaelxysg@gmail.com)  
📄 [xinyi-xu.com](http://xinyi-xu.com)  
🏠 🎓



## Research & Work Experience

- 2025/6 – **Meta**, *Research scientist*, Menlo Park, CA, U.S., on auctions and ads demand efficiency.
- 2024/10 – **Meta, Central Applied Science**, *Research scientist intern*, Menlo Park, CA, U.S., in the team of [Will Bullock](#), on markets algorithms, auctions, market efficiency and fairness.
- 2025/3
- 2024/2 – **University of California, Berkeley (UCB), EECS**, *Visiting Scholar*, Berkeley, CA, U.S., advised by [Michael I. Jordan](#), on Incentives, Data Sharing and Marketplaces.
- 2024/7
- 2022/8 – **Carnegie Mellon University (CMU), ECE**, *Visiting Scholar*, Pittsburgh, PA, U.S., advised by [Giulia Fanti](#), on Data Sharing, Federated Learning, Privacy Enhancing Technologies.
- 2023/4
- 2020/2 – **National University of Singapore (NUS)**, *Research assistant*, Singapore, advised by [Leong Tze-Yun](#), on Hierarchical Reinforcement Learning with Human Expertise in StarCraft II.
- 2020/7

## Education Background

- 2020–2025 **National University of Singapore (NUS)**, *Ph.D., Computer Science*, supported by [A\\*STAR ACIS Scholarship](#), advised by [Chuan Sheng Foo](#) and [Low, Bryan Kian Hsiang](#).
- 2015–2019 **National University of Singapore**, *Bachelor of Computing, Computer Science*, 1st Class Honours, Turing Programme, supported by [NUS Science & Technology Scholarship](#).
- 2015–2019 **National University of Singapore**, *Bachelor of Science, Applied Mathematics*.

## Research Interests & Technical Skills

Data-centric, trustworthy and fair machine learning & federated learning; data evaluation and optimization for LLMs. Market algorithms, fairness, differential privacy, game theory, Bayesian methods.

## Honors and Awards

- 2024 **Dean's Graduate Research Excellence Award** 2023/24 Semester 2, School of Computing, NUS (1 of 11 awarded).
- 2024 **Research Achievement Award** 2023/24 Semester 1, School of Computing, NUS.
- 2023 **Teaching Fellowship Scheme** of School of Computing, NUS (1 of 5 selected CS Ph.D. students).
- 2023 1 of 5 nominated for **Apple Scholars** in AI/ML 2023 Ph.D. Fellowship by School of Computing, NUS.
- 2022 **Research Achievement Award** 2021/22 Semester 1, School of Computing, NUS.
- 2021 **Honor List of Student Tutor** 2020/21 Semester 1 (list in pdf), School of Computing, NUS.
- 2020 **Best Paper Award** at International Workshop on Federated Learning for User Privacy and Data Confidentiality in Conjunction with IJCAI 2020 (FL-IJCAI'20).
- 2020–2024 **A\*STAR Computing and Information Science (ACIS) Scholarship** of Agency for Science, Technology and Research (A\*STAR), as a Ph.D. research student in School of Computing, NUS.
- 2015–2019 **Science and Technology Scholarship** of NUS, Singapore, as an undergraduate at NUS.

2010–2014 **School-based Scholarship** (formerly SM1 scholarship) of Ministry of Education, Singapore, at Hwa Chong Institution high school and junior college.

## Professional Service

Conference Reviewer for ICML 2022-2024, NeurIPS 2022-2024, ICLR 2023-2025, AAMAS 2023-2025, AISTATS 2023-2025, AAAI 2024 2025, IJCAI 2024. Post-conference Workshop Proposal Reviewer for ICML 2024.

**Outstanding Reviewer** (Top 10%) for ICML 2022; **Top Reviewer** for NeurIPS 2022, NeurIPS 2023; **Top Reviewer** for AISTATS 2023 (Top 10%).

## Invited Talks

2022 Google - Collaborative machine learning/Federated learning [[Slides](#)].

2022 2022 Workshop on Federated Learning and Analytics by Google - Beyond Standard FL [[Slides](#)].

