# Comment on "NIST SP 800-226: Guidelines for Evaluating Differential Privacy Guarantees"

Rachel Cummings[1], Shlomi Hod[1,2], Gabriel Kaptchuk[2,3], Priyanka Nanayakkara[4], Jayshree Sarathy[1], and Jeremy Seeman[5]

[1]Columbia University, {rac2239,js6514}@columbia.edu
[2]Boston University, shlomi@bu.edu
[3]University of Maryland, College Park, kaptchuk@umd.edu
[4]Northwestern University, priyankan@u.northwestern.edu
[5]University of Michigan, jhseeman@umich.edu

March 2024

Differential privacy (DP) [14] has become a widely accepted framework for preserving privacy of data subjects while enabling statistical analyses. The rapid rise of DP, and its quick transition from theory to practice, has exposed the many open questions around how to deploy this framework *effectively*. In particular, existing deployments have faced several hurdles and raised questions related to risk assessment, parameter selection, communication, and interaction with other privacy and security practices. Without finding compelling ways to answer these questions, we risk DP failing to realize its promise, or, even worse, being used to *privacy wash* data systems without actually offering data subjects with meaningful privacy guarantees.

We are excited to see NIST's willingness to engage with these important questions within the special publication, "NIST SP 800-226: Guidelines for Evaluating Differential Privacy Guarantees." We believe that the draft guidelines do an excellent job identifying many of these questions and providing a structure within which readers can start to answer them. In particular, the guidelines' explicit characterization of *privacy hazards* provides a set of actionable best practices that can guide future choices. The presentation of these hazards (as short entry points into more complex conversations) also make the document highly approachable.

In this comment, we identify opportunities for improvement within the draft document and provide specific recommendations on how the guidelines can be improved. We believe the the potential impact of this document is significant, as it can serve as the de-facto starting point for individuals, organizations, and government groups interested in exploring differential privacy. Our recommendations are aimed at maximizing this potential. Specifically, we suggest (1) expanding the discussion around $\epsilon$, (2) clarifying the intended use of the privacy pyramid, (3) replacing the running example with something more compelling, (4) engaging with the social dimensions of utility, (5) reworking the discussion on "human bias", and (6) highlighting the importance of communication in differential privacy deployments.

### Themes Within Our Recommendations

While the suggestions that we make are specific—and, we hope, actionable— we first identify several high-level themes that ground our recommendations. In doing so, we hope that these perspectives can be applied iteratively as the draft guidelines continue to develop.

(1) **Understanding data as social artifacts.** A tremendous amount of research within science and technology studies (and related disciplines such as human-centered computing, sociology, and communications) has documented the ways in which the collection, processing, and meaning-making of data are inherently social processes [5, 6, 13, 18, 22, 32]. For example, social processes shape fundamental decisions such as what data is collected and how that data is collected, as well as seemingly minute decisions such as the formats in which data are released. Similarly, the data shapes social contexts in which it is embedded. In other words, data and its use are not neutral, but are laden with social values. Within the context of this comment, we find embracing this framing provides an opportunity to re-evaluate the approach of the "Utility and Accuracy" section (Section 3.2) and "Human Bias" section (Section 3.3.2). Additionally, this framing highlights the need for social engagement *around* the collection of data and the use of differentially private statistics—*even when the technical aspects of the privacy protections are approached competently.*

(2) **Developing vs. evaluating differential privacy deployments.** The current draft of the guidelines is ambitious in its scope, in that it aims to "help practitioners of all backgrounds—including business owners, product managers, software engineers, data scientists, and academics—better understand how to think about differentially private software solutions." While all these parties might have a vested interest in differential privacy, we note that the point at which their interest is piqued might be *very* different. Specifically, the process by which a differentially private data release is *developed* is very different than the process of *evaluating* an existing differentially private data release. While the current version of the draft highlights concepts that will be valuable to individuals participating in both processes, this may create some confusion. For example, the current presentation of the pyramid implicitly encourages a reader to think about a deployment by *sequentially* iterating up through the levels of the pyramid. While this might be appropriate when evaluating an existing deployment, it obscures the fact that a more iterative approach might be necessary when developing a new deployment.

(3) **The need for compelling case studies.** Throughout the existing draft, we found the opportunity for integrating two types of case studies. The first type of case study would serve to illustrate the real-world impact associated with deploying differential privacy. This is not achieved in the current draft; unfortunately the existing running example surrounding pumpkin spice lattes risks minimizes the risks associated with real-world data processing pipelines and misses an opportunity to present the stakes involved with privacy-sensitive systems. Case studies in which potential harms are described will make the existing text more compelling. The second type of case study highlights the *use* of the tools and strategies discussed in the guidelines. Case studies that *show* how to apply the guidelines will lower the barriers associated with their use.

