Kumar Kshitij Patel

kumarkshitij.patel@yale.edu • kkpatel.ttic.edu • Google Scholar

Research Interests: I develop machine learning methods with rigorous guarantees for heterogeneous and sensitive data, with applications including healthcare. My research spans federated learning, optimization, differential privacy, game theory, and data markets. My recent work also studies private diffusion models and efficient LLM training.

Education

- **Toyota Technological Institute at Chicago** — **Chicago, USA**
  Ph.D. in Computer Science, Advisors: Nathan Srebro, Lingxiao Wang — 2019–2025
  Thesis: What Makes Local Updates Effective: The Role of Data Heterogeneity and Smoothness
- **Indian Institute of Technology Kanpur** — **Kanpur, India**
  B.Tech. in Computer Science and Engineering — 2015–2019
  Year-long academic exchange at École Polytechnique Fédérale de Lausanne (EPFL), 2017–2018

Academic and Research Positions

- **Postdoctoral Associate, Yale University** — **New Haven, USA**
  Faculty host: Daniel Spielman, Institute for Foundations of Data Science — Jul 2025–Jun 2028
  As a Yale FDS postdoc, I develop differentially private methods for learning from distributed and sensitive healthcare data; study the theory of data valuation, efficient optimization of LLMs, and distributed sampling with collaborators across Yale.
- **Research Fellow, Simons Institute for the Theory of Computing** — **Berkeley, USA**
  Research Program on Federated and Collaborative Learning — Jan 2026–May 2026
  Studied federated learning, data valuation, and privacy during the semester-long program; co-organized a workshop, reading groups, and the Data Markets Seminar.
- **Research Intern, EPFL** — **Lausanne, Switzerland**
  Advisors: Aymeric Dieuleveut, Martin Jaggi — 2017–2018
  Studied local and post-local SGD in the machine learning and optimization laboratory, analyzing trade-offs among communication, computation, and generalization.

Awards, Honors, and Distinctions

- **Distinguished Paper Award**, IJCAI [Award] 2024
  - *Online Combinatorial Optimization with Group Fairness Constraints*; one of three papers selected conference-wide from 5,500+ submissions
- **Best Student Paper Honorable Mention** [Award] 2023
  - *On the Still Unreasonable Effectiveness of Federated Averaging for Heterogeneous Distributed Learning*; FL@ICML Workshop
- **Top Reviewer Awards**
  - AISTATS (2023); ICML (2022); NeurIPS (2022, 2025); ICLR (2022, 2026)
- **TTIC Student Workshop Awards**
  - Best Poster Presentation Award (2023); Best Talk Award (2022)
- **Honda Young Engineer and Scientist's (YES) Fellowship** [Award] 2017
  - Awarded to 14 students across India for academic and research work
- **Selected Delegate**, Government of India's Youth Delegation to Nepal 2016

- **Academic Excellence Award**, IIT Kanpur 2015
- **National Examination Distinctions** 2015
  - JEE Advanced: All India Rank 200
  - JEE Main: 99.9th percentile among 1.3 million candidates
  - Indian National Chemistry Olympiad: qualified after placing in the top 1%

