# IsyE 8813: Algorithmic Foundations of Ethical ML
# **Differential Privacy Reading**

Summarize 1 paper on privacy if working alone, x if working in a group of x people. For each paper, please i) provide context and motivation for the studied problem, ii) summarize the main results of the paper, iii) provide an overview of the key technical tools use to achieve these results, and iv) find a few questions left open by the paper.

If you are summarizing 2 papers, please try to pick papers that are related to each other, and to discuss how these papers interact with each other (is one a follow-up to the other? Do they complete each other? Do they study the same problem under different assumptions? Do they use the same techniques but apply them to different problems?). Feel free to pick topics or papers of interest that are not included in the list below.

Your paper summary should be emailed to jziani3@gatech.edu. You should also prepare a 15 to 20-minute class presentation about the paper you have read if working alone, and a 30 to 40 minute presentation if working in a group of 2.

## List of Topics

**Beyond Worst-Case Sensitivity, and Personalized Privacy Guarantees:** How to calibrate sensitivity to the average-case or specific current instance rather than to the worst-case over databases?

How to give different "individual" sensitivities to different agents so that we can personalize the level of privacy given to each, instead of calibrating the noise to and giving everyone the worst-case desired privacy level across agents?

- Smooth Sensitivity and Sampling in Private Data Analysis. Kobbi Nissim, Sofya Raskhodnikova, Adam Smith.
- Differential Privacy and Robust Statistics. Cynthia Dwork, Jing Lei.
- Individual Sensitivity Preprocessing for Data Privacy. Rachel Cummings, David Durfee.

**How Can Privacy Help Mechanism Design/Equilibrium Computation:** Differential privacy can be used as a tool to incentivize certain behaviors among strategic agents:

- Approximately Optimal Mechanism Design via Differential Privacy. Kobbi Nissim, Rann Smorodinsky, Moshe Tennenholtz.
- Is privacy compatible with truthfulness? David Xiao.
- Mechanism Design in Large Games: Incentives and Privacy. Michael Kearns, Mallesh Pai, Aaron Roth, Jon Ullman.

- Asymptotically Truthful Equilibrium Selection in Large Congestion Games. Ryan Rogers, Aaron Roth.
- Privacy and Truthful Equilibrium Selection for Aggregative Games. Rachel Cummings, Michael Kearns, Aaron Roth, Zhiwei Steven Wu.

**Mechanism Design for Data Collection with Privacy-Aware Agents:** How do we deal with the optimization and mechanism design issues that come with collection data from privacy-aware agents, i.e.: how to decide what data to get from what agent, how to incenvitize agents to behave truthfully, how to compensate agents for any incurred privacy costs, and how to privately use/aggregate this data?

- Truthful Mechanisms for Agents that Value Privacy. Yiling Chen, Stephen Chong, Ian A. Kash, Tal Moran, Salil Vadhan.
- Privacy-Aware Mechanism Design. Kobbi Nissim, Claudio Orlandi, Rann Smorodinsky.
- Take It or Leave It: Running a Survey when Privacy Comes at a Cost. Katrina Ligett, Aaron Roth.
- Selling Privacy at Auction. Arpita Ghosh, Aaron Roth.
- Redrawing the boundaries on purchasing data from privacy-sensitive individuals. Kobbi Nissim, Salil Vadhan, David Xiao.
- Buying Private Data without Verification. Arpita Ghosh, Katrina Ligett, Aaron Roth, Grant Schoenebeck.
- Approximately optimal auctions for selling privacy when costs are correlated with data. Lisa Fleischer, Yu-Han Lyu.
- Accuracy for Sale: Aggregating Data with a Variance Constraint. Rachel Cummings, Katrina Ligett, Aaron Roth, Zhiwei Steven Wu, Juba Ziani.
- Privacy and coordination: computing on databases with endogenous participation. Arpita Ghosh, Katrina Ligett.

**Privacy and Generalization in Machine Learning:** How one can use differential privacy to provide stability and strong generalization guarantees in adaptive machine learning.

- Generalization in Adaptive Data Analysis and Holdout Reuse. Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, Aaron Roth.
- Preserving Statistical Validity in Adaptive Data Analysis. Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, Aaron Roth.

- Algorithmic Stability for Adaptive Data Analysis. Raef Bassily, Kobbi Nissim, Adam Smith, Thomas Steinke, Uri Stemmer, Jonathan Ullman.

- A New Analysis of Differential Privacy's Generalization Guarantees. Christopher Jung, Katrina Ligett, Seth Neel, Aaron Roth, Saeed Sharifi-Malvajerdi, Moshe Shenfeld.

- Adaptive Learning with Robust Generalization Guarantees. Rachel Cummings, Katrina Ligett, Kobbi Nissim, Aaron Roth, Zhiwei Steven Wu.

**DP Learning ⇔ Online Learning:** How does offline DP learning maps/connects to non-private online learning:

- Online learning via differential privacy. Jacob D. Abernethy, Chansoo Lee, Audra McMillan, and Ambuj Tewari.

- The price of differential privacy for online learning. Naman Agarwal and Karan Singh.

- Private PAC learning implies finite Littlestone dimension. Noga Alon, Roi Livni, Maryanthe Malliaris, and Shay Moran.

- An Equivalence Between Private Classification and Online Prediction. Mark Bun, Roi Livni, Shay Moran.

**In-between Local and Central DP:** Introducing models that interpolate between the better accuracy guarantees of central privacy, and the more decentralized approach of local privacy (that does not require trusting a central curator).

