Christopher A. Choquette-Choo

🌐 christopherchoquette.com
✉ choquette.christopher@gmail.com
in christopher-choquette-choo

📞 +1 408-442-7846
⭘ cchoquette
🏴 CA, USA

Research Scientist   *with 40+ papers*
as well as direct experience deploying my work into 6+ products and indirectly into 50+ downstream products.
I am a scientist and engineer. I enjoy defining then solving tough problems, and deploying the solutions.

Research Experience

Google Brain & Google DeepMind   *Mountain View, CA, USA*
Research Scientist   *2024 – Present*
Machine Learning Researcher   *2022 – 2023*

- Lead memorization analysis in large language models. Research how memorization manifests.
- Analyze and ensure product compliance of LLMs for release—enabled 8+ launches and 100s of products, including Gemini (+1.5 Pro/Flash), Gemma (+CodeGemma), GBoard, and PaLM 2.
- Research privacy/security vulnerabilities and auditing of machine learning and language models.
- Research and develop state-of-the-art differential privacy mechanisms for machine learning.
- Lead research into compression in federated learning.
- Deploy my techniques for compression, memorization analysis, and differential privacy into production.
- 8 spot bonuses for exceptional work, including LLM releases like Gemini (+1.5 Pro), Gemma, and PaLM 2, as well as impactful reserach like DP-FTRL and attacking SOTA models like GPT-3.
- 500+ CLs, 1 competition, 20+ papers released to date.

Google Research, Cerebra team   *New York, NY, USA*
Brain Resident   *2020 – 2022*

- Investigated concept interpretability of acoustic models. Presented at Google Research Conference.
- Led research into optimal privacy-communication-accuracy tradeoffs with sparsity in federated learning.
- Researched differentially private multi-winner voting mechanisms for machine learning.
- Guided and advise project into private semi-supervised learning for federated learning in dermatology.

Vector Institute, with Professor Nicolas Papernot   *Toronto, ON, Canada*
Research Assistant   *Sept 2019 – Oct. 2020*

- Led research into differentially private collaborative algorithms.
- Led Privacy-preserving machine learning.

Georgian Partners   *Toronto, ON, Canada*
Research Engineer   *Apr. 2019 – Aug. 2019*

- Owned development of a differentially private ML model, to guarantee user data privacy, in collaboration with Google's top machine learning library, TensorFlow/Privacy, which is used by 1000 people.
- Designed an AutoML package to intelligently tune an ML model on any dataset; used by 25+ people.

Vector Institute, with Professor Aspuru-Guzik   *Toronto, ON, Canada*
Undergraduate Researcher   *Apr. 2019 – Aug. 2019*

- Researched machine learning for molecular discovery via Gaussian processes and active learning.

Intel Corp.   *Toronto, ON, Canada*
Research Engineer   *May 2018 – May 2019*

- Spearheaded SOTA ML bug triager with 55% accuracy on 2000+ engineers and 76% on 500+ teams.
- Productionized triager with an engineering efficiency improvement of 25% and savings of >$10M annually.

Institute of Biomaterials and Biomedical Engineering with Professor Paul Santerre   *Toronto, ON, Canada*
Undergraduate Researcher   *Apr. 2016 – Sept. 2016*

- Studied mechanical properties of polyurethane scaffolds and dental resin composites. Used in patents.

Research and Papers

[X] = First or Co-First Author. To date, I've first or co-first authored 13 papers.

