Charlie Hou
AI Research Scientist/Engineer

houcharlie.github.io | LinkedIn | Twitter
hou.charlie2@gmail.com

Summary

I am an AI research scientist at Meta. I work on topics including (1) developing LLM reasoning models for trust and safety, and (2) improving the efficiency and quality of text extraction models (OCR/ASR). This work has been launched, leading to robust time spent wins in Instagram.

I obtained my Ph.D. from CMU advised by Prof. Giulia Fanti. There I led the development of POPri, a step-change improvement in synthetic data quality generation under privacy constraints. I also co-led SquirRL which pioneered the use of reinforcement learning (RL) for identifying security weaknesses in blockchain protocols.

Work Experience

Meta (Bellevue, WA) — AI Research Scientist (Full-time) *Jun 2024 – Present*

- Launched a high-quality OCR feature that increased Instagram (IG) time spent by 0.1% (for context, IG time spent in total increased by 6% in Q2 2025).
- Launched retrieval model pretraining which increased Instagram impressions by 0.2%.
- Deployed an LLM reasoning classifier that achieves 90% recall and 95% precision on highly problematic disgusting/unsettling content for trust and safety, unblocking 1.23% time spent for Instagram US new users.

Meta (Redmond, WA) — AI Research Scientist (Intern) *Jun 2023 – Sep 2023*

- Published PrE-Text (ICML 2024 Oral) paper on synthetic data generation under privacy constraints.

Meta (Redmond, WA) — AI Research Scientist (Intern) *Sep 2022 – Feb 2023*

- Published FrED paper on privately choosing pretraining datasets for federated learning.

Amazon (Palo Alto, CA) — Applied Scientist (Intern) *Jun 2022 – Sep 2022*

- Published ranking pretraining paper which was accepted to TMLR.

Talks/Media

2025 Aug	Invited talk on POPri at **OpenAI**
2025 May	Invited talk on POPri at **Amazon**
2025 Apr	Oral presentation on POPri at SynthData @ ICLR 2024 workshop
2024 Jul	Oral presentation on PrE-Text at ICML 2024
2023 Jul	Oral presentation on pretraining in LTR at MFPL @ ICML 2023 workshop
2021 Jul	Oral presentation on FedChain at FL @ ICML 2021 workshop

Publications (selected)

[ICML 2025] **POPri: Private Federated Learning using Preference-Optimized Synthetic Data**.
Charlie Hou, Mei-Yu Wang, Yige Zhu, Daniel Lazar, Giulia Fanti.
paper | code

[ICML 2024] **PrE-Text: Training Language Models on Private Federated Data in the Age of LLMs**.
(ORAL, TOP 1% PAPER) Charlie Hou, Akshat Shrivastava, Hongyuan Zhan, Rylan Conway, Trang Le, Adithya Sagar, Giulia Fanti, Daniel Lazar.
paper | code

UNDER REVIEW	**On the Convergence of Differentially-Private Fine-tuning: To Linearly Probe or to Fully Fine-tune?** Shuqi Ke, Charlie Hou, Giulia Fanti, Sewoong Oh. paper
[TMLR]	**Pretrained deep models outperform GBDTs in Learning-To-Rank under label scarcity**. Charlie Hou, Kiran Thekumparampil, Michael Shavlovsky, Giulia Fanti, Yesh Dattatreya, Sujay Sanghavi. paper
[ICLR 2022]	**FedChain: Chained Algorithms for Near-Optimal Communication Cost in Federated Learning**. Charlie Hou, Kiran Thekumparampil, Giulia Fanti, Sewoong Oh. *arXiv*, 2021. paper
[NDSS 2021]	**SquirRL: Automating Attack Analysis on Blockchain Incentive Mechanisms with Deep RL**. Charlie Hou, Mingxun Zhou, Yan Ji, Phil Daian, Florian Tramèr, Giulia Fanti, Ari Juels. paper | code
WORKSHOP	**Privately Customizing Prefinetuning to Better Match User Data in Federated Learning**. Charlie Hou, Hongyuan Zhan, Sid Wang, Aleksandr Livshits, Giulia Fanti, Daniel Lazar. paper

Awards

- Best paper runner-up at SynthData@ICLR 2025 workshop for POPri
- Best paper honorable mention at ICLR 2024 PrivML workshop for PrE-Text
- Google Collabs 2022 Research Award ($100k grant)
- Winner of Tiger Chef challenge at Princeton

Education

Carnegie Mellon University *2019 – 2024*
Ph.D., Electrical and Computer Engineering
Advisor: Prof. Giulia Fanti

Princeton University *2015 – 2019*
B.S.E., Operations Research and Financial Engineering