

Virginia Smith

Employment

- 2024– **Associate Professor (pre-tenure)**, *Carnegie Mellon University*
- 2018–2024 **Assistant Professor**, *Carnegie Mellon University*
- 2021 **Visiting Researcher**, *Google*, Federated Learning Team
- 2017–2018 **Postdoctoral Researcher**, *Stanford University*, Advisor: Christopher Ré
- 2012–2017 **Research Assistant**, *UC Berkeley*

Education

- 2012–2017 **MS & PhD, Computer Science**, *University of California, Berkeley*
Advisors: Michael I. Jordan and David Culler
Thesis: *System-Aware Optimization for Machine Learning at Scale*
- 2008–2012 **BA, Mathematics & BA, Computer Science, Highest Distinction**, *University of Virginia*

Awards

- 2024 Alfred P. Sloan Research Fellowship
- 2023– Leonardo Career Development Chair
- 2023 Samsung AI Researcher of the Year Award
- 2023 MLSys Outstanding Paper Award, for “Validating Large Language Models with ReLM”
- 2023 UK-US Privacy Enhancing Technologies Prize Challenge, First Place Solution
- 2022 Intel Rising Star Award
- 2022 Apple Privacy-Preserving Machine Learning Award
- 2022 Meta Privacy Enhancing Technologies Research Award
- 2022 NSF CAREER Award
- 2021 MIT Technology Review’s 35 Innovators Under 35 Award
- 2021 Google Research Scholar Award
- 2020 Facebook Faculty Research Award
- 2019 Carnegie Bosch Institute Research Award
- 2018 Google Faculty Research Award
- 2017 Outstanding Graduate Student Instructor Award
- 2016 Rising Stars in EECS, Invited Participant
- 2015 MLconf Industry Impact Student Research Award
- 2014–2017 National Science Foundation Graduate Research Fellowship
- 2014 National Defense Science and Engineering Graduate Fellowship
- 2014 Tong Leong Lim Pre-Doctoral Prize
- 2014 Google Anita Borg Memorial Scholarship
- 2012–2014 UC Berkeley Chancellor’s Fellowship
- 2012 UC Berkeley College of Engineering Fellowship
- 2012 UC Berkeley Department of Electrical Engineering and Computer Sciences Excellence Award

- 2012 CRA Outstanding Undergraduate Research Award, Honorable Mention
- 2012 Rader Undergraduate Research Award (UVA Top Undergraduate CS Research)
- 2009–2012 College Science Scholar (UVA Research Program)
- 2008–2012 Echols Scholar (UVA Honors Program)
- 2008 J. L. Wang Memorial Mathematics Scholarship

Publications

Conference or Journal

- P. Thaker, S. Hu, N. Kale, Y. Maurya, Z. S. Wu, and V. Smith, “Position: Llm unlearning benchmarks are weak measures of progress,” *IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, 2025.
- P. Thaker, A. Setlur, Z. S. Wu, and V. Smith, “Leveraging public representations for private transfer learning,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
- Q. Pang, S. Hu, W. Zheng, and V. Smith, “No free lunch in llm watermarking: Trade-offs in watermarking design choices,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
- A. Setlur, S. Garg, X. Geng, N. Garg, V. Smith, and A. Kumar, “Rl on incorrect synthetic data scales the efficiency of llm math reasoning by eight-fold,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
- A. Muhamed, O. Li, D. Woodruff, M. Diab, and V. Smith, “Grass: Compute efficient low-memory llm training with structured sparse gradients,” in *Empirical Methods in Natural Language Processing (EMNLP)*, 2024.
- A. Setlur, S. Garg, V. Smith, and S. Levine, “Prompting is a double-edged sword: Improving worst-group robustness of foundation models,” in *International Conference on Machine Learning (ICML)*, 2024.
- Y. J. Cho, D. Jhunjhunwala, T. Li, V. Smith, and G. Joshi, “Maximizing global model appeal in federated learning,” *Transactions on Machine Learning Research (TMLR)*, 2024.
- S. Hu, Z. S. Wu, and V. Smith, “Fair federated learning via bounded group loss,” in *IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, 2024.
- D. Dennis, A. Shetty, A. Sevekari, K. Koishida, and V. Smith, “Progressive knowledge distillation: Constructing ensembles for efficient inference,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- O. Li, J. Harrison, J. Sohl-Dickstein, V. Smith, and L. Metz, “Variance-reduced gradient estimation via noise-reuse in online evolution strategies,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- S. Garg, A. Setlur, Z. C. Lipton, S. Balakrishnan, V. Smith, and A. Raghunathan, “Complementary benefits of contrastive learning and self-training under distribution shift,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.
- T. Li*, A. Beirami*, M. Sanjabi, and V. Smith, “On tilted losses in machine learning: Theory and applications,” *Journal of Machine Learning Research*, vol. 24, no. 142, pp. 1–79, 2023.
- S. Hu, Z. S. Wu, and V. Smith, “Private multi-task learning: Formulation and applications to federated learning,” *Transactions on Machine Learning Research (TMLR)*, 2023.
- M. Kuchnik, V. Smith, and G. Amvrosiadis, “Validating large language models with ReLM,” in *Conference on Machine Learning Systems (MLSys)*, 2023.
- K. Kuo, P. Thaker, M. Khodak, J. Ngyuen, D. Jiang, A. Talwalkar, and V. Smith, “On noisy evaluation in federated hyperparameter tuning,” in *Conference on Machine Learning Systems (MLSys)*, 2023.
- T. Li, M. Zaheer, K. Z. Liu, S. J. Reddi, H. B. McMahan, and V. Smith, “Differentially private adaptive optimization with delayed preconditioners,” in *International Conference on Learning Representations (ICLR)*, 2023.

