CS208 Spring 2022 Annotated Bibliography

James Honaker, Salil Vadhan, and Wanrong Zhang

January 25, 2022

- Background Material
  - Discrete math and proofs: Solow [2013], Rosen [2012]
  - Algorithms and complexity: Cormen et al. [2009], Mitzenmacher and Upfal [2005]
  - Basic Probability and statistics: Ross [1998]
- General References
  - Many videos of talks on recent developments in the theory and applications of differential privacy: `https://simons.berkeley.edu/programs/privacy2019`
  - Tutorial on "DP in the Wild": Machanavajjhala et al. [2017] (see also slides online)
  - A list of real-world uses of differential privacy: Desfontaines [2021]
  - Lecture Notes on Privacy in Machine Learning and Statistics: Smith and Ullman [2022]
- Reidentification Attacks
  - (assigned) Forbes article on Sweeney's reidentification of Personal Genome Project participants: Tanner [2013]
  - (assigned) New York Times article on reidentification from AOL Search Log release: Barbaro and Zeller [2006]
  - (assigned) Narayanan-Shmatikov opinion piece on the concept of PII: Narayanan and Shmatikov [2010]
  - Sweeney's original re-identification: Sweeney [1997]
  - Statistics on reidentification by DOB, ZIP, and Sex: Sweeney [2000], Golle [2006]
  - Paper on the Personal Genome Project reidentification: Sweeney et al. [2013]
  - Paper introducing k-anonymity: Sweeney [2002]
  - Composition attack on k-anonymity: Ganta et al. [2008]
  - Biases introduced by deidentification of EdX data: Daries et al. [2014]
  - Netflix reidentification: Narayanan and Shmatikov [2008]
  - Cancellation of 2nd Netflix Challenge after Lawsuit: Singel [2010]
  - Cohen's downcoding attacks and EdX reidentification: Cohen [2021]
  - Defenses of de-identification: Cavoukian and Castro [2014], Cavoukian and El Emam [2014]
- Reconstruction Attacks
  - Linear programming attack on Diffix: Cohen and Nissim [2018]
  - SAT Solver attack on Census data: Garfinkel et al. [2018a]

- Survey paper on attacks on aggregate statistics: Dwork et al. [2017, §1,2]
- Paper introducing reconstruction attacks: Dinur and Nissim [2003]
- Differencing attack on Israeli Census: Ziv [2013]

- Membership Attacks
  - P3G Consortium responses to membership attacks on genomic data: Consortium et al. [2009]
  - Privacy attacks on microtargeted ads: Korolova [2011, §1,4,6,8]
  - Survey paper on attacks on aggregate statistics: Dwork et al. [2017, §3]
  - Membership attack on means in genomic data: Homer et al. [2008]
  - Membership attack on noisy means: Dwork et al. [2015b]
  - Membership attack on ML as a Service: Shokri et al. [2017]
  - Attribute inference attacks on ML: Fredrikson et al. [2014]
  - Blog post in response to inference attacks on ML: McSherry [2016]

- Foundations of Differential Privacy
  - Primer for non-technical audiences: Wood et al. [2018]
  - A book about differential privacy, for programmers: Near and Abuah [2021]
  - The standard textbook: Dwork and Roth [2013]
  - Survey on complexity-theoretic aspects of differential privacy: Vadhan [2017]
  - The papers leading up to and culminating in the definition of differential privacy and the first mechanisms (Laplace, histograms, implementing the SQ model): Dinur and Nissim [2003], Dwork and Nissim [2004], Blum et al. [2005], Dwork et al. [2016].
  - Attacks on floating-point implementations of differential privacy and remedies: Mironov [2012], Balcer and Vadhan [2018]
  - The geometric mechanism: Ghosh et al. [2012]
  - A Bayesian interpretation of approximate DP: Kasiviswanathan and Smith [2014]
  - A survey on differential privacy for social networks: Raskhodnikova and Smith [2014]
  - The advanced and "optimal" composition theorems for approximate DP: Dwork et al. [2010], Kairouz et al. [2017], Murtagh and Vadhan [2018]
  - Other variants of DP that compose more cleanly than approximate DP: Dwork and Roth [2013], Bun and Steinke [2016], Mironov [2017], Bun et al. [2018]
  - Differential privacy and the Statistical Query model for machine learning: Blum et al. [2005], Kasiviswanathan et al. [2011]
  - The paper that introduced the exponential mechanism: McSherry and Talwar [2007]
  - Another mechanism for the median (via smooth sensitivity): Kasiviswanathan et al. [2013]
  - Survey of approaches to add noise closer to the local sensitivity: [Vadhan, 2017, Ch. 3]

- Implementing Differential Privacy: One-Shot Releases
  - The stability-based histogram and other histogram algorithms for large data universes: Korolova et al. [2009], Balcer and Vadhan [2018]
  - Early applications of DP synthetic data to commuting patterns and mobility data: Machanavajjhala et al. [2008], Mir et al. [2013]

- (required or slides covered in class) Census Bureau's adoption of DP: Garfinkel et al. [2018b], Garfinkel [2018]
- Other papers and talks on the Census Bureau's adoption of DP: Abowd [2018], Kifer [2019], Dajani et al. [2017]
- Private Multiplicative Weights: Hardt and Rothblum [2010]. (See also sections of Dwork and Roth [2013], Vadhan [2017].)
- (excerpts required) DualQuery: Gaboardi et al. [2017]
- Another algorithm for synthetic data generation (MWEM): Hardt et al. [2012]
- Worst-case hardness of differentially private synthetic data generation: Dwork et al. [2009], Ullman and Vadhan [2011] (See also sections of Vadhan [2017].)
- (excerpts required) The Opportunity Atlas and the underlying privacy mechanism: Chetty et al. [2018], Chetty and Friedman [2019]
- The Matrix Mechanism: Li et al. [2015]
- The Hierarchical Mechanism for Range Queries: Hay et al. [2010]
- How to compare DP algorithms: Hay et al. [2016]

