

Anvith Thudi

anvith.com

✉️ anvith.thudi@mail.utoronto.ca

Education

University of Toronto

Ph.D. in Computer Science

Toronto, ON, Canada

Sep. 2022 - ongoing

- Advisors: Nicolas Papernot and Chris Maddison

University of Toronto

B.Sc in Mathematics, Spent Fall 2020 in Engineering Science

Toronto, ON, Canada

Sep. 2020 - May 2022

- GPA: 3.92/4.0

Simon Fraser University

Concurrent Studies Student (attended while in highschool)

Burnaby, BC, Canada

Sep. 2017 - May 2020

- GPA: 4.09/4.33

Awards and Honours

Notable Reviewer: ICLR 2025

2023 Canada Graduate Scholarship-Doctoral: NSERC

- *declined due to Vanier*

2023 Vanier Canada Graduate Scholarship: NSERC

- *Rank 1/173 of national round nominees (Ph.D. students in the Natural Sciences or Engineering)*

Doctoral Entrance Scholarship: UofT Department of Computer Science

Doctoral Recruitment Award: UofT Faculty of Arts and Science

Galois Award: University College UofT

Dean's List Scholar: UofT

Dean's Honours List: UofT

2020 Loran Scholarship National Finalist: Loran Scholar's Foundation

- *Top 88 highschool students in Canada*

Publications

Journal Proceedings

"k-Nearest Neighbour Adaptive Sampling (kNN-AS), a Simple Tool to Efficiently Explore Conformational Space": *Evianne M. Rovers, Anvith Thudi, Jérôme Hénin, Chris Maddison, Matthieu Schapira. Journal of Chemical Theory and Computation*

"From Differential Privacy to Bounds on Membership Inference: Less can be More": *Anvith Thudi, Ilia Shumailov, Franziska Boenisch, Nicolas Papernot. Transactions on Machine Learning Research*

"Selective Classification via Neural Training Dynamics": *Stephan Rabanser, Anvith Thudi, Kimia Hamidieh, Adam Dziedzic, Nicolas Papernot. Transactions on Machine Learning Research*

Conference Proceedings

"Gauss-Newton Unlearning for the LLM Era": *Lev McKinney, Anvith Thudi, Juhan Bae, Tara Rezaei Kheirkhah, Nicolas Papernot, Sheila A. McIlraith, Roger Baker Grosse. Proceedings of the 4th IEEE Conference on Secure and Trustworthy Machine Learning*

"Leveraging Per-Instance Privacy for Machine Unlearning": Nazanin Mohammadi Sepahvand, Anvith Thudi, Berivan Isik, Ashmita Bhattacharyya, Nicolas Papernot, Eleni Triantafillou, Daniel M. Roy, Gintare Karolina Dziugaite. Proceedings of the 42nd International Conference on Machine Learning. Oral at TPDP workshop 2025

"Fast Exact Unlearning for In-context Learning Data for LLMs": Andrei Muresanu, Anvith Thudi, Michael R. Zhang, Nicolas Papernot. Proceedings of the 42nd International Conference on Machine Learning

"MixMin: Finding Data Mixtures via Convex Minimization": Anvith Thudi, Eviianne Rovers, Yangjun Ruan, Tristan Thrush, Chris J. Maddison. Proceedings of the 42nd International Conference on Machine Learning

"MixMax: Distributional Robustness in Function Space via Optimal Data Mixtures": Anvith Thudi, Chris J. Maddison. Proceedings of the 13th International Conference on Learning Representations

"Gradients Look Alike: Sensitivity is Often Overestimated in DP-SGD": Anvith Thudi, Hengrui Jia, Casey Meehan, Ilia Shumailov, Nicolas Papernot. Proceedings of the 33rd USENIX Security Symposium, 2024

"Better Sparsifiers for Directed Eulerian Graphs": Sushant Sachdeva, Anvith Thudi, Yibin Zhao. Proceedings of the 51st EATCS International Colloquium on Automata, Languages and Programming

"Training Private Models That Know What They Don't Know": Stephan Rabanser, Anvith Thudi, Abhradeep Thakurta, Krishnamurthy Dvijotham, Nicolas Papernot. Proceedings of the 37th Conference on Neural Information Processing Systems

"Proof-of-Learning is Currently More Broken Than You Think": Congyu Fang, Hengrui Jia, Anvith Thudi, Mohammad Yaghini, Christopher A. Choquette-Choo, Natalie Dullerud, Varun Chandrasekaran, Nicolas Papernot. Proceedings of the 8th IEEE European Symposium on Security and Privacy, 2023

"On the Necessity of Auditable Algorithmic Definitions for Machine Unlearning": Anvith Thudi, Hengrui Jia, Ilia Shumailov, Nicolas Papernot. Proceedings of the 31st USENIX Security Symposium, 2022

"Unrolling SGD: Understanding Factors Influencing Machine Unlearning": Anvith Thudi, Gabriel Deza, Varun Chandrasekaran, Nicolas Papernot. Proceedings of the 7th IEEE European Symposium on Security and Privacy, 2022

"Proof of Learning: Definitions and Practice": Hengrui Jia, Mohammad Yaghini, Christopher A. Choquette-Choo, Natalie Dullerud, Anvith Thudi, Varun Chandrasekaran, Nicolas Papernot. Proceedings of the 42nd IEEE Symposium on Security and Privacy, 2021

Preprints

"Efficient Public Verification of Private ML via Regularization": Zoë Ruha Bell, Anvith Thudi, Olive Franzese-McLaughlin, Nicolas Papernot, Shafi Goldwasser

"SoK: Machine Learning Governance": Varun Chandrasekaran, Hengrui Jia, Anvith Thudi, Adelin Travers, Mohammad Yaghini, Nicolas Papernot

Experience

Microsoft Research Cambridge
Ph.D. Research Intern

Cambridge, UK
May. 2023 - July 2023

Talks

"Leveraging Per-Instance Privacy for Machine Unlearning": Vector's ICML 2025 Conference Highlights

"Leveraging Per-Instance Privacy for Machine Unlearning": Google DeepMind

"Making Datasets from Multiple Data Distributions": Social Foundations of Computation at MPI Tübingen

"Making Datasets from Multiple Data Distributions": University of British Columbia

"Unlearning Can Be Easy": University of Wisconsin-Madison Security and Privacy Seminar

"Datapoints that are Easy to Unlearn": Google DeepMind

"Gradients Look Alike: Sensitivity is Often Overestimated in DP-SGD": USENIX Security 24'

"Datapoints that are Easy to Unlearn": Harvard Efficient ML Seminar

"The Unlearning Problem(s)": CS 562 at University of Illinois Urbana-Champaign

"The Unlearning Problem(s)": The Alan Turing Institute

"The Unlearning Problem(s)": Cambridge

"The Unlearning Problem(s)": Google

"The Unlearning Problem(s)": EPFL

"The Unlearning Problem(s)": ETH Zurich

"On the Necessity of Auditable Algorithmic Definitions for Machine Unlearning": USENIX Security 22'

"Unrolling SGD: Understanding Factors Influencing Machine Unlearning": Euro S&P 22'

"The Unlearning Problem(s)": Meta

Service

Reviewer: Euro S&P (2022), ICLR (2025, 2026), ICML (2025, 2026), L2M2 Workshop at ACL (2025), Neurips (2025)

Subreviewer: IEEE S&P (2024), CCS (2023), Neurips (2022)

Panel: Neurips 2023 Unlearning Competition

Organizer: ML Lunch Talks at Vector (2024 - ongoing)