- A. Setlur, D. Dennis, B. Eysenbach, A. Raghunathan, C. Finn, V. Smith, and S. Levine, "Bitrate-constrained dro: Beyond worst case robustness to unknown group shifts," in *International Conference on Learning Representations (ICLR)*, 2023.
- Z. Liu, S. Hu, Z. S. Wu, and V. Smith, "On privacy and personalization in cross-silo federated learning," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.
- A. Setlur, B. Eysenbach, V. Smith, and S. Levine, "Adversarial unlearning: Reducing confidence along adversarial directions," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.
- T. Li, M. Zaheer, S. Reddi, and V. Smith, "Private adaptive optimization with side information," in *International Conference on Machine Learning (ICML)*, 2022.
- O. Li, J. Sun, X. Yang, W. Gao, H. Zhang, J. Xie, V. Smith, and C. Wang, "Label leakage and protection in two-party split learning," in *International Conference on Learning Representations (ICLR)*, 2022.
- R. Balakrishnan, T. Li, T. Zhou, N. Himayat, V. Smith, and J. Bilmes, "Diverse client selection for federated learning via submodular maximization," in *International Conference on Learning Representations (ICLR)*, 2022.
- M. Kuchnik, A. Klimovic, J. Simsa, V. Smith, and G. Amvrosiadis, "Plumber: Diagnosing and removing performance bottlenecks in machine learning data pipelines," in *Conference on Machine Learning and Systems (MLSys)*, 2022.
- Z. Charles, Z. Garrett, Z. Huo, S. Shmulyian, and V. Smith, "On large-cohort training for federated learning," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- A. Setlur*, O. Li*, and V. Smith, "Two sides of meta-learning evaluation: In vs. out of distribution," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- M. Khodak, R. Tu, T. Li, L. Li, M.-F. Balcan, V. Smith, and A. Talwalkar, "Federated hyperparameter tuning: Challenges, baselines, and connections to weight-sharing," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- M. Kuchnik, G. Amvrosiadis, and V. Smith, "Progressive compressed records: Taking a byte out of deep learning data," in *Conference on Very Large Databases (VLDB)*, 2021.
- T. Li, S. Hu, A. Beirami, and V. Smith, "Ditto: Fair and robust federated learning through personalization," in *International Conference on Machine Learning (ICML)*, 2021.
- D. K. Dennis, T. Li, and V. Smith, "Heterogeneity for the win: One-shot federated clustering," in *International Conference on Machine Learning (ICML)*, 2021.
- T. Li, A. Beirami, M. Sanjabi, and V. Smith, "Tilted empirical risk minimization," in *International Conference on Learning Representations (ICLR)*, 2021.
- T. Li, A. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine, Special Issue on Distributed Machine Learning*, 2020.
- T. Li, M. Sanjabi, M. Zaheer, and V. Smith, "Fair resource allocation in federated learning," in *International Conference on Learning Representations (ICLR)*, 2020.
- T. Li, A. Sahu, M. Sanjabi, M. Zaheer, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in *Conference on Machine Learning and Systems (MLSys)*, 2020.
- T. Yu, T. Li, Y. Sun, S. Nanda, V. Smith, V. Sekar, and S. Seshan, "Learning context-aware policies from multiple smart homes via federated multi-task learning," in *ACM/IEEE Conference on Internet of Things Design and Implementation*, 2020.
- T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Feddan: A federated newton-type method," in *Asilomar Conference on Signals, Systems, and Computers*, 2019.
- T. Dao, A. Gu, A. Ratner, V. Smith, C. D. Sa, and C. Re, "A kernel theory of modern data augmentation," in *International Conference on Machine Learning (ICML)*, 2019.