• Implementing Differential Privacy: Programming Frameworks and Query Systems

- PinQ and its formal verification: McSherry [2010], Ebadi and Sands [2017]
- εktelo: Zhang et al. [2018]
- Differentially Private SQL: Johnson et al. [2018], Kotsogiannis et al. [2019]
- Differentially Private MapReduce: Roy et al. [2010]
- Side-channel attacks on implementations of DP: Haeberlen et al. [2011], Mironov [2012]
- Survey on formal verification of DP and recent developments: Barthe et al. [2016], Zhang and Kifer [2017], Albarghouthi and Hsu [2017]
- DP Query Systems that Budget via Accuracy: Mohan et al. [2012], Gaboardi et al. [2016]

• The Local and Multiparty Models of Differential Privacy, and Combining Cryptography and DP

- Tutorial: Cormode et al. [2018], see also videos online
- Survey talk by Adam Smith: http://www.bu.edu/hic/files/2018/06/2018-06-05-Adam.Smith_.pptx (Change file extension to .pdf to open.)
- History of randomized response in the survey literature, and some current applications: Gingerich [2015, 2010], Blair et al. [2015]
- Equivalence of local DP and the SQ model: Kasiviswanathan et al. [2011]
- More on models for interactive and multiparty DP: Vadhan [2017, Chs. 9-10]
- Composition when privacy parameters are chosen adaptively: Rogers et al. [2016]
- Local DP with anonymous/shuffled data subjects: Bittau et al. [2017], Cheu et al. [2019], Erlingsson et al. [2019], Balle et al. [2019]
- Differential Privacy meets Multiparty Computation workshop: http://www.bu.edu/hic/dpmpc-2018/
- Recent papers on combining DP and secure multiparty computation: He et al. [2017], Archer et al. [2018]
- Google's RAPPOR: Erlingsson et al. [2014]
- Apple's "learning with privacy at scale": https://machinelearning.apple.com/2017/12/06/learning-with-privacy-at-scale.html

  - Microsoft's "Collecting telemetry data privately": https://www.microsoft.com/en-us/research/blog/collecting-telemetry-data-privately/, Ding et al. [2017]
  - Critiques of deployments of local DP: https://www.wired.com/story/apple-differential-privacy-shortcom Tang et al. [2017]
  - Local DP for Evolving Data: Joseph et al. [2018]

- Machine Learning and Statistical Inference with DP
  - Bibliography for Adam Smith's Fall 2018 course CS 591 at BU: https://docs.google.com/document/d/1jsZLEd3ZM-ZWdNAjNRI4_bgPysRUsKQDHvy4VKgtzJ8/edit#heading=h.6a7pxu1gz13i
  - Tutorial at NeurIPS 2017: https://nips.cc/Conferences/2017/Schedule?showEvent=8732
  - Workshop at NeurIPS 2018: https://ppml-workshop.github.io/ppml/
  - TensorFlow Privacy: https://medium.com/tensorflow/introducing-tensorflow-privacy-learning-with-d
  - Background on Deep Learning: Stanford cs231 lecture notes,
  - DP as a protection against overfitting: Dwork et al. [2015a], Bassily et al. [2016]
  - Output perturbation and objective perturbation: Chaudhuri et al. [2011].
  - Differentially private PAC learning, the exponential mechanism for differentially private learning, and the equivalence between SQ learning and local DP learning: Vadhan [2017, Ch. 8], Kasiviswanathan et al. [2011].
  - Negative results for differentially private PAC learning (requires finite data universes even for simple models like threshold functions, can require computing time exponential in dimensionality): Bun and Zhandry [2016], Alon et al. [2018]
  - Deep nets can memorize their training data: Zhang et al. [2017], Carlini et al. [2018] (See also Membership Inference attacks on ML from the Attacks section of the course.)
  - The $\|\cdot\|$-norm mechanism: Hardt and Talwar [2010]
  - Concentrated differential privacy and variants: Dwork and Rothblum [2016], Bun and Steinke [2016], Mironov [2017], Abadi et al. [2016]
  - Differentially private gradient descent and stochastic gradient descent in the centralized and local models: Williams and Mcsherry [2010], Jain et al. [2012], Song et al. [2013], Bassily et al. [2014], Abadi et al. [2016], Duchi et al. [2014], Smith et al. [2017] (The theorems about utility are for convex loss functions, but the algorithms are DP even for non-convex loss functions.)
  - Thorough experimental evaluation and critique of differentially private machine learning and attacks: Jayaraman and Evans [2019].
  - Background on machine learning (without privacy): Kearns and Vazirani [1994], Stanford cs231 lecture notes, Deep learning tutorial, Tensorflow visual demo

- Software
  - OpenDP: http://opendp.org/
  - DualQuery: https://github.com/ejgallego/dualquery
  - MWEM: https://github.com/mrtzh/PrivateMultiplicativeWeights.jl
  - PinQ: https://www.microsoft.com/en-us/download/details.aspx?id=52363
  - εktelo: https://ektelo.github.io/
  - TensorFlow Privacy: https://github.com/tensorflow/privacy
  - FLEX (SQL, deployed by Uber): http://www.uvm.edu/~jnear/elastic/
  - PSI: http://psiprivacy.org/about/