2024 Federated Learning One World Seminar (FLOW) - Fairness and incentives in data sharing & collaborative learning [[Slides](#)].

## Refereed Conference Publications

\* indicates equal contributions. For most updated list, please refer to my [Google scholar](#).

- [C1] Arun Verma, Zhaoxuan Wu, ... **Xinyi Xu**, and ... . “[Position Paper: Uncover Scaling Laws for Large Language Models via Inverse Problems](#)”. In: *Proceedings of the 2025 Empirical Methods in Natural Language Processing (EMNLP-25) Findings, Suzhou, China, Nov. 5, 2025*. 2025.
- [C2] Xiaoqiang Lin, **Xinyi Xu**, See-Kiong Ng, and Bryan Kian Hsiang Low. “[Top- \$m\$  Data Values Identification](#)”. In: *The Thirteenth International Conference on Learning Representations (ICLR-25) Singapore, Apr. 24-28, 2025*. (Acceptance rate: 31%).
- [C3] **Xinyi Xu**, Shuaiqi Wang, Chuan-Sheng Foo, Bryan Kian Hsiang Low, and Giulia Fanti. “[Data Distribution Valuation](#)”. In: *Advances in Neural Information Processing Systems 37: 38th Annual Conference on Neural Information Processing Systems (NeurIPS’24), Vancouver, Dec. 9-15, 2024*. (Acceptance rate: 25.8%).
- [C4] Zijian Zhou, Xiaoqiang Lin, **Xinyi Xu**, Alok Prakash, Daniela Rus, and Bryan Kian Hsiang Low. “[DETAIL: Task DEmonsTration Attribution for Interpretable In-context Learning](#)”. In: *Advances in Neural Information Processing Systems 37: 38th Annual Conference on Neural Information Processing Systems (NeurIPS’24), Vancouver, Dec. 9-15, 2024*. (Acceptance rate: 25.8%).
- [C5] **Xinyi Xu** et al. “[Data-Centric AI in the Age of Large Language Models](#)”. In: *Proceedings of the 2024 Empirical Methods in Natural Language Processing (EMNLP-24) Findings, Miami, Florida, Nov. 12-16, 2024*. 2024.
- [C6] Xiaoqiang Lin, **Xinyi Xu**, Zhaoxuan Wu, See-Kiong Ng, and Bryan Kian Hsiang Low. “[Distributionally Robust Data Valuation](#)”. In: *Proceedings of the 41st International Conference on Machine Learning (ICML-24), Vienna, Jul. 21-27, 2024*. (Acceptance rate: 27.5%).
- [C7] **Xinyi Xu**, Thanh Lam, Chuan-Sheng Foo, and Bryan Kian Hsiang Low. “[Model Shapley: Equitable Model Valuation with Black-box Access](#)”. In: *Advances in Neural Information Processing Systems 36: 37th Annual Conference on Neural Information Processing Systems (NeurIPS’23), New Orleans, Dec. 10-16, 2023*. (Acceptance rate: 26%).
- [C8] Rachael Hwee Ling Sim, Yehong Zhang, Trong Nghia Hoang, **Xinyi Xu**, Bryan Kian Hsiang Low, and Patrick Jaillet. “[Incentives in Private Collaborative Machine Learning](#)”. In: *Advances in Neural Information Processing Systems 36: 37th Annual Conference on Neural Information Processing Systems (NeurIPS’23), New Orleans, Dec. 10-16, 2023*. (Acceptance rate: 26%).
- [C9] Rui Qiao, **Xinyi Xu**, and Bryan Kian Hsiang Low. “[Collaborative Causal Inference with Fair Incentives](#)”. In: *Proceedings of the 40th International Conference on Machine Learning (ICML-23), Hawaii, Jul. 23-29, 2023*. (Acceptance rate: 28%).