### Questions with which our recommendations engage.

The draft document contains six explicit questions for reviewer comment. While we do not organize our suggestions specifically around these questions, many of our suggestions engage directly with them. For convenience, we repeat these questions here with numbering system and annotate the beginning of each suggestion with the questions to which our suggestion is most related.

**(Q1)** *Does this publication have a clear and appropriate scope?*

**(Q2)** *Is this publication understandable for the intended audience?*

**(Q3)** *Does publication provide a conceptual framework for understanding the uses and pitfalls of differential privacy? Is there any guidance that is not well-founded?*

**(Q4)** *Is the differential privacy pyramid a helpful conceptual device?*

**(Q5)** *Are the privacy hazards described accurately? Should additional hazards be added?*

**(Q6)** *For topics where the research is inconclusive, were any key points missed from the literature?*

**About the Authors**

We are a group of researchers at US universities who study differential privacy from a wide array of perspectives. Our work tackles issues across the pipeline of differential privacy deployments, including conceptual frameworks [43], algorithmic design and statistical inference [2,12,39], communication [7,10,27,28], practical hurdles [35,38], and policy perspectives [4,37], including conducting a real-world DP release [23]. Collectively, we have expertise in security & privacy, cryptography, statistics, machine learning, human computer interaction, communication, and science & technology studies. We are particularly interested in understanding the barriers to deployment of DP and advocating for holistic, context-aware approaches to overcome these barriers.

## Suggestion 1: Expand the discussion around $\epsilon$ to engage with real-world risks

### (Relevant to Q3 and Q6)

Differential privacy is an established approach to address the privacy risks associated with statistical analysis. Contextual and robust evaluation of the differential privacy guarantees requires grounding them within these risk. This calls for (1) a clear and well-defined articulation of risks; and (2) an interpretation of the guarantees in terms of the risks.. Without that, the evaluation might render incomplete.

Privacy risks are formulated differently in different fields, such as law and computer science. We suggest beginning with the definition of privacy risks as described in the technical literature on differential privacy, characterized by attacks like reconstruction, attribute inference, singling-out, and membership inference [34]. Articulating threat models and attack specifics within a concrete deployment context can help clarify the implications of the differential privacy guarantee. This particularly applies to the assessment of privacy loss parameters (e.g., $\varepsilon$, $\delta$, $\rho$). Evidence suggests that these parameters can be challenging to interpret [11,28], especially in rich and complex releases, such as those that involve many queries or complex models. While the document discusses different aspects such as the threat model (global/local, neighborhood datasets), we encourage NIST to engage more deeply in with the meaning of $\varepsilon$, especially because there is no established consensus.

More concretely, we have two suggestions. First, we propose including a comprehensive and structured discussion of those technical privacy risks identified in the literature. Second, we believe that practitioners would benefit from guidance on how to relate privacy loss parameters to attack parameters. Although this question is an active of research, we can still outline two emerging threads on how to do that: (1) deriving theoretical bounds on attack success from privacy loss parameters (see [3,9,29,34] for some examples); and (2) conducting empirical estimation of attack success based on assumption on the adversary [25,26,29].

**Case study of choosing privacy parameters or determining if a given privacy parameter is appropriate**

Initiating a discussion on the meaning of $\varepsilon$ within the publication would already provide value to practitioners. However, to make these considerations more actionable and comprehensible, we believe that a case study would be highly beneficial. Specifically, such a case study could illustrate how to conduct a technical privacy risk analysis of a deployment in terms of attacks and select which available tools (theoretical or empirical) could assist in interpreting the privacy loss parameters.

## Suggestion 2: Clarify intended use of the privacy pyramid

### (Relevant to Q1 and Q4)

In interrogating the privacy pyramid in Figure 1, we observe that it is easy to imagine it being used as a conceptual device within multiple different settings:

- A *pedagogical tool* for a readers first learning about the multiple, interlocking assumptions that underpin a differential privacy guarantee;
- A *process oriented tool* that is intended to guide relevant parties as they are designing a new system which they intend to provide an differential privacy guarantee;
- An *evaluative tool* that should be filled out when attempting to understand the differential privacy guarantee provided by an existing data release; or
- A *comparative* tool that should support someone attempting to choose between two different differential privacy guarantees.