- LightDP: https://github.com/RyanWangGit/lightdp
- RAPPOR: https://github.com/google/rappor
- Prochlo: https://github.com/google/prochlo
- DPComp (for comparing DP algorithms): https://www.dpcomp.org/
- Membership Inference Attacks: https://www.comp.nus.edu.sg/~reza/files/datasets.html
- DiffPriv (Easy Differential Privacy): https://cran.r-project.org/web/packages/diffpriv/index.html
- DPML (Differentially Private Convex Optimization, including SGD): https://github.com/sunblaze-ucb/dpml-benchmark

References

Martin Abadi, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, pages 308–318, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4139-4. doi: 10.1145/2976749.2978318. URL http://doi.acm.org/10.1145/2976749.2978318.

John M. Abowd. The u.s. census bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, pages 2867–2867, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5552-0. doi: 10.1145/3219819.3226070. URL https://digitalcommons.ilr.cornell.edu/ldi/49/.

Aws Albarghouthi and Justin Hsu. Synthesizing coupling proofs of differential privacy. *Proc. ACM Program. Lang.*, 2(POPL):58:1–58:30, December 2017. ISSN 2475-1421. doi: 10.1145/3158146. URL http://doi.acm.org/10.1145/3158146.

Noga Alon, Roi Livni, Maryanthe Malliaris, and Shay Moran. Private PAC learning implies finite littlestone dimension. *CoRR*, abs/1806.00949, 2018. URL http://arxiv.org/abs/1806.00949.

David W. Archer, Dan Bogdanov, Liina Kamm, Y. Lindell, Kurt Nielsen, Jakob Illeborg Pagter, Nigel P. Smart, and Rebecca N. Wright. From keys to databases – real-world applications of secure multi-party computation. Cryptology ePrint Archive, Report 2018/450, 2018. https://eprint.iacr.org/2018/450.

Victor Balcer and Salil Vadhan. Differential Privacy on Finite Computers. In Anna R. Karlin, editor, *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*, volume 94 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 43:1–43:21, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 978-3-95977-060-6. doi: 10.4230/LIPIcs.ITCS.2018.43. URL http://drops.dagstuhl.de/opus/volltexte/2018/8353.

Borja Balle, James Bell, Adrià Gascón, and Kobbi Nissim. The privacy blanket of the shuffle model. *CoRR*, abs/1903.02837, 2019. URL http://arxiv.org/abs/1903.02837.

Michael Barbaro and Tom Zeller, Jr. A face is exposed for AOL searcher no. 4417749. *The New York Times*, 9 August 2006. URL https://www.nytimes.com/2006/08/09/technology/09aol.html.

Gilles Barthe, Marco Gaboardi, Justin Hsu, and Benjamin Pierce. Programming language techniques for differential privacy. *ACM SIGLOG News*, 3(1):34–53, February 2016. ISSN 2372-3491. doi: 10.1145/2893582.2893591. URL http://doi.acm.org/10.1145/2893582.2893591.

Raef Bassily, Adam Smith, and Abhradeep Thakurta. Private empirical risk minimization: efficient algorithms and tight error bounds. In *55th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2014*, pages 464–473. IEEE Computer Soc., Los Alamitos, CA, 2014. doi: 10.1109/FOCS.2014.56. URL http://dx.doi.org/10.1109/FOCS.2014.56.

Raef Bassily, Kobbi Nissim, Adam Smith, Thomas Steinke, Uri Stemmer, and Jonathan Ullman. Algorithmic stability for adaptive data analysis. In *48th Annual Symposium on the Theory of Computing (STOC'16)*, June 2016. Preliminary version available at http://arxiv.org/abs/1511.02513.

Andrea Bittau, Úlfar Erlingsson, Petros Maniatis, Ilya Mironov, Ananth Raghunathan, David Lie, Mitch Rudominer, Ushasree Kode, Julien Tinnes, and Bernhard Seefeld. Prochlo: Strong privacy for analytics in the crowd. In *Proceedings of the 26th Symposium on Operating Systems Principles*, SOSP '17, pages 441–459, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-5085-3. doi: 10.1145/3132747.3132769. URL http://doi.acm.org/10.1145/3132747.3132769.

Graeme Blair, Kosuke Imai, and Yang-Yang Zhou. Design and analysis of the randomized response technique. *Journal of the American Statistical Association*, 110(511):1304–1319, 2015.

Avrim Blum, Cynthia Dwork, Frank McSherry, and Kobbi Nissim. Practical privacy: the SuLQ framework. In *Proceedings of the 24th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 128–138. ACM, 2005.

Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. *CoRR*, abs/1605.02065, 2016. URL http://arxiv.org/abs/1605.02065.

Mark Bun and Mark Zhandry. Order-revealing encryption and the hardness of private learning. In *Theory of Cryptography Conference (TCC '16A)*, pages 176–206. Springer, 10–13 January 2016. Full version available at https://eprint.iacr.org/2015/417.

Mark Bun, Cynthia Dwork, Guy N. Rothblum, and Thomas Steinke. Composable and versatile privacy via truncated cdp. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, pages 74–86, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5559-9. doi: 10.1145/3188745.3188946. URL http://doi.acm.org/10.1145/3188745.3188946.