Selected Talks and Tutorials

- **Personalized Federated Training of Diffusion Models with Privacy Guarantees**
  - SIAM Conference on Mathematics of Data Science (MDS26) [Program] Nov 2026
  - Information Security Group, University College London [Seminar] Oct 2026
  - MRC Biostatistics Unit Seminar, University of Cambridge [Video] Jul 2026
  - Department of EEE, Imperial College London Jul 2026
  - SAYED Systems Group, Queen Mary University of London Jul 2026
  - Trust in Decentralized Systems Workshop, Simons Institute [Video] Mar 2026
- **Local SGD: Convergence, Personalization, and Second-Order Heterogeneity**
  - ML Theory Seminar, University of Cambridge Oct 2026
  - Statistics Group, London School of Economics Oct 2026
  - Federated Learning Seminar, Vector Institute Mar 2026
  - Google Federated Learning Seminar Jul 2024
  - CISPA Helmholtz Center for Information Security Jul 2024
  - Federated Learning One World (FLOW) Seminar [Video] Jun 2024
- **When Do Score-Based Data Valuation Methods Work, and Why?**
  - Agency in Collaborative Learning Workshop, Simons Institute [Video] Apr 2026
  - Machine Learning Seminar, University of Illinois Urbana-Champaign Mar 2026
- **Reinventing Machine Learning for Multiple Distributions: Optimization, Privacy, and Incentives**
  - Department of Computer Science, University of Texas at Dallas Feb 2025
  - Center for Computational Mathematics, Flatiron Institute Jan 2025
  - Yale Institute for Foundations of Data Science (FDS) Seminar Dec 2024
  - Yu Group, Departments of Statistics and EECS, UC Berkeley Sep 2024
- **Online Combinatorial Optimization with Group Fairness Constraints**
  - INFORMS Annual Meeting; Invited session [Program] Oct 2024
  - International Joint Conference on Artificial Intelligence (IJCAI); paper received a **Distinguished Paper Award** Aug 2024
  - INFORMS Annual Meeting; Invited session [Program] Oct 2023
- **On the Effect of Defections in Federated Learning and How to Prevent Them**
  - Workshop on Federated Learning for Foundation Models at IJCAI [Program] Aug 2024
- **Online Optimization Meets Federated Learning** *(tutorial with Aadirupa Saha)*
  - Conference on Uncertainty in Artificial Intelligence (UAI) [Program] Jul 2023
- **Towards Optimal Communication Complexity in Distributed Non-Convex Optimization**
  - SIAM Conference on Optimization (OP23) [Program] Jun 2023
  - Federated Learning One World (FLOW) Seminar [Video] Sep 2022
  - CISPA Helmholtz Center for Information Security Jun 2022

Industry Research Experience

- **Research Intern, Sony AI** — **Tokyo, Japan**
  Advisors: Nidham Gazagnadou, Lingjuan Lyu — Jun 2023–Sep 2023
  In the privacy-preserving ML team, I developed theory and experiments for personalization in federated learning, including new convergence guarantees for personalized local SGD.
- **Applied Scientist Intern, Amazon Web Services** — **Seattle, USA**
  Advisors: Omer Tripp, Srinivasan Sengamedu — Jun 2020–Sep 2020
  In the CodeGuru Team, I developed deep-learning methods, combined with program analysis, to detect sensitive information leakage in Java code, supporting a customer-facing feature.

Publications

[*] Equal contribution throughout.

Refereed Conference and Journal Publications

- **Patel K. K.**[*], Islamov R.[*], Stich S. U., Lucchi A., Gorbunov E. and Wang L., What's in a Smoothness Constant? Tighter Rates for Local SGD with Bounded Second-Order Heterogeneity. **NeurIPS 2026**. [PDF]
- **Patel K. K.**, Karimireddy S. P., Fernandez R. C. and Zampetakis M., Limits and Potential of Score-Based Data Valuation: Redundancy, Complementarity, and Non-Monotonicity. **NeurIPS 2026**.
- **Patel K. K.**[*], Jiang B.[*], Kabir A. F. M.[*], Zhang W., Zou D. and Wang L., Personalized Federated Training of Diffusion Models with Privacy Guarantees. **CVPR 2026**. [PDF]
- Kundu R., Salvatore M., **Patel K. K.**, Ohno-Machado L., Cho H., Shi X. and Mukherjee B., Privacy-Enhancing Sequential Learning under Heterogeneous Selection Bias in Multi-Site Electronic Health Records Data. **JAMIA, 2026**. [Article]
- Manoj N. S. and **Patel K. K.**, Distributionally Robust Linear Regression with Block Lewis Weights. **ICLR 2026**. [PDF]
- **Patel K. K.**, Zindari A., Wang L. and Stich S. U., Revisiting Consensus Error: A Fine-Grained Analysis of Local SGD under Second-Order Data Heterogeneity. **NeurIPS 2025**. [PDF]
- Golrezaei N.[*], Niazadeh R.[*], **Patel K. K.**[*] and Susan F.[*], Online Combinatorial Optimization with Group Fairness Constraints. **IJCAI 2024; Distinguished Paper Award**. [PDF]
- **Patel K. K.**, Glasgow M., Zindari A., Wang L., Stich S. U., Cheng Z., Joshi N. and Srebro N., The Limits and Potentials of Local SGD for Distributed Heterogeneous Learning with Intermittent Communication. **COLT 2024**. [PDF]
- **Patel K. K.**, Wang L., Saha A. and Srebro N., Federated Online and Bandit Convex Optimization. **ICML 2023**. [PDF]
- **Patel K. K.**, Wang L., Woodworth B., Bullins B. and Srebro N., Towards Optimal Communication Complexity in Distributed Non-Convex Optimization. **NeurIPS 2022**. [PDF]
- Bullins B.[*], **Patel K. K.**[*], Shamir O.[*], Srebro N.[*] and Woodworth B.[*], A Stochastic Newton Algorithm for Distributed Convex Optimization. **NeurIPS 2021**. [PDF]
- Woodworth B., **Patel K. K.** and Srebro N., Minibatch vs Local SGD for Heterogeneous Distributed Learning. **NeurIPS 2020**. [Talk] [PDF]
- Woodworth B., **Patel K. K.**, Stich S. U., Dai Z., Bullins B., McMahan B., Shamir O. and Srebro N., Is Local SGD Better than Minibatch SGD? **ICML 2020**. [Talk] [PDF]
- Lin T., Stich S. U., **Patel K. K.** and Jaggi M., Don't Use Large Mini-batches, Use Local SGD. **ICLR 2020**. [Talk] [PDF]
- **Patel K. K.**[*] and Dieuleveut A.[*], Communication Trade-offs for Synchronized Distributed SGD with Large Step-size. **NeurIPS 2019**. [PDF]