Peer-Reviewed Conference and Journal Proceedings

[42] *Auditing Private Prediction* Link   *2024*
Proceedings of the 41st International Conference on Machine Learning (ICML)

   Karan Chadha, Matthew Jagielski, Nicolas Papernot, **Christopher A. Choquette-Choo**, Milad Nasr

[41] *Privacy Side-Channels in Machine Learning* Link   *2024*
USENIX Security Symposium (USENIX)

   Edoardo Debenedetti, Giorgio Severi, Milad Nasr, **Christopher A. Choquette-Choo**, Matthew Jagielski, Eric Wallace, Nicholas Carlini, Florian Tramèr

[40] *Privacy Amplification for Matrix Mechanisms* Link   *2024*
(Spotlight) International Conference on Learning Representations (ICLR)

   Christopher A. Choquette-Choo, Arun Ganesh, Thomas Steinke, Abhradeep Guha Thakurta

[39] *Correlated Noise Provably Beats Independent Noise for Differentially Private Learning* Link   *2024*
International Conference on Learning Representations (ICLR)

   Christopher A. Choquette-Choo, Krishnamurthy Dj Dvijotham, Krishna Pillutla, Arun Ganesh, Thomas Steinke, Abhradeep Guha Thakurta

[38] *Teach LLMs to Phish: Stealing Private Information from Language Models* Link   *2024*
International Conference on Learning Representations (ICLR)

   Ashwinee Panda, **Christopher A. Choquette-Choo**, Zhengming Zhang, Yaoqing Yang, Prateek Mittal

[37] *Poisoning web-scale training datasets is practical* Link   *2024*
(Oral) IEEE Symposium on Security and Privacy (IEEE S&P)

   Nicholas Carlini, Matthew Jagielski, **Christopher A. Choquette-Choo**, Daniel Paleka, Will Pearce, Hyrum Anderson, Andreas Terzis, Kurt Thomas, Florian Tramèr.

[36] *(Amplified) Banded Matrix Factorization: A unified approach to private training* Link   *2023*
Thirty-seventh Conference on Neural Information Processing Systems (Neurips)

   Christopher A. Choquette-Choo, Arun Ganesh, Ryan McKenna, H. Brendan McMahan, Keith Rush, Abhradeep Guha Thakurta, Zheng Xu.

[35] *Are aligned neural networks adversarially aligned?* Link   *2023*
Thirty-seventh Conference on Neural Information Processing Systems (Neurips)

   Nicholas Carlini, Milad Nasr, **Christopher A. Choquette-Choo**, Matthew Jagielski, Irena Gao, Anas Awadalla, Pang Wei Koh, Daphne Ippolito, Katherine Lee, Florian Tramèr, Ludwig Schmidt.

[34] *Students Parrot Their Teachers: Membership Inference on Model Distillation* Link   *2023*
(Oral) Thirty-seventh Conference on Neural Information Processing Systems (Neurips)

   Matthew Jagielski, Milad Nasr, Katherine Lee, **Christopher A. Choquette-Choo**, Nicholas Carlini.

[33] *MADLAD-400: Multilingual And Document-Level Large Audited Dataset* Link   *2023*
Thirty-seventh Conference on Neural Information Processing Systems (Neurips)

   Sneha Kudugunta, Isaac Caswell, Biao Zhang, Xavier Garcia, **Christopher A. Choquette-Choo**, Katherine Lee, Derrick Xin, Aditya Kusupati, Romi Stella, Ankur Bapna, Orhan Firat

[32] *Robust and Actively Secure Serverless Collaborative Learning* Link   *2023*
Thirty-seventh Conference on Neural Information Processing Systems (Neurips)

Nicholas Franzese, Adam Dziedzic, **Christopher A. Choquette-Choo**, Mark R. Thomas, Muhammad Ahmad Kaleem, Stephan Rabanser, Congyu Fang, Somesh Jha, Nicolas Papernot, Xiao Wang

[31] *Multi-epoch matrix factorization mechanisms for private machine learning* Link   *2023*
(Oral) Proceedings of the 40th International Conference on Machine Learning (ICML)

   Christopher A. Choquette-Choo, H. Brendan McMahan, Keith Rush, Abhradeep Thakurta.