Nicholas Carlini, Chang Liu, Jernej Kos, Úlfar Erlingsson, and Dawn Song. The secret sharer: Measuring unintended neural network memorization & extracting secrets. *CoRR*, abs/1802.08232, 2018. URL http://arxiv.org/abs/1802.08232.

Ann Cavoukian and Daniel Castro. Big data and innovation, setting the record straight: De-identification *does* work. Information and Privacy Commissioner, Ontario, Canada, June 2014. https://itif.org/publications/2014/06/16/setting-record-straight-de-identification-does-work/.

Ann Cavoukian and Khaled El Emam. De-identification protocols: Essential for protecting privacy. Information and Privacy Commissioner, Ontario, Canada, June 2014. https://www.ipc.on.ca/resource/de-identification-protocols-essential-for-protecting-privacy/.

Kamalika Chaudhuri, Claire Monteleoni, and Anand D. Sarwate. Differentially private empirical risk minimization. *J. Mach. Learn. Res.*, 12:1069–1109, July 2011. ISSN 1532-4435. URL http://dl.acm.org/citation.cfm?id=1953048.2021036.

Raj Chetty and John Friedman. A practical method to reduce privacy loss when disclosing statistics based on small samples. *American Economic Review Papers and Proceedings*, May 2019. URL https://opportunityinsights.org/paper/. To appear.

Raj Chetty, John Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. The opportunity atlas: Mapping the childhood roots of social mobility. Working Paper, October 2018. URL https://opportunityinsights.org/paper/.

Albert Cheu, Adam Smith, Jonathan Ullman, David Zeber, and Maxim Zhilyaev. Distributed differential privacy via shuffling. Cryptology ePrint Archive, Report 2019/245, 2019. https://eprint.iacr.org/2019/245.

Aloni Cohen. Attacks on deidentification's defenses. Working paper, October 2021. https://aloni.net/research/.

Aloni Cohen and Kobbi Nissim. Linear program reconstruction in practice. *CoRR*, abs/1810.05692, 2018. URL http://arxiv.org/abs/1810.05692.

P3G Consortium, George Church, Catherine Heeney, Naomi Hawkins, Jantina de Vries, Paula Boddington, Jane Kaye, Martin Bobrow, and Bruce Weir. Public access to genome-wide data: Five views on balancing research with privacy and protection. *PLOS Genetics*, 5(10):1–4, 10 2009. doi: 10.1371/journal.pgen.1000665. URL https://doi.org/10.1371/journal.pgen.1000665.

Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, 3rd edition, 2009. ISBN 0262033844, 9780262033848. URL https://mitpress.mit.edu/books/introduction-algorithms-third-edition.

Graham Cormode, Somesh Jha, Tejas Kulkarni, Ninghui Li, Divesh Srivastava, and Tianhao Wang. Privacy at scale: Local differential privacy in practice. In *Proceedings of the 2018 International Conference on Management of Data*, SIGMOD '18, pages 1655–1658, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-4703-7. doi: 10.1145/3183713.3197390. URL http://doi.acm.org/10.1145/3183713.3197390.

Aref N. Dajani, Amy D. Lauger, Phyllis E. Singer, Daniel Kifer, Jerome P. Reiter, Ashwin Machanavajjhala, Simson L. Garfinkel, Scot A. Dahl, Matthew Graham, Vishesh Karwa, Hang Kim, Philip Leclerc, Ian M. Schmutte, William N. Sexton, Katherine J. Thompson, Lars Vilhuber, and John M. Abowd. The modernization of statistical disclosure limitation at the u.s. census bureau. UNECE/EUROSTAT Work Session on Data Confidentiality, September 2017. URL https://www.unece.org/index.php?id=43931.

Jon P. Daries, Justin Reich, Jim Waldo, Elise M. Young, Jonathan Whittinghill, Andrew Dean Ho, Daniel Thomas Seaton, and Isaac Chuang. Privacy, anonymity, and big data in the social sciences. *Communication of the ACM*, 57(9):56–63, September 2014. ISSN 0001-0782. doi: 10.1145/2643132. URL http://doi.acm.org/10.1145/2643132.

Damien Desfontaines. A list of real-world uses of differential privacy. https://desfontain.es/privacy/real-world-differential-privacy.html, 2021. Blog post (last updated 10/21).

Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, pages 3574–3583, 2017. URL http://papers.nips.cc/paper/6948-collecting-telemetry-data-privately.

Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *Proceedings of the Twenty-second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, PODS '03, pages 202–210, New York, NY, USA, 2003. ACM. doi: 10.1145/773153.773173.

John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. Privacy aware learning. *Journal of the ACM*, 61(6):Art. 38, 57, 2014. ISSN 0004-5411. doi: 10.1145/2666468. URL http://dx.doi.org/10.1145/2666468.

Cynthia Dwork and Kobbi Nissim. Privacy-preserving datamining on vertically partitioned databases. In *Advances in Cryptology–CRYPTO 2004*, pages 528–544. Springer, 2004.

Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2013. ISSN 1551-305X. doi: 10.1561/0400000042. URL http://dx.doi.org/10.1561/0400000042.

Cynthia Dwork and Guy N. Rothblum. Concentrated differential privacy. *CoRR*, abs/1603.01887, 2016. URL http://arxiv.org/abs/1603.01887.

Cynthia Dwork, Moni Naor, Omer Reingold, Guy N. Rothblum, and Salil Vadhan. On the complexity of differentially private data release: Efficient algorithms and hardness results. In *Proceedings of the Forty-first Annual ACM Symposium on Theory of Computing*, STOC '09, pages 381–390, New York, NY, USA, 2009. ACM.