M. Kuchnik and V. Smith, “Efficient augmentation via data subsampling,” in *International Conference on Learning Representations (ICLR)*, 2019.

V. Smith, S. Forte, C. Ma, M. Takac, M. I. Jordan, and M. Jaggi, “CoCoA: A general framework for communication-efficient distributed optimization,” *Journal of Machine Learning Research*, 2018.

V. Smith, C.-K. Chiang, M. Sanjabi, and A. Talwalkar, “Federated multi-task learning,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

C. Ma, J. Konecny, M. Jaggi, V. Smith, M. I. Jordan, P. Richtarik, and M. Takac, “Distributed optimization with arbitrary local solvers,” *Optimization Methods and Software*, 2017.

C. Ma*, V. Smith*, M. Jaggi, M. I. Jordan, P. Richtarik, and M. Takac, “Adding vs. averaging in distributed primal-dual optimization,” in *International Conference on Machine Learning (ICML)*, 2015.

V. Smith, M. Connor, and I. Stanton, “Going in-depth: Finding longform on the web,” in *Conference on Knowledge Discovery and Data Mining (KDD)*, 2015.

M. Jaggi*, V. Smith*, M. Takac, J. Terhorst, S. Krishnan, T. Hofmann, and M. I. Jordan, “Communication-efficient distributed dual coordinate ascent,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2014.

E. Sparks, A. Talwalkar, V. Smith, X. Pan, J. Gonzalez, T. Kraska, M. I. Jordan, and M. J. Franklin, “MLI: An API for user-friendly distributed machine learning,” in *International Conference on Data Mining*, 2013.

J. Taneja, V. Smith, D. Culler, and C. Rosenberg, “A comparative study of high renewables penetration electricity grids,” in *IEEE International Conference on Smart Grid Communications*, 2013.

A. Aswani, N. Master, J. Taneja, V. Smith, A. Krioukov, D. Culler, and C. Tomlin, “Identifying models of HVAC systems using semiparametric regression,” in *American Control Conference (ACC)*, 2012.

V. Smith, T. Sookoor, and K. Whitehouse, “Modeling building thermal response to HVAC zoning,” *ACM SIGBED Review*, vol. 9, no. 3, 2012.

Selected Refereed Workshop

S. Wu, T. Li, Z. Charles, Y. Xiao, Z. Liu, Z. Xu, and V. Smith, “Motley: Benchmarking heterogeneity and personalization in federated learning,” *Workshop on Federated Learning at NeurIPS*, 2022.

S. Caldas, P. Wu, T. Li, J. Konečný, H. B. McMahan, V. Smith, and A. Talwalkar, “Leaf: A benchmark for federated settings,” *Workshop on Federated Learning for Data Privacy and Confidentiality at NeurIPS*, 2019.

N. Guha, A. Talwalkar, and V. Smith, “One-shot federated learning,” in *Machine Learning on Devices Workshop at NeurIPS*, 2018.

Invited Talks

- 2024 NeurIPS Workshop on Statistical Frontiers of LLMs
- 2024 Boston-Area Charles River Privacy Day
- 2024 UC Berkeley CLIMB Seminar
- 2024 Stanford ML Seminar
- 2024 Conference on the Mathematical Theory of Deep Neural Networks (DeepMath)
- 2024 CVPR Workshop on Federated Learning for Vision
- 2024 FLOWER Summit, Keynote
- 2023 NeurIPS Optimization for Machine Learning Workshop
- 2023 NeurIPS Competition Workshop on Document Intelligence and Privacy
- 2023 Princeton S. S. Wilks Memorial Seminar
- 2023 ETH Zurich Scalable Computing Lab Seminar
- 2023 RISE Research Institutes of Sweden Learning Machines Seminar
- 2023 IJCAI Workshop on Trustworthy Federated Learning