[30] *Private Federated Learning with Autotuned Compression* Link   *2023*
Proceedings of the 40th International Conference on Machine Learning (ICML)

   Enayat Ullah*, **Christopher A. Choquette-Choo***, Peter Kairouz*, Sewoong Oh*.
   *Equal contribution

[29] *Federated Learning of Gboard Language Models with Differential Privacy* Link   *2023*
The 61st Annual Meeting of the Association for Computational Linguistics

   Zheng Xu, Yanxiang Zhang, Galen Andrew, **Christopher A. Choquette-Choo**, Peter Kairouz, H. Brendan McMahan, Jesse Rosenstock, Yuanbo Zhang.

[28] *Preventing verbatim memorization in language models gives a false sense of privacy* Link   *2023*
(Runner-up Best Paper) Proceedings of the 16th International Natural Language Generation Conference

   Daphne Ippolito, Florian Tramèr*, Milad Nasr*, Chiyuan Zhang*, Matthew Jagielski*, Katherine Lee*, **Christopher A. Choquette-Choo***, Nicholas Carlini.
   *Equal contribution, random ordering.

[27] *Proof-of-Learning is Currently More Broken Than You Think* Link   *2023*
(Oral) IEEE 8th European Symposium on Security and Privacy (EuroS&P). IEEE Computer Society

   Congyu Fang*, Hengrui Jia*, Anvith Thudi, Mohammad Yaghini, **Christopher A. Choquette-Choo**, Natalie Dullerud, Varun Chandrasekaran, Nicolas Papernot.
   *Equal contribution, alphabetical ordering.

[26] *Private Multi-Winner Voting for Machine Learning* Link   *2023*
Proceedings on 23rd Privacy Enhancing Technologies Symposium (PETS)

   Adam Dziedzic, **Christopher A. Choquette-Choo**, Natalie Dullerud, Vinith Menon Suriyakumar, Ali Shahin Shamsabadi, Muhammad Ahmad Kaleem, Somesh Jha.

[25] *The fundamental price of secure aggregation in differentially private federated learning* Link   *2022*
(Spotlight) International Conference on Machine Learning. PMLR

   Wei-ning Chen*, **Christopher A. Choquette-Choo***, Peter Kairouz*, Ananda Theertha Suresh*.
   *Equal contribution, alphabetical ordering.

[24] *Label-Only Membership Inference Attacks* Link   *2021*
(Spotlight) International Conference on Machine Learning (ICML)

   Christopher A. Choquette-Choo, Florian Tramer, Nicholas Carlini, Nicolas Papernot.

[23] *Entangled Watermarks as a Defense against Model Extraction* Link   *2021*
(Oral) USENIX Security Symposium (USENIX)

   Hengrui Jia, **Christopher A. Choquette-Choo**, Varun Chandrasekaran, Nicolas Papernot.

[22] *Proof of Learning: Definitions and Practice* Link   *2021*
(Oral) IEEE Symposium on Security and Privacy (IEEE S&P)

   Hengrui Jia*, Mohammad Yaghini*, **Christopher A Choquette-Choo**,^Natalie Dullerud,^Anvith Thudi,^ Varun Chandrasekaran, Nicolas Papernot.
   *,Êqual contribution, alphabetical ordering.

[21] *Machine Unlearning* Link   *2021*
(Oral) IEEE Symposium on Security and Privacy (IEEE S&P)

   Lucas Bourtoule*, Varun Chandrasekaran*, **Christopher A. Choquette-Choo***, Hengrui Jia*, Adelin Travers*, Baiwu Zhang*, David Lie, Nicolas Papernot.
   *Equal contribution, alphabetical ordering.

[20] *CaPC Learning: Confidential and Private Collaborative Learning* Link   *2021*
International Conference on Learning Representations (ICLR)

> **Christopher A. Choquette-Choo***, Natalie Dullerud*, Adam Dziedzic*, Yunxiang Zhang*, Somesh Jha, Nicolas Papernot, Xiao Wang.
> *Equal contribution, alphabetical ordering.