Cynthia Dwork, Guy Rothblum, and Salil Vadhan. Boosting and differential privacy. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS '10)*, pages 51–60. IEEE, 23–26 October 2010.

Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Roth. The reusable holdout: Preserving validity in adaptive data analysis. *Science*, 349(6248):636–638, 2015a. ISSN 0036-8075. doi: 10.1126/science.aaa9375. URL https://science.sciencemag.org/content/349/6248/636.

Cynthia Dwork, Adam Smith, Thomas Steinke, Jonathan Ullman, and Salil Vadhan. Robust traceability from trace amounts. In *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, FOCS '15, pages 650–669, Washington, DC, USA, 2015b. IEEE Computer Society. ISBN 978-1-4673-8191-8. doi: 10.1109/FOCS.2015.46. URL http://dx.doi.org/10.1109/FOCS.2015.46.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. *Journal of Privacy and Confidentiality*, 7(3), 2016. To appear. Preliminary version in *Proc. TCC '06*.

Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman. Exposed! a survey of attacks on private data. *Annual Review of Statistics and Its Application*, 4(1):61–84, 2017. doi: 10.1146/annurev-statistics-060116-054123. URL https://doi.org/10.1146/annurev-statistics-060116-054123.

Hamid Ebadi and David Sands. Featherweight PINQ. *Journal of Privacy and Confidentiality*, 7(2), 2017. doi: 10.29012/jpc.v7i2.653. URL https://doi.org/10.29012/jpc.v7i2.653.

Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, CCS '14, pages 1054–1067, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2957-6. doi: 10.1145/2660267.2660348. URL http://doi.acm.org/10.1145/2660267.2660348.

Úlfar Erlingsson, Vitaly Feldman, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Abhradeep Thakurta. Amplification by shuffling: From local to central differential privacy via anonymity. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, pages 2468–2479, Philadelphia, PA, USA, 2019. Society for Industrial and Applied Mathematics. URL http://dl.acm.org/citation.cfm?id=3310435.3310586.

Matthew Fredrikson, Eric Lantz, Somesh Jha, Simon Lin, David Page, and Thomas Ristenpart. Privacy in pharmacogenetics: An end-to-end case study of personalized warfarin dosing. In Kevin Fu and Jaeyeon Jung, editors, *Proceedings of the 23rd USENIX Security Symposium, San Diego, CA, USA, August 20-22, 2014.*, pages 17–32. USENIX Association, 2014. URL https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/fredrikson_matthew.

Marco Gaboardi, James Honaker, Gary King, Kobbi Nissim, Jonathan Ullman, and Salil P. Vadhan. PSI (Ψ): a private data sharing interface. *CoRR*, abs/1609.04340, 2016. URL http://arxiv.org/abs/1609.04340.

Marco Gaboardi, Emilio Gallego Arias, Justin Hsu, Aaron Roth, and Zhiwei Wu. Dual query: Practical private query release for high dimensional data. *Journal of Privacy and Confidentiality*, 7(2), 2017. doi: 10.29012/jpc.v7i2.650. URL https://doi.org/10.29012/jpc.v7i2.650.

Srivatsava Ranjit Ganta, Shiva Prasad Kasiviswanathan, and Adam Smith. Composition attacks and auxiliary information in data privacy. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, pages 265–273, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-193-4. doi: 10.1145/1401890.1401926. URL http://doi.acm.org/10.1145/1401890.1401926.

Simson Garfinkel, John M. Abowd, and Christian Martindale. Understanding database reconstruction attacks on public data. *Queue*, 16(5):50:28–50:53, October 2018a. ISSN 1542-7730. doi: 10.1145/3291276.3295691. URL http://doi.acm.org/10.1145/3291276.3295691.

Simson L. Garfinkel. Challenges and experiences adapting differentially private mechanisms to the 2020 census. Federal Committee on Statistical Methodology (FCSM) Research and Policy Conference, March 2018. URL http://www.copafs.org/seminars/fcsm2018.aspx.

Simson L. Garfinkel, John M. Abowd, and Sarah Powazek. Issues encountered deploying differential privacy. In *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*, WPES'18, pages 133–137, New York, NY, USA, 2018b. ACM. ISBN 978-1-4503-5989-4. doi: 10.1145/3267323.3268949. URL http://doi.acm.org/10.1145/3267323.3268949.

A. Ghosh, T. Roughgarden, and M. Sundararajan. Universally utility-maximizing privacy mechanisms. *SIAM Journal on Computing*, 41(6):1673–1693, 2012. doi: 10.1137/09076828X. URL https://doi.org/10.1137/09076828X.

Daniel Gingerich. Randomized response: Foundations and new developments. *Comparative Politics Newsletter (The Organized Section in Comparative Politics of the American Political Science Association)*, 25(1):16–27, 2015.

Daniel W Gingerich. Understanding off-the-books politics: Conducting inference on the determinants of sensitive behavior with randomized response surveys. *Political Analysis*, 18(3):349–380, 2010.

Philippe Golle. Revisiting the uniqueness of simple demographics in the us population. In *Proceedings of the 5th ACM Workshop on Privacy in Electronic Society*, WPES '06, pages 77–80, New York, NY, USA, 2006. ACM. ISBN 1-59593-556-8. doi: 10.1145/1179601.1179615. URL http://doi.acm.org/10.1145/1179601.1179615.

Andreas Haeberlen, Benjamin C. Pierce, and Arjun Narayan. Differential privacy under fire. In *Proceedings of the 20th USENIX Conference on Security*, SEC'11, pages 33–33, Berkeley, CA, USA, 2011. USENIX Association. URL http://dl.acm.org/citation.cfm?id=2028067.2028100.