- Distributed Differential Privacy via Shuffling. Albert Cheu, Adam Smith, Jonathan Ullman, David Zeber, Maxim Zhilyaev.

- Amplification by Shuffling: From Local to Central Differential Privacy via Anonymity. Úlfar Erlingsson, Vitaly Feldman, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, Abhradeep Thakurta.

- Pan-Private Streaming Algorithms. Cynthia Dwork, Moni Naor, Toniann Pitassi, Guy N. Rothblum, Sergey Yekhanin.

- Connecting Robust Shuffle Privacy and Pan-Privacy. Victor Balcer, Albert Cheu, Matthew Joseph, Jieming Mao.

- The Limits of Pan Privacy and Shuffle Privacy for Learning and Estimation. Albert Cheu, Jonathan Ullman.

**A Relaxation of Differential Privacy: Concentrated Differential Privacy** A definition of differential privacy that lies somewhere between $(\varepsilon, 0)$ and $(\varepsilon, \delta)$-DP. It guarantees, unlike $(\epsilon, \delta)$-DP, that in the small $\delta$ probability cases where $\epsilon$-DP does not hold, we still have privacy guarantees that degrade gracefully (versus in that case, $(\epsilon, \delta)$-DP guarantees nothing: the probability $\delta$ event could be extremely bad privacy-wise):

- Concentrated Differential Privacy. Cynthia Dwork, Guy N. Rothblum.
- Concentrated Differential Privacy: Simplifications, Extensions, and Lower Bounds. Mark Bun, Thomas Steinke.
- Renyi Differential Privacy. Ilya Mironov.

**Lower Bounds and Reconstruction Attacks:**

- Revealing Information while Preserving Privacy. Irit Dinur, Kobbi Nissim.
- The Power of Linear Reconstruction Attacks. Shiva Prasad Kasiviswanathan, Mark Rudelson, Adam Smith.
- Robust De-anonymization of Large Sparse Datasets. Arvind Narayanan, Vitaly Shmatikov.
- Linear Program Reconstruction in Practice. Aloni Cohen and Kobbi Nissim.
- Lower bounds in differential privacy. Anindya De.
- On the Complexity of Differentially Private Data Release: Efficient Algorithms and Hardness Results. Cynthia Dwork, Moni Naor, Omer Reingold, Guy Rothblum, Salil Vadhan.
- Robust Traceability from Trace Amounts. Cynthia Dwork, Adam Smith, Thomas Steinke, Jonathan Ullman, Salil Vadhan.

**Privacy in Graphs:**

- The Johnson-Lindenstrauss Transform Itself Preserves Differential Privacy. Jeremiah Blocki, Avrim Blum, Anupam Datta, Or Sheffet.
- Analyzing Graphs with Node Differential Privacy. Shiva Prasad Kasiviswanathan, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith.
- Efficient Lipschitz Extensions for High-Dimensional Graph Statistics and Node Private Degree Distributions. Sofya Raskhodnikova, Adam Smith.
- Efficiently Estimating Erdos-Renyi Graphs with Node Differential Privacy. Adam Sealfon, Jonathan Ullman.

**Privacy and the Law:**

- Towards Formalizing the GDPR Notion of Singling Out. Aloni Cohen, Kobbi Nissim.
- What a Hybrid Legal-Technical Analysis Teaches Us About Privacy Regulation: The Case of Singling Out. Micah Altman, Aloni Cohen, Kobbi Nissim, Alexandra Wood.
- Data Protection's Composition Problem. Aaron Fluitt, Aloni Cohen, Micah Altman,Kobbi Nissim, Salome Viljoen, Alexandra Wood.

**Other optim/ML applications:** This includes papers that develop differential privacy techniques for more complex optim and ML classes than the ones seen in class:

- Deep Learning with Differential Privacy. Martín Abadi, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, Li Zhang.
- Differentially Private Online Learning. Prateek Jain, Pravesh Kothari, Abhradeep Thakurta.
- Private Matchings and Allocations. Justin Hsu, Zhiyi Huang, Aaron Roth, Tim Roughgarden, Zhiwei Steven Wu.
- Jointly Private Convex Programming. Justin Hsu, Zhiyi Huang, Aaron Roth, Zhiwei Steven Wu.
- Differentially Private Combinatorial Optimization. Anupam Gupta, Katrina Ligett, Frank McSherry, Aaron Roth, Kunal Talwar.
- How to Use Heuristics for Differential Privacy. Seth Neel, Aaron Roth, Zhiwei Steven Wu.

**Differential Privacy in Practice:** Covers some of the recent applications of differential privacy, both by the public (US Census) and private (Google, Apple, Microsoft, etc.) sectors:

- Census TopDown: Differentially Private Data, Incremental Schemas, and Consistency with Public Knowledge. John Abowd, Robert Ashmead, Simson Garfinkel, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, William Sexton, Brett Moran.
- Census TopDown: The Impacts of Differential Privacy on Redistricting. Aloni Cohen, Moon Duchin, JN Matthews, Bushan Suwal.
- Prochlo: Strong privacy for analytics in the crowd. Andrea Bittau, Ulfar Erlingsson, Petros Maniatis, Ilya Mironov, Ananth Raghunathan, David Lie, Mitch Rudominer, Ushasree Kode, Julien Tinnes, and Bernhard Seefeld.
- Collecting telemetry data privately. Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin.

- Rappor: Randomized aggregatable privacy-preserving ordinal response. Ulfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova.

- Learning with privacy at scale. Apple Differential Privacy Team.