[19] *A Multi-label, Dual-Output Deep Neural Network for Automated Bug Triaging* Link   *2019*
International Conference on Machine Learning and Applications (ICMLA)

> **Christopher A. Choquette-Choo**, David Sheldon, Jonny Proppe, John Alphonso-Gibbs, Harsha Gupta.

Peer-Reviewed Workshop Proceedings

[18] *Communication Efficient Federated Learning with Secure Aggregation and Differential Privacy* Link   *2021*
the Neural Information Processing Systems (NeurIPS) workshop on Privacy in Machine Learning

> Wei-ning Chen*, Christopher A. Choquette-Choo*, Peter Kairouz*.
> *Equal contribution, alphabetical ordering.

[17] *User Inference Attacks on Large Language Models* Link   *2023*
Socially Responsible Language Modelling Research (SoLaR)

> Nikhil Kandpal, Krishna Pillutla, Alina Oprea, Peter Kairouz, **Christopher A. Choquette-Choo**, Zheng Xu

[16] *Correlated Noise Provably Beats Independent Noise for Differentially Private Learning* Link   *2023*
International Workshop on Federated Learning in the Age of Foundation Models (FL@FM-NeurIPS'23)

> **Christopher A. Choquette-Choo**, Krishnamurthy Dj Dvijotham, Krishna Pillutla, Arun Ganesh, Thomas Steinke, Abhradeep Guha Thakurta

[15] *User Inference Attacks on Large Language Models* Link   *2023*
International Workshop on Federated Learning in the Age of Foundation Models in Conjunction with NeurIPS (FL@FM-NeurIPS'23)

> Nikhil Kandpal, Krishna Pillutla, Alina Oprea, Peter Kairouz, **Christopher A. Choquette-Choo**, Zheng Xu

Reports

[14] *CodeGemma: Open Code Models Based on Gemma* Link   *2024*
arXiv

> ..., **Christopher A. Choquette-Choo***, ...
> *Contributor.

[13] *Gemma: Open Models Based on Gemini Research and Technology* Link   *2024*
arXiv

> ..., **Christopher A. Choquette-Choo***, ...
> *Contributor. Led memorization efforts.

[12] *Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context* Link   *2024*
arXiv

> ..., **Christopher A. Choquette-Choo***, ...
> *Contributor. Led memorization testing.

[11] *Gemini: A Family of Highly Capable Multimodal Models* Link   *2023*
arXiv

> Anil, R., ..., **Christopher A. Choquette-Choo***, ..., & Vinyals, O.
> *Contributor. Led memorization efforts.

[10] *Palm 2 technical report* Link   *2023*
arXiv

Anil, R., Dai, A. M., Firat, O., Johnson, M., Lepikhin, D., Passos, A., ..., **Christopher A. Choquette-Choo***, ..., & Wu, Y.
*Core contributor. Led memorization efforts.

[9] *Report of the 1st Workshop on Generative AI and Law* Link   *2023*
arXiv

A. Feder Cooper*, Katherine Lee*, James Grimmelmann, Daphne Ippolito, Christopher Callison-Burch, **Christopher A. Choquette-Choo**, ...
*Equal contribution, alphabetical ordering.

Pre-Prints (arXiv)

[8] *Fine-tuning with differential privacy necessitates an additional hyperparameter search* Link   *2022*
arXiv

Yannis Cattan, **Christopher A Choquette-Choo**, Nicolas Papernot, Abhradeep Thakurta

[7] *Scalable Extraction of Training Data from (Production) Language Models* Link   *2023*
arXiv

Milad Nasr, Nicholas Carlini, Jonathan Hayase, Matthew Jagielski, A. Feder Cooper, Daphne Ippolito, **Christopher A. Choquette-Choo**, Eric Wallace, Florian Tramèr, Katherine Lee

[6] *Optimal Rates for DP-SCO with a Single Epoch and Large Batches* Link   *2024*
arXiv

Christopher A. Choquette-Choo, Arun Ganesh, Abhradeep Thakurta

[5] *Phantom: General Trigger Attacks on Retrieval Augmented Language Generation* Link   *2024*
arXiv