Moritz Hardt and Guy N. Rothblum. A multiplicative weights mechanism for privacy-preserving data analysis. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 61–70, Oct 2010. doi: 10.1109/FOCS.2010.85.

Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *STOC'10—Proceedings of the 2010 ACM International Symposium on Theory of Computing*, pages 705–714. ACM, New York, 2010.

Moritz Hardt, Katrina Ligett, and Frank McSherry. A simple and practical algorithm for differentially private data release. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'12, pages 2339–2347, USA, 2012. Curran Associates Inc. URL http://dl.acm.org/citation.cfm?id=2999325.2999396.

Michael Hay, Vibhor Rastogi, Gerome Miklau, and Dan Suciu. Boosting the accuracy of differentially private histograms through consistency. *Proc. VLDB Endow.*, 3(1-2):1021–1032, September 2010. ISSN 2150-8097. doi: 10.14778/1920841.1920970. URL http://dx.doi.org/10.14778/1920841.1920970.

Michael Hay, Ashwin Machanavajjhala, Gerome Miklau, Yan Chen, and Dan Zhang. Principled evaluation of differentially private algorithms using dpbench. In *Proceedings of the 2016 International Conference on Management of Data*, SIGMOD '16, pages 139–154, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-3531-7. doi: 10.1145/2882903.2882931. URL http://doi.acm.org/10.1145/2882903.2882931.

Xi He, Ashwin Machanavajjhala, Cheryl Flynn, and Divesh Srivastava. Composing differential privacy and secure computation: A case study on scaling private record linkage. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, pages 1389–1406, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4946-8. doi: 10.1145/3133956.3134030. URL http://doi.acm.org/10.1145/3133956.3134030.

Nils Homer, Szabolcs Szelinger, Margot Redman, David Duggan, Waibhav Tembe, Jill Muehling, John V Pearson, Dietrich A Stephan, Stanley F Nelson, and David W Craig. Resolving individuals contributing trace amounts of dna to highly complex mixtures using high-density snp genotyping microarrays. *PLoS genetics*, 4(8):e1000167, 2008. URL https://doi.org/10.1371/journal.pgen.1000167.

Prateek Jain, Pravesh Kothari, and Abhradeep Thakurta. Differentially private online learning. In Shie Mannor, Nathan Srebro, and Robert C. Williamson, editors, *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pages 24.1–24.34, Edinburgh, Scotland, 25–27 Jun 2012. PMLR. URL http://proceedings.mlr.press/v23/jain12.html.

Bargav Jayaraman and David Evans. When relaxations go bad: "differentially-private" machine learning. *CoRR*, abs/1902.08874, 2019. URL http://arxiv.org/abs/1902.08874.

Noah Johnson, Joseph P. Near, and Dawn Song. Towards practical differential privacy for sql queries. *Proc. VLDB Endow.*, 11(5):526–539, January 2018. ISSN 2150-8097. doi: 10.1145/3187009.3177733. URL https://doi.org/10.1145/3187009.3177733.

Matthew Joseph, Aaron Roth, Jonathan Ullman, and Bo Waggoner. Local differential privacy for evolving data. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 2375–2384. Curran Associates, Inc., 2018. URL http://papers.nips.cc/paper/7505-local-differential-privacy-for-evolving-data.pdf.

P. Kairouz, S. Oh, and P. Viswanath. The composition theorem for differential privacy. *IEEE Transactions on Information Theory*, 63(6):4037–4049, June 2017. ISSN 0018-9448. doi: 10.1109/TIT.2017.2685505.

Shiva P. Kasiviswanathan and Adam Smith. On the 'semantics' of differential privacy: A bayesian formulation. *Journal of Privacy and Confidentiality*, 6(1), 2014.

Shiva Prasad Kasiviswanathan, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. What can we learn privately? *SIAM J. Comput.*, 40(3):793–826, 2011. doi: 10.1137/090756090.

Shiva Prasad Kasiviswanathan, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Analyzing graphs with node differential privacy. In *TCC*, pages 457–476, 2013. doi: 10.1007/978-3-642-36594-2_26. URL http://dx.doi.org/10.1007/978-3-642-36594-2_26.

Michael J. Kearns and Umesh V. Vazirani. *An introduction to computational learning theory*. MIT Press, Cambridge, MA, 1994. ISBN 0-262-11193-4.

Dan Kifer. Consistency with external knowledge: The topdown algorithm. Simons Insitute Workshop on "Data Privacy: From Foundations to Applications", March 2019. URL https://simons.berkeley.edu/talks/tba-31.

Aleksandra Korolova. Privacy violations using microtargeted ads: A case study. *Journal of Privacy and Confidentiality*, 3, 2011. URL https://doi.org/10.29012/jpc.v3i1.594.

Aleksandra Korolova, Krishnaram Kenthapadi, Nina Mishra, and Alexandros Ntoulas. Releasing search queries and clicks privately. In *Proceedings of the 18th International Conference on World Wide Web*, WWW '09, pages 171–180, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-487-4. doi: 10.1145/1526709.1526733. URL http://doi.acm.org/10.1145/1526709.1526733.

Ios Kotsogiannis, Yuchao Tao, Ashwin Machanavajjhala, Gerome Miklau, and Michael Hay. Architecting a differentially private SQL engine. In *CIDR 2019, 9th Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, January 13-16, 2019, Online Proceedings*. www.cidrdb.org, 2019. URL http://cidrdb.org/cidr2019/papers/p125-kotsogiannis-cidr19.pdf.

