Roy Rinberg

www.royrinberg.com
royrinberg+CV@gmail.com
Github: RoyRin

EDUCATION		
Harvard University, Cambridge, MA		2023 - PRESENT
PhD. Computer Science. Advisors: Prof. Seth Neel and Prof. Salil Vadhan		
Columbia University, New York, NY		2021 - 2023
MS in Computer Science [Thesis Track]. Advisors: Prof. Rachel Cummings and Prof. Steven Bellovin		
New York University, New York, NY		2014 - 2018
B.A. Computer Science, Physics, Minor: Math.		
Thomas Jefferson High School for Science and Technology, Alexandria, VA		2010 - 2014

RESEARCH EXPERIENCE

Harvard University, Cambridge, MA — AUG. 2023 - PRESENT
Trustworthy Machine Learning [Advisors: Prof. Seth Neel and Prof. Salil Vadhan]
- Research on foundations of machine learning and fundamentals of Differential Privacy (DP).
- Research on Machine Unlearning.

Columbia University, New York, NY — AUG. 2021 - AUG. 2023
Privacy in ML [Advisors: Prof. Rachel Cummings and Prof. Steven Bellovin]
- Extensions of Gaussian & Laplace DP primitives, and their application to ML. *In Submission.*
- Research on Catered PATE - PATE in the presence of heterogenous data (link). *On-going.*
- Research on how ML algorithms memorize training data.

Vector Institute, Toronto, ON — MAY 2022 - SEP. 2022
Privacy in Machine Learning [Advisor: Prof. Nicolas Papernot]
- Research on Individualization of PATE (PoPETs 2023) and DP-SGD (Neurips 2024).
- Research on reducing distributional and user-preference-level assumptions in private ML.

New York University, New York, NY — FEB. 2017 - MAY 2018
Evolution of Language Models within Social Networks [Advisor: Prof. Bud Mishra]
- Studied the development of echo chambers within social networks using TDA to study distances between Word2Vec models trained on Reddit text. Preprint on arXiv.

PAPERS

1. **R. Rinberg**, K. Georgiev, S. Park, S. Garg, A. Ilyas, A. Madry, S. Neel. *Attribute-to-Delete: Machine Unlearning via Datamodel Matching.* (2024). ICLR 2025.
2. **R. Rinberg**, Ilia Shumailov, Rachel Cummings, Nicolas Papernot. *Beyond Laplace and Gaussian: Exploring the Generalized Gaussian Mechanism for Private Machine Learning.* Preprint.
3. F. Boenisch, C Mühl, A. Dziedzic, **R. Rinberg**, N. Papernot. *Have it your way: Individualized Privacy Assignment for DP-SGD.* Accepted to Neurips 2023.
4. F. Boenisch, C Mühl, **R. Rinberg**, J. Ihrig, A. Dziedzic. *Individualized PATE: Differentially Private Machine Learning with Individual Privacy Guarantees.* Accepted to PoPETs 2023.
5. **R. Rinberg**, N. Agarwal. *Privacy when Everyone is Watching: An SOK on Anonymity on the Blockchain.* ePrint.
6. A. Tamaskar, **R. Rinberg**, S. Chakraborty, B. Mishra. *Creolizing the Web.* arXiv.

WORKSHOPS

1. **R. Rinberg** and M. Pawelczyk. *When is Differentially Private Finetuning Actually Private?* In: SFLLM (NeurIPS 2024 Workshop). 2024. workshop paper and associated Blog.

TEACHING AND SERVICE

1. Harvard - CS2881 (AI Safety) Teaching Fellow	FALL 2025
2. Harvard - CS1200 (Intro to Algorithms) Head Teaching Fellow	FALL 2024
3. OSDI (USENIX) 2023 - Assistant organizer for PC	SPRING 2023
4. NYU - General Physics I and II Tutor	2017 - 2018

ML Alignment and Theory Scholar (MATS), Berkeley, CA　　　SUMMER 2025
Research Scholar
MATS is an independent research program for research on AI safety
- Worked with Keri Warr at Anthropic on detecting and preventing model weight exfiltration.

Shelton AI, New York, NY　　　JAN. 2022 - JUN. 2022
Lead Software Engineer
Shelton AI leverages machine learning to help pension funds manage investments in private equity.
- Worked with CEO to develop fintech product to manage 10s of millions of dollars.
- Developed core AWS infrastructure for NLP document processing pipeline.

Ouster, San Francisco, CA　　　JUN. 2018 - JUL. 2021
Software Engineer
Ouster is a company developing lidar sensors. I was an early-stage engineer working on lidar-based collision-avoidance systems.
- Led development of on-edge computing for live predictions about dangerous driving.
- Developed platforms for evaluating algorithms on historical lidar data and monitoring live data.
- *Internship Project:* Produced open-source C++ lidar point-cloud data visualizer (Github link).

Career Copilots, San Francisco, CA　　　MAY 2020 - AUG. 2020
Software Engineer
Career Copilots is a startup seeking to help individuals find jobs using Linkedin data.
- Developed web-scraping data-exploration pipeline of jobs-data to help users find relevant roles.

Knight First Amendment Institute, New York, NY　　　SEPT. 2022 - MAY 2023
Algorithmic Amplification in Society [Advisor: Professor Arvind Narayaran]
KFAI works to protect digital freedoms through strategic litigation, research, and education.
- Work with Professor Arvind Narayaran to develop essays, videos, and interactives for explaining how algorithmic amplification can affect speech online.

Student Leaders in AI (Berkman Klein Center)	2025
Columbia Advanced Master's Research Specialization	2022-2023
Presidential Honors Scholar (NYU)	2015 - 2018
Dean's List (NYU)	2014 - 2018
Sigma Pi Sigma (Physics Honor Society) (NYU)	INDUCTED 2018
HPC for Undergraduates - Conference Scholarship for SC'17	FALL 2017
DURF & Research+ for Housing and Stipend (NYU)	SUMMER 2017

Technically Private (Organizer)　　　2021-PRESENT
- Organize group of inter-university graduate students in the privacy and security spaces, across legal, policy, and technical domains.

Mentor Ukraine (Mentor)　　　2022-2023
- I mentored three Ukrainian students and assisted them in applying to university in the US and Europe. One of whom received a full ride to Bard University.

Project BEST (Building Excitement for Science and Tech) (Cofounder)　　　2011-2014
- Fundraised and grew organization to 25 chapters across 3 states, reaching 3000+ students.
- Led two full-day STEM programs for 100+ students, and co-led team of 20 volunteers.

Ouster Community Work　　　2018-2020
- Advocated management to donate $6k to 6 public-interest orgs and institute paid volunteer-day.