Harsh Chaudhari, Giorgio Severi, John Abascal, Matthew Jagielski, **Christopher A. Choquette-Choo**, Milad Nasr, Cristina Nita-Rotaru, Alina Oprea

Under Review (and not yet released)

[4] *Privacy Auditing of Large Language Models* Link   *2024*
under review

Ashwinee Panda, Xinyu Tang, Milad Nasr, **Christopher A. Choquette-Choo**, Prateek Mittal

[3] *The Last Iterate Advantage: Empirical Auditing and Principled Heuristic Analysis of Differentially Private SGD*
Link   *2024*
under review

Milad Nasr, Thomas Steinke, Borja Balle, **Christopher A. Choquette-Choo**, Arun Ganesh, Matthew Jagielski, Jamie Hayes, Abhradeep Thakurta, Adam Smith, Andreas Terzis

[2] *POST: A Framework for Privacy of Soft-prompt Transfer* Link   *2024*
under review

Xun Wang, Jing Xu, **Christopher A. Choquette-Choo**, Adam Dziedzic, Franziska, Boenisch

[1] *Data Source Attribution in Diffusion Models* Link   *2024*
under review

Matthew Jagielski, Milad Nasr, Nicholas Carlini, **Christopher A. Choquette-Choo**, A. Feder Cooper, Katherine Lee, Andreas Terzis, Georgina Evans, Chiyuan Zhang, Avijit Ghosh, Florian Tramèr

Talks

Invited Talks

DP-Follow-The-Regularized-Leader: State-of-the-art Optimizers for Private Machine Learning.   *2024*
Institute of Science and Technology Austria (ISTA) for Prof. Christoph Lampert   Slides available upon request.

DP-Follow-The-Regularized-Leader: State-of-the-art Optimizers for Private Machine Learning. 2024
"Federated Learning on the Edge" AAAI Spring 2024 Symposium. *Slides available upon request.*

Host of "Private Optimization with Correlated Noise" invited session and co-presented first talk 2024
Information Theory and Applications (ITA) *Slides available upon request.*

Poisoning Web-Scale Training Datasets is Practical 2024
Guest talk for Prof. Varun Chandrasekaran at University of Illinois *Slides available upon request.*

The Privacy Considerations of Production Machine Learning 2021
MLOps New York Area Summit *Slides available upon request.*

Adversarial Machine Learning: Ensuring Security and Privacy of ML Models and Sensitive Data 2019
REWORK Responsible AI Summit *Available as a part of the Privacy and Security in Machine Learning package*

Paper Presentations

Multi-Epoch Matrix Factorization Mechanisms for Private Machine Learning *Oral presentation at ICML 2023*

The Fundamental Price of Secure Aggregation in Differentially Private Machine Learning *Spotlight at ICML 2022*

Label-Only Membership Inference Attacks *Spotlight at ICML 2021*

Proof-of-Learning Definitions and Practice *Oral presentation at IEEE S&P 2021*

Machine Unlearning *Oral presentation at IEEE S&P 2021*

Professional Activities

Program Committee

IEEE Security and Privacy (S&P) conference *2025*

IEEE Security and Privacy (S&P) conference *2024*

Generative AI + Law (GenLaw)'23 Workshop at ICML *2023*

Area Chair

Neural Information Processing Systems (NeurIPS) *2024*

Session Chair

DL: Robustness at International Conference on Machine Learning (ICML) *2022*

Reviewer

International Conference on Machine Learning (ICML) *2024*

International Conference on Learning Representations (ICLR) *2024*

Google Research Scholar *2023-2024*

Nature Machine Intelligence Journal *2023*

Neural Information Processing Systems (NeurIPS) *2023*

International Conference on Machine Learning (ICML) *2023*

Neural Information Processing Systems (NeurIPS)	*2022*
Nature Machine Intelligence Journal	*2022*
International Conference on Machine Learning (ICML) + Outstanding	*2022*
IEEE Transactions on Emerging Topics in Computing	*2022*
Machine Learning for the Developing World (ML4D) workshop at NeurIPS	*2021*
Journal of Machine Learning Research	*2021*
Machine Learning for the Developing World (ML4D) workshop at NeurIPS	*2020*