Chao Li, Gerome Miklau, Michael Hay, Andrew McGregor, and Vibhor Rastogi. The matrix mechanism: Optimizing linear counting queries under differential privacy. *The VLDB Journal*, 24(6):757–781, December 2015. ISSN 1066-8888. doi: 10.1007/s00778-015-0398-x. URL http://dx.doi.org/10.1007/s00778-015-0398-x.

Ashwin Machanavajjhala, Daniel Kifer, John Abowd, Johannes Gehrke, and Lars Vilhuber. Privacy: Theory meets practice on the map. In *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*, ICDE '08, pages 277–286, Washington, DC, USA, 2008. IEEE Computer Society. ISBN 978-1-4244-1836-7. doi: 10.1109/ICDE.2008.4497436. URL https://doi.org/10.1109/ICDE.2008.4497436.

Ashwin Machanavajjhala, Xi He, and Michael Hay. Differential privacy in the wild: A tutorial on current practices & open challenges. In *Proceedings of the 2017 ACM International Conference on Management of Data*, SIGMOD '17, pages 1727–1730, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-4197-4. doi: 10.1145/3035918.3054779. URL http://doi.acm.org/10.1145/3035918.3054779.

Frank McSherry. Privacy integrated queries: An extensible platform for privacy-preserving data analysis. *Communications of the ACM*, 53(9):89–97, September 2010. ISSN 0001-0782. doi: 10.1145/1810891.1810916. URL http://doi.acm.org/10.1145/1810891.1810916.

Frank McSherry. Statistical inference considered harmful. Blog post, 14 June 2016. URL https://github.com/frankmcsherry/blog/blob/master/posts/2016-06-14.md.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '07, pages 94–103, Washington, DC, USA, 2007. IEEE Computer Society. doi: 10.1109/FOCS.2007.41.

D. J. Mir, S. Isaacman, R. Cáceres, M. Martonosi, and R. N. Wright. Dp-where: Differentially private modeling of human mobility. In *2013 IEEE International Conference on Big Data*, pages 580–588, Oct 2013. doi: 10.1109/BigData.2013.6691626.

Ilya Mironov. On significance of the least significant bits for differential privacy. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, CCS '12, pages 650–661, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1651-4. doi: 10.1145/2382196.2382264. URL http://doi.acm.org/10.1145/2382196.2382264.

Ilya Mironov. Rényi differential privacy. In *30th IEEE Computer Security Foundations Symposium, CSF 2017, Santa Barbara, CA, USA, August 21-25, 2017*, pages 263–275. IEEE Computer Society, 2017. ISBN 978-1-5386-3217-8. doi: 10.1109/CSF.2017.11. URL https://doi.org/10.1109/CSF.2017.11.

Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, New York, NY, USA, 2nd edition, 2005. ISBN 0521835402. URL https://doi.org/10.1017/CBO9780511813603.

Prashanth Mohan, Abhradeep Thakurta, Elaine Shi, Dawn Song, and David Culler. Gupt: Privacy preserving data analysis made easy. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, SIGMOD '12, pages 349–360, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1247-9. doi: 10.1145/2213836.2213876. URL http://doi.acm.org/10.1145/2213836.2213876.

Jack Murtagh and Salil Vadhan. The complexity of computing the optimal composition of differential privacy. *Theory of Computing*, 14(8):1–35, 2018. doi: 10.4086/toc.2018.v014a008. URL http://www.theoryofcomputing.org/articles/v014a008.

Arvind Narayanan and Vitaly Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (S&P 2008), 18-21 May 2008, Oakland, California, USA*, pages 111–125. IEEE Computer Society, 2008. ISBN 978-0-7695-3168-7. doi: 10.1109/SP.2008.33. URL http://dx.doi.org/10.1109/SP.2008.33.

Arvind Narayanan and Vitaly Shmatikov. Myths and fallacies of "personally identifiable information". *Communications of the ACM*, 53(6):24–26, June 2010. ISSN 0001-0782. doi: 10.1145/1743546.1743558. URL http://doi.acm.org/10.1145/1743546.1743558.

Joseph P. Near and Chiké Abuah. *Programming Differential Privacy*. 2021. https://programming-dp.com/.

Sofya Raskhodnikova and Adam Smith. Private analysis of graph data. In Ming-Yang Kao, editor, *Encyclopedia of Algorithms*, pages 1–6. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. ISBN 978-3-642-27848-8. doi: 10.1007/978-3-642-27848-8_549-1. URL http://dx.doi.org/10.1007/978-3-642-27848-8_549-1.

Ryan Rogers, Aaron Roth, Jonathan Ullman, and Salil Vadhan. Privacy odometers and filters: Pay-as-you-go composition. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, pages 1929–1937, USA, 2016. Curran Associates Inc. ISBN 978-1-5108-3881-9. URL http://dl.acm.org/citation.cfm?id=3157096.3157312.

Kenneth Rosen. *Discrete Mathematics and its Applications*. McGraw Hill Education, 7th edition, 2012. URL https://www.mheducation.com/highered/product/discrete-mathematics-applications-rosen/M9780073383095.html.

Sheldon M. Ross. *A First Course in Probability*. Fifth edition, 1998.

Indrajit Roy, Srinath T. V. Setty, Ann Kilzer, Vitaly Shmatikov, and Emmett Witchel. Airavat: Security and privacy for mapreduce. In *Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation*, NSDI'10, pages 20–20, Berkeley, CA, USA, 2010. USENIX Association. URL http://dl.acm.org/citation.cfm?id=1855711.1855731.

Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. Membership inference attacks against machine learning models. In *2017 IEEE Symposium on Security and Privacy, SP 2017, San Jose, CA, USA, May 22-26, 2017*, pages 3–18. IEEE Computer Society, 2017. ISBN 978-1-5090-5533-3. doi: 10.1109/SP.2017.41. URL https://doi.org/10.1109/SP.2017.41.

Ryan Singel. Netflix cancels recommendation contest after privacy lawsuit. *Wired*, 12 March 2010. URL https://www.wired.com/2010/03/netflix-cancels-contest/.

A. Smith, A. Thakurta, and J. Upadhyay. Is interaction necessary for distributed private learning? In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 58–77, May 2017. doi: 10.1109/SP.2017.35.

Adam Smith and Jonathan Ullman. Privacy in statistics and machine learning. https://dpcourse.github.io/, 2022.

Daniel Solow. *How to Read and Do Proofs: An Introduction to Mathematical Thought Processes*. Wiley, 6th edition, 2013. URL https://www.wiley.com/en-us/How+to+Read+and+Do+Proofs%3A+An+Introduction+to+Mathematical+Thought+Processes%2C+6th+Edition-p-9781118164020.

S. Song, K. Chaudhuri, and A. D. Sarwate. Stochastic gradient descent with differentially private updates. In *2013 IEEE Global Conference on Signal and Information Processing*, pages 245–248, Dec 2013. doi: 10.1109/GlobalSIP.2013.6736861.

Latanya Sweeney. Weaving technology and policy together to maintain confidentiality. *The Journal of Law, Medicine & Ethics*, 25(2-3):98–110, 1997. doi: 10.1111/j.1748-720X.1997.tb01885.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1748-720X.1997.tb01885.x.

Latanya Sweeney. Simple demographics often identify people uniquely. Technical report, Technical report, Carnegie Mellon University, 2000. URL https://dataprivacylab.org/projects/identifiability/.

Latanya Sweeney. K-anonymity: A model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 10(5):557–570, October 2002. ISSN 0218-4885. doi: 10.1142/S0218488502001648. URL http://dx.doi.org/10.1142/S0218488502001648.

Latanya Sweeney, Akua Abu, and Julia Winn. Identifying participants in the personal genome project by name. Whitepaper 1021-1, Harvard University Data Privacy Lab, 13 April 2013. URL https://dataprivacylab.org/projects/pgp/.

Jun Tang, Aleksandra Korolova, Xiaolong Bai, Xueqiang Wang, and XiaoFeng Wang. Privacy loss in apple's implementation of differential privacy on macos 10.12. *CoRR*, abs/1709.02753, 2017. URL http://arxiv.org/abs/1709.02753.

Adam Tanner. Harvard professor re-identifies anonymous volunteers in DNA study. *Forbes*, 25 April 2013. URL https://www.forbes.com/sites/adamtanner/2013/04/25/harvard-professor-re-identifies-anonymous-volunteers-in-dna-study/#7bfba06192c9.

Jonathan Ullman and Salil Vadhan. PCPs and the hardness of generating private synthetic data. In *Theory of Cryptography*, pages 400–416. Springer, 2011.

Salil P. Vadhan. The complexity of differential privacy. In Yehuda Lindell, editor, *Tutorials on the Foundations of Cryptography — Dedicated to Oded Goldreich*, pages 347–450. Springer, 2017. ISBN 978-3-319-57047-1. doi: 10.1007/978-3-319-57048-8_7. URL https://doi.org/10.1007/978-3-319-57048-8_7.

Oliver Williams and Frank Mcsherry. Probabilistic inference and differential privacy. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 2451–2459. Curran Associates, Inc., 2010. URL http://papers.nips.cc/paper/3897-probabilistic-inference-and-differential-privacy.pdf.

Alexandra Wood, Micah Altman, Aaron Bembenek, Mark Bun, Marco Gaboardi, James Honaker, Kobbi Nissim, David R. O'Brien, Thomas Steinke, and Salil Vadhan. Differential privacy: A primer for a non-technical audience. *Vanderbilt Journal of Entertainment & Technology Law*, 21(1): 209–275, 2018. URL http://www.jetlaw.org/journal-archives/volume-21/volume-21-issue-1/differential-privacy-a-primer-for-a-non-technical-audience/. Preliminary version workshopped at PLSC 2017.

Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. Understanding deep learning requires rethinking generalization. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. URL https://openreview.net/forum?id=Sy8gdB9xx.

Dan Zhang, Ryan McKenna, Ios Kotsogiannis, Michael Hay, Ashwin Machanavajjhala, and Gerome Miklau. Ektelo: A framework for defining differentially-private computations. In *Proceedings of the 2018 International Conference on Management of Data*, SIGMOD '18, pages 115–130, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-4703-7. doi: 10.1145/3183713.3196921. URL http://doi.acm.org/10.1145/3183713.3196921.

Danfeng Zhang and Daniel Kifer. Lightdp: Towards automating differential privacy proofs. *SIGPLAN Notices*, 52(1):888–901, January 2017. ISSN 0362-1340. doi: 10.1145/3093333.3009884. URL http://doi.acm.org/10.1145/3093333.3009884.

Amitai Ziv. Israel's 'anonymous' statistics surveys aren't so anonymous. *Haaretz*, 7 January 2013. URL https://www.haaretz.com/surveys-not-as-anonymous-as-respondents-think-1.5288950.