- Kapoor S., **Patel K. K.** and Kar P., Corruption-Tolerant Bandit Learning. **Springer, Machine Learning, 2018**. [PDF]

Manuscripts Under Review and In Submission

- **Patel K. K.**, El Gourari A. and Koloskova A., Differentially Private SGD without the Bounded Gradients Assumption: The Boon of Interpolation. In submission.
- Kabir A. F. M., **Patel K. K.**, Zou D. and Wang L., Privacy-Preserving Personalized Federated Latent Diffusion Models. In submission.
- **Patel K. K.** and Wang L., Sharper Guarantees for Distributed Bandit Convex Optimization. In submission.
- Han M.[*], **Patel K. K.**[*], Shao H.[*] and Wang L.[*], On the Effect of Defections in Federated Learning and How to Prevent Them. In submission. [PDF]

Preprints

- Yunis D., **Patel K. K.**, Wheeler S., Savarese P., Vardi G., Livescu K., Maire M. and Walter M., Approaching Deep Learning through the Spectral Dynamics of Weights. Preprint, 2024. [PDF]
- Wang L., Zou D., **Patel K. K.**, Wu J. and Srebro N., Private Overparameterized Linear Regression Without Suffering in High Dimensions. Preprint, 2024. [PDF]
- **Patel K. K.**, Gazagnadou N., Wang L. and Lyu L., Personalization Mitigates the Perils of Local SGD for Heterogeneous Distributed Learning. Preprint, 2023. [PDF]

Refereed Workshop Papers and Extended Abstracts

- **Patel K. K.**[*] and Manoj N. S.[*], Attention Geometry Helps Early: Mirror-Descent Updates for Transformer Query and Key Matrices. OPT Workshop at NeurIPS 2026.
- **Patel K. K.**, Karimireddy S. P., Fernandez R. C. and Zampetakis M., When Do Score-Based Data Valuation Methods Work, and Why? DATA-FM Workshop at ICLR 2026. [PDF]
- **Patel K. K.** and Murhekar A., Performative Personalization Incentivizes Truthfulness in Federated Learning. AIMS Workshop at ICLR 2026. [PDF]
- Manoj N. S. and **Patel K. K.**, Distributionally Robust Linear Regression with Block Lewis Weights. OPT Workshop at NeurIPS 2024. [PDF]
- Han M.[*], **Patel K. K.**[*], Shao H.[*] and Wang L.[*], On the Effect of Defections in Federated Learning and How to Prevent Them. FL@FM Workshop at IJCAI 2024. [PDF]
- Kabra A.[*] and **Patel K. K.**[*], The Limitations of Model Retraining in the Face of Performativity. Humans, Algorithmic Decision-Making and Society Workshop at ICML 2024. [PDF]
- Wang L., Zhou X., **Patel K. K.**, Tang L. and Saha A., Efficient Private Federated Non-Convex Optimization with Shuffled Model. Privacy Regulation and Protection in ML Workshop at ICLR 2024. [PDF]
- Yunis D., **Patel K. K.**, Wheeler S., Savarese P., Vardi G., Livescu K., Maire M. and Walter M., Grokking, Rank Minimization and Generalization in Deep Learning. Mechanistic Interpretability Workshop at ICML 2024. [PDF]
- Yunis D., **Patel K. K.**, Wheeler S., Savarese P., Vardi G., Frankle J., Livescu K., Maire M. and Walter M., Rank Minimization, Alignment and Weight Decay in Neural Networks. High-Dimensional Learning Dynamics Workshop at ICML 2024. [PDF]
- **Patel K. K.**, Wang L., Saha A. and Srebro N., Federated Online and Bandit Convex Optimization. FL@FM Workshop at NeurIPS 2023. [PDF]
- **Patel K. K.**[*], Glasgow M.[*], Wang L., Joshi N. and Srebro N., On the Still Unreasonable Effectiveness of Federated Averaging for Heterogeneous Distributed Learning. Federated Learning Workshop at ICML 2023; **Best Student Paper Honorable Mention**. [PDF]
- **Patel K. K.**, Saha A., Srebro N. and Wang L., Distributed Online and Bandit Convex Optimization. OPT Workshop at NeurIPS 2022. [PDF]

