Ashwinee Panda

website Email

Education

20 - 24	PhD **Princeton University** advised by Prateek Mittal
20	M.S. + B.S. **UC Berkeley** advised by Joey Gonzalez

Work

24 - 25	**Postdoc** at UMD College Park with Tom Goldstein
24	Applied Research Intern at Capital One Research
18 - 22	CEO at DiscreetAI

Research (selected 1st-author conference publications)

Phishing	**Ashwinee Panda**, Christopher Choquette-Choo, Zhengming Zhang, Yaoqing Yang, Prateek Mittal Teach LLMs to Phish: Stealing Private Information from Language Models ICLR 2024 Poster
DP-ICL	Tong Wu*, **Ashwinee Panda***, Tianhao Wang*, Prateek Mittal Privacy-Preserving In-Context Learning for Large Language Models ICLR 2024 Poster
DP-RandP	Xinyu Tang*, **Ashwinee Panda***, Prateek Mittal Differentially Private Image Classification by Learning Priors from Random Processes NeurIPS 2023 Spotlight
Neurotoxin	Zhengming Zhang*, **Ashwinee Panda***, Linyue Song, Yaoqing Yang, Prateek Mittal, Joseph Gonzalez, Kannan Ramchandran, Michael Mahoney NeuroToxin: Durable Backdoors in Federated Learning ICML 2022 Oral
SparseFed	**Ashwinee Panda**, Saeed Mahloujifar, Arjun Bhagoji, Supriyo Chakraborty, Prateek Mittal SparseFed: Mitigating Model Poisoning Attacks in Federated Learning via Sparsification AISTATS 2022 Poster
FetchSGD	Daniel Rothchild*, **Ashwinee Panda***, Enayat Ullah, Nikita Ivkin, Ion Stoica, Vladimir Braverman, Joseph Gonzalez, Raman Arora FetchSGD: Communication-Efficient Federated Learning with Sketching ICML 2020 Poster

Invited Talks

Nov '23	New Privacy Attacks on Large Language Models *Sun Lab, Berkeley*
Nov '23	Challenges in Data-Driven Alignment of Large Language Models *SPYLab, ETH Zurich*
Oct '23	New Directions in Differentially Private Machine Learning *Meta CAS*
Sep '23	Challenges in Data-Driven Alignment of Large Language Models *University of Maryland, College Park*
Sep '23	Challenges in Augmenting Large Language Models with Private Data *SL^2 Lab, UIUC*
Sep '23	Improving the Privacy Utility Tradeoff in Differentially Private Machine Learning with Prior Information *SECRIT, University of Michigan*
Apr '23	Improving the Privacy Utility Tradeoff in Differentially Private Machine Learning with Public Data *Apple*
Mar '23	Google Privacy Seminar (click for talk recording) *Google*
Jun '22	Challenges and Directions in Privacy Preserving Machine Learning *Microsoft Research Cambridge*
May '22	Towards Trustworthy Machine Learning *Meta AI*
Jan '22	Federated Learning for Forecasting *Ohmconnect*
Nov '21	Building Federated Learning Systems at Scale *Liftoff AI*
Nov '21	Practical Defenses Against Model Poisoning Attacks (click for talk) *Google Federated Learning Workshop*

Research (Advised and workshop papers)

DP-ZO	Xinyu Tang*, **Ashwinee Panda***, Milad Nasr, Saeed Mahloujifar, Prateek Mittal Private Fine-tuning of Large Language Models with Zeroth-order Optimization
AdvVLM	Xiangyu Qi*, Kaixuan Huang*, **Ashwinee Panda**, Mengdi Wang, Prateek Mittal Introducing Vision into Large Language Models Expands Attack Surfaces and Failure Implications At *Thirty-Eighth AAAI Conference on Artificial Intelligence*
Phishing	**Ashwinee Panda**, Zhengming Zhang, Yaoqing Yang, Prateek Mittal Teach GPT to Phish: Neural Phishing Attacks on Large Language Models At *40th International Conference on Machine Learning* AdvML Workshop
DP-Diffusion	Vikash Sehwag*, **Ashwinee Panda***, Ashwini Pokle, Xinyu Tang, Saeed Mahloujifar, Mung Chiang, J Zico Kolter, Prateek Mittal Differentially Private Generation of High Fidelity Samples From Diffusion Models At *40th International Conference on Machine Learning* GenAI Workshop
DP-ICL	**Ashwinee Panda***, Tong Wu*, Tianhao Wang*, Prateek Mittal Differentially Private In-Context Learning At *NAACL 2023* TrustNLP Workshop
DP-Lin	**Ashwinee Panda***, Xinyu Tang*, Vikash Sehwag, Saeed Mahloujifar, Prateek Mittal A New Linear Scaling Rule for Differentially Private Hyperparameter Optimization *Submitted to NeurIPS 2023*
SoftPBT	**Ashwinee Panda**, Eric Liang, Richard Liaw, Joey Gonzalez SoftPBT: Leveraging Experience Replay for Efficient Hyperparameter Schedule Search *Submitted to NeurIPS 2019*

Service

Teaching

2023	Teaching Assistant for COS/ECE 432 at Princeton
2019	Course Staff for CS 189 (Machine Learning) at UC Berkeley
2018	Undergraduate Student Instructor for CS 70 (Probability and Discrete Mathematics) and Course Staff for CS 189 at UC Berkeley
2017	Course Staff for CS 70 at UC Berkeley

Peer Reviewing

2023	SATML 2023, ACL 2023, ICML 2023, NeurIPS 2023, TMLR
2022	ICML 2022, AISTATS 2022
2021	ICML 2021, NeurIPS 2021
2020	ICML 2020
2019	ICLR 2019, NeurIPS 2019