- [C10] Xiaoqiang Lin\*, **Xinyi Xu\***, See-Kiong Ng, Chuan Sheng Foo, and Bryan Kian Hsiang Low. “[Fair yet Asymptotically Equal Collaborative Learning](#)”. In: *Proceedings of the 40th International Conference on Machine Learning (ICML-23)*, Hawaii, Jul. 23-29, 2023. (Acceptance rate: 28%).
- [C11] **Xinyi Xu**, Zhaoxuan Wu, Arun Verma, Chuan Sheng Foo, and Bryan Kian Hsiang Low. “[FAIR: Fair Collaborative Active Learning with Individual Rationality for Scientific Discovery](#)”. In: *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS-23)*, Valencia, Spain, Apr. 25 - 27, 2023. (Acceptance rate: 29%).
- [C12] Zijian Zhou\*, **Xinyi Xu\***, Rachael Hwee Ling Sim, Chuan Sheng Foo, and Bryan Kian Hsiang Low. “[Probably Approximate Shapley Fairness with Applications in Machine Learning](#)”. In: *Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI-23) Oral Presentation*, Washington DC, Feb. 7-14, 2023. (Acceptance rate: 19.6%).
- [C13] Lucas Agussurja, **Xinyi Xu**, and Bryan Kian Hsiang Low. “[On the Convergence of the Shapley Value in Parametric Bayesian Learning Games](#)”. In: *Proceedings of the 39th International Conference on Machine Learning (ICML-22)*, Baltimore, Jul. 17-23, 2022. (Acceptance rate: 21.9%).
- [C14] Rachael Hwee Ling Sim\*, **Xinyi Xu\***, and Bryan Kian Hsiang Low. “[Data Valuation in Machine Learning: Ingredients, Strategies, and Open Challenges](#)”. In: *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI-22)*, Vienna, Jul. 23-29, 2022. (Acceptance rate: 18.2%). Survey track.
- [C15] Sebastian Sheng Hong Tay, **Xinyi Xu**, Chuan Sheng Foo, and Bryan Kian Hsiang Low. “[Incentivizing Collaboration in Machine Learning via Synthetic Data Rewards](#)”. In: *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI-22) Oral Presentation*, Feb. 22-Mar. 1, 2022. (Acceptance rate: 4%).
- [C16] **Xinyi Xu\***, Lingjuan Lyu\*, Xingjun Ma, Chenglin Miao, Chuan Sheng Foo, and Bryan Kian Hsiang Low. “[Gradient Driven Rewards to Guarantee Fairness in Collaborative Machine Learning](#)”. In: *Advances in Neural Information Processing Systems 34: 35th Annual Conference on Neural Information Processing Systems (NeurIPS'21)*, Dec. 6-14, 2021. (Acceptance rate: 26%).
- [C17] **Xinyi Xu\***, Zhaoxuan Wu\*, Chuan Sheng Foo, and Bryan Kian Hsiang Low. “[Validation Free and Replication Robust Volume-based Data Valuation](#)”. In: *Advances in Neural Information Processing Systems 34: 35th Annual Conference on Neural Information Processing Systems (NeurIPS'21)*, Dec. 6-14, 2021. (Acceptance rate: 26%).

## Refereed Journal/Workshop Publications

- [J1] **Xinyi Xu**, Shuaiqi Wang, Chuan-Sheng Foo, Bryan Kian Hsiang Low, and Giulia Fanti. “[Data Distribution Valuation](#)”. In: *Data-centric Machine Learning Research (DMLR) Workshop at ICLR (DMLR@ICLR'24)*. 2024.
- [J2] **Xinyi Xu** and Lingjuan Lyu. “[A Reputation Mechanism Is All You Need: Collaborative Fairness and Adversarial Robustness in Federated Learning](#)”. In: *International Workshop on Federated Learning for User Privacy and Data Confidentiality in Conjunction with ICML 2021 (FL-ICML'21) Oral Presentation*. 2021.
- [J3] Lingjuan Lyu, **Xinyi Xu**, and Qian Wang. “[Collaborative Fairness in Federated Learning](#)”. In: *IJCAI 2020. International Workshop on Federated Learning for User Privacy and Data Confidentiality in Conjunction (FL-IJCAI'20) Best Paper Award*. 2020.