- Yunis D., **Patel K. K.**, Savarese P., Vardi G., Livescu K. and Walter M., On Convexity and Linear Mode Connectivity. OPT Workshop at NeurIPS 2022. [PDF]

Teaching

- **Guest Lecturer**, Big Data Summer Immersion at Yale (BDSY) [Slides] Jul 2026
  Delivered an introductory lecture for undergraduates on differential privacy and private diffusion models.
- **Instructor**, graduate mini-course on Algorithmic Game Theory Fall 2024
  Designed and delivered the graduate mini-course at Toyota Technological Institute in Nagoya, Japan (TTIJ).
- **Teaching Assistant**, Convex Optimization, TTIC Winter 2022, Winter 2024
  Co-developed lecture slides, reviewed and finalized each lecture with the instructor, and delivered two lectures for the first offering; led tutorials and held office hours for the second offering; helped design assignments and exams, coordinated graders and resolved regrade requests, and graded final exams across both offerings.
- **Teaching Assistant**, Mathematical Toolkit, TTIC Spring 2023
  Co-developed and graded all assignments, led weekly tutorials, and held office hours.
- **Teaching Assistant**, Research at TTIC Colloquium Fall 2020, Winter 2021
  Supervised reports from first-year PhD students and introduced several colloquium speakers.

Professional and Institutional Service

Workshop Organization

- **Learning from Heterogeneous Sources**, Simons Institute [Link] Feb 2026
  Co-organized the workshop alongside the organizers of the semester-long Federated and Collaborative Learning program; invited speakers, coordinated talks around the theme, chaired sessions, and designed and led the open-problems session.
- **Incentives for Collaborative Learning and Data Sharing** [Link] Aug 2025
  New Frontiers in Federated Learning [Link] Sep 2023
  Co-developed successful proposals for TTIC's Summer Workshop program and co-led both workshops end to end: recruited speakers and participants, developed the programs and schedules, organized contributed talks and posters, allocated over $10,000 in participant support, coordinated travel and accommodation, and chaired sessions and discussions.

Peer Reviewing and Area Chairing

- **Area Chair:** ICLR'27.
- **Journals:** JMLR; TMLR; Machine Learning; IEEE TPAMI; JAIR.
- **Conferences:** STOC; FOCS; COLT; ALT; ICML; NeurIPS; ICLR; CVPR; AISTATS; AAAI; IJCAI; ECCV.

Seminar and Reading Group Organization

- Heterogeneity Reading Group and Data Markets Seminar, Simons Institute, 2026.
- Machine Learning and Optimization Reading Group, TTIC, 2021–2022.
- TTIC/UChicago Student Theory Seminar, 2020.

Other Professional and Institutional Service

- AIhub Ambassador and volunteer at IJCAI 2024; interviewed keynote speakers at the conference.
- Session chair, ICML 2022 and NeurIPS 2022; volunteer, NeurIPS 2019.
- Student member, Sexual Misconduct Policy Committee, TTIC, 2021.