2023 SIGIR Workshop on Federated Learning for Information Retrieval
2023 MLSys Workshop on Federated Learning Systems
2023 Intel Rising Star Tech Talk
2023 Amazon Data-Centric AI Seminar Series
2023 Google Federated Learning Talk Series
2022 Ohio State University AI Seminar
2022 NeurIPS Data Compression with Machine Learning, Panelist
2022 NeurIPS Decentralization and Trustworthy ML Workshop
2022 Argonne AI Distinguished Lecture
2022 ACL Federated Learning for NLP Workshop
2022 ICLR Socially Responsible Machine Learning Workshop
2022 Apple Workshop on Privacy Preserving ML
2022 ELLIS Talk Series at IST Austria
2021 MIT OPTML++ Seminar
2021 New Frontiers in Federated Learning Workshop at NeurIPS
2021 Workshop on Distributed Machine Learning, CoNEXT, Keynote
2021 Microsoft Research Summit: Workshop on Federated Learning
2021 Facebook AI Research
2021 Google Research
2021 TWIML AI Podcast
2021 Oracle Machine Learning Seminar Series
2021 FLOWER Summit, Keynote
2021 NSF-TRIPODS Workshop on Communication-Efficient Distributed Optimization
2021 Oracle Machine Learning Seminar
2021 Amazon Alexa AI
2020 Workshop on Scalability, Privacy, and Security in Federated Learning at NeurIPS
2020 Stanford MLSys Seminar
2020 UT Austin ML Seminar
2020 Google Workshop on Federated Learning and Analytics, Keynote
2018 Machine Learning on Consumer Devices Workshop at NeurIPS
2018 Microsoft Research Faculty Summit
2017 ML Systems Workshop at NeurIPS
2017 Google Seattle
2017 Carnegie Mellon University
2017 Massachusetts Institute of Technology
2017 University of Washington
2017 University of California, Los Angeles
2017 Harvey Mudd College
2017 Microsoft Research, New England
2017 Microsoft Research, NYC
2017 Microsoft Research and MSR-NExT, Seattle
2017 SIAM Conference on Optimization (SIOPT)
2016 The Machine Learning Conference (MLconf)
2016 ML Systems Workshop at ICML
2015 The Machine Learning Conference (MLconf)
2015 Modeling and Optimization: Theory and Applications Conference (MOPTA)
2014 Distributed Machine Learning and Matrix Computations Workshop at NeurIPS

Teaching

- Fall 2024 **Instructor**, 10-605/10-805: *Machine Learning with Large Datasets*
Spring 2024 **Instructor**, 10-718: *Machine Learning in Practice*
Fall 2023 **Instructor**, 10-719: *Federated and Collaborative Learning*
Spring 2022 **Instructor**, 10-405/10-605: *Machine Learning with Large Datasets*
Fall 2021 **Instructor**, 10-605/10-805: *Machine Learning with Large Datasets*
Spring 2021 **Co-Instructor**, 10-405/10-605: *Machine Learning with Large Datasets*
Fall 2020 **Co-Instructor**, 10-605/10-805: *Machine Learning with Large Datasets*
Spring 2020 **Co-Instructor**, 10-405/10-605: *Machine Learning with Large Datasets*
Fall 2018 **Co-Instructor**, 18-461/18-661: *Introduction to ML for Engineers*

Service and Activities

- 2024–2025 **Program Chair**, *International Conference on Machine Learning*
2020–2024 **Co-Organizer**, *Federated Learning One World Seminar*
2021–2023 **Workshops Co-Chair**, *International Conference on Machine Learning*
2023 **Co-Organizer**, *Simons Institute Workshop on Federated & Collaborative Learning*
2023 **Co-Organizer**, *MLSys Workshop on Decentralized and Collaborative Learning*
2023 **Co-Organizer**, *MLSys Workshop on Resource-Constrained Learning in Wireless Networks*
2019–2022 **Board Member**, *MLSys: Conference on Machine Learning and Systems*
2020 **Co-Organizer**, *NeurIPS Tutorial on Federated Learning*
2019 **Co-Organizer**, *NeurIPS Workshop on Federated Learning for Data Privacy and Confidentiality*
2019 **Session Chair**, *Asilomar Session on Machine Learning and Optimization in Distributed Networks*
2018–2019 **Program Chair**, *MLSys: Conference on Machine Learning and Systems*
2017–2018 **Co-Organizer**, *MLSys: Conference on Machine Learning and Systems*
2015–2017 **Co-Founder and Organizer**, *Women in Technology Leadership Round Table (wit.berkeley.edu)*
2014–2015 **Co-President**, *UC Berkeley Women in Computer Science and Engineering (WICSE)*
2012–2014 **Officer**, *UC Berkeley Women in Computer Science and Engineering (WICSE)*

Reviewing

- Area Chair Neural Information Processing Systems (NeurIPS), International Conference on Machine Learning (ICML), International Conference on Learning Representations (ICLR), Transactions on Machine Learning Research
Reviewer Journal of Machine Learning Research, Conference on Machine Learning and Systems (MLSys), Foundations and Trends in Machine Learning