External Reviewer

USENIX Security Symposium	*2022*
IEEE Symposium on Security and Privacy	*2022*
International Conference on Machine Learning (ICML)	*2021*
USENIX Security Symposium	*2021*
IEEE Symposium on Security and Privacy	*2021*

Education

Bachelor of Applied Science in Engineering Science   *University of Toronto*
Major in Robotics Engineering   *2015-2020*

Thesis: Label-Only Membership Inference Attacks as Realistic Privacy Threats
Graduation with Honors (cGPA 3.73/4.00)

Honors and Awards

Schulich Leaders Full Scholarship   *University of Toronto*
$100,000 Value   *2015-2020*
Awarded on the basis of academic achievement and leadership to students pursuing a STEM degree.

Class of 9T7 Award   *University of Toronto*
$4000 Value   *2017*
Awarded on the basis of academic achievement and leadership.

Director's Summer Research Opportunities   *University of Toronto*
$5000 Value   *2016*
Awarded to fund a summer research opportunity in Canada at the Institute for Biomaterials and Biomedical Engineering.

Burger King Scholarship   *University of Toronto*
$1500 Value   *2015*
Awarded on the basis of academic achievement and leadership.

University of Toronto Scholarship   *University of Toronto*
$6000 Value   *2015*
Awarded on the basis of academic achievement.

Competitions

Undergraduate Science Case Competition (SCINAPSE)   *Western University*
(Finalist of 2) of 250+ teams. Upper Year Division.   *2017*

Microsoft Azure Machine Learning Case Competition   *University of Toronto*
(1st) of 20+ teams.   *2017*

UTEK Consulting Competition
(Semi-Finalist) of 20+ teams.

University of Toronto
2016

The Game, Engineering Design Competition
(1st) of 10+ teams. $10,000 value.

University of Toronto
Sept. 2015 - Mar. 2016

Community Outreach

Public Software

Google Research: Main Owner of Multi-Epoch Matrix Factorization package
2023

Google Research: Owner of Private Linear Compression
2022

TensorFlow Privacy: Sole Contributor of Bolt-On Method for Differentially Private Training
2019

CleverHans Blog

Arbitrating the integrity of stochastic gradient descent with proof-of-learning
2021

Beyond federation: collaborating in ML with confidentiality and privacy
2021

Teaching Machines to Unlearn
2020

Personal Blog

How to do Machine Unlearning
2021

Teaching Machines to Unlearn
2020

Community Service and Leadership

University of Toronto Consulting Association, University of Toronto
Director of Volunteer Consulting Group

University of Toronto
2017-2018

FoodSkrap Startup
Co-Founder, CEO, and Software Developer

Own Incorporation
2016-2017

You're Next Career Network
Director of Business Development, Startup

University of Toronto
2016-2017

Board of Directors
Youth Advisor

Plan Canada
2015-2017

Youth Advisory Council
Member

Plan Canada
2014-2017

Technical skills

Proficient in:	Python, C
Familiar with:	Java, MATLAB, Perl, SQL, Elasticsearch, JavaScript
Python libraries:	TensorFlow, Jax, Pax, SeqIO, T5X, PyTorch, NumPy, Pandas, Matplotlib, Scikit-learn, TensorFlow Federated, TensorFlow Privacy

Soft skills

Communication	I focus on communicating complex ideas in a way anyone can understand.
Teamwork	I care about being considerate and sharing responsibility in effective ways. Evidenced by 10 peer bonuses and 1 kudos at Google.
Leadership	I believe that identifying strengths and clearing runways enables success.