However, it is current form, it does not appear to be well suited to all of these tasks. For example, while the privacy pyramid skillfully illustrates the different technical components of differential privacy implementations and the degree to which differential privacy's formal guarantees rely on these components' technical features, it does not lend itself to non-pedagogical uses. We recommend ensuring, both in the text and in the figure, that the pyramid is a descriptor of the technical guarantee and not an exhaustive description of every policy dimension needed to implement differential privacy. Moreover, NIST may also want to consider being more explicit in how it expects readers to user this conceptual tool or expand this conceptual tool to be valuable in multiple of these settings.

**Differentiate between technical dependencies and order of operations**

Settling on a particular differential privacy guarantee is often a consequence of considering different possible "elements" of the pyramid as policy alternatives. These decisions are often made in an iterative fashion (for example, when the Census Bureau changed from $(\epsilon, \delta)$-DP to $\rho$-zCDP, or when they changed which queries would be answered at all). The structure of the pyramid may, however, imply that decisions need to be made in a particular order (i.e., moving "up" the pyramid). In practice, organizations may be required to reconsider alternative strategies on lower, previously "established," levels of the the pyramid to meet privacy and/or utility needs. Possible changes to the pyramid include:

- *Placement of "utility and bias"*: while these evaluations are certainly important from an implementation perspective, they do not imply any particular technical dependencies for the DP guarantee. Do these necessarily belong as part of the pyramid?
- *Ordering of "algorithms & correctness" and "unit of privacy"*: once a threat model is specified, different algorithms can satisfy DP guarantees for different units of privacy (for example,

by group composition properties); however, what an algorithm does theoretically could be divorced from its correct implementation. Does it make sense to separate theoretical algorithm descriptions versus particular implementations on actual systems with pre-specified security architectures, floating-point precisions, etc.?

**Introduce a *process oriented* conceptual tool**
The modifications that we proposed to the pyramid suggest that there are separate policy dimensions to the different elements of the pyramid (i.e., comparing multiple pyramids with components interchanged). These decision-making properties are themselves important for the pyramid; for example, selecting a privacy loss budget based on a public datasource versus the confidential data in question introduces a trade-off between a potential side-channel vulnerability and the ability to contextually identify privacy, utility, and bias specific to the data processing task at hand. Similarly, different modes of communicating about privacy loss and/or the decision-making process can affect the pyramid as well (for example, does a particular privacy guarantee sufficiently meet user expectations for data protection?). Incorporating dimensions like these suggests that the pyramid could be "annotated" at each level with considerations about how an organization might choose between different versions of each pyramid layer. Moreover, by framing the considerations as specific policy questions, the guidance provided by NIST will remain sufficiently objective for comparing different DP guarantees to one another and the degree to which they align with particular formal privacy and/or data utility goals.

**Case study of using the privacy pyramid**
Finally, we observe that there is an opportunity to include a case study in the document on how to "apply" the privacy pyramid to a particular deployment of differential privacy. This case study could either be a real-world deployment if there exists a real-world deployment that NIST feels comfortable including in the document. More likely, a synthetic use case could be fabricated that is similar but distinct to an existing deployment. A version of the pyramid in which the concrete choice made for each brick could be shown. If NIST intends for the pyramid to be valuable in the *comparative* setting, then a second version of the pyramid could be prepared and a comparison could be done.

## Suggestion 3: Make the running example more compelling

### (Relevant to Q3)

Within the cryptographic community there is an increased effort to shift the descriptions of technologies from toy examples (e.g., "Alice wants to send a message to Bob") into more plausible examples that make the stakes of a technological deployment more tangible. This shift is commonly attributed to Phil Rogaway's On The Moral Character of Cryptographic Work [33] and mirrors other scholarship highlighting the importance of emotional context when it comes to legitimizing the importance of privacy [40]. Rogaway highlights that the ways in which we frame problems has led to a shift in the *types* of problems on which the cryptographic community works. Heeding this call to action there have been a number of efforts to shift the norms on the examples used within the cryptographic community. For example, Glencora Boradille has recently written a new cryptography textbook titled Defend Dissent, which integrates examples of the ways in which cryptography can be used to protect the rights of people around the world. There are other similar shifts playing out in the security research community, including the ACM CCS suggesting that authors use names that reflect a greater array of the people in the world in their submissions (e.g., instead of Alice and Bob, use Alvarez and Bano), which can also be an opportunity to highlight communities who are most impacted by surveillance's harms [1].

It is in this spirit that we encourage a reworking of the current document to use a more concrete running example of a differential privacy deployment that engages with real harm. The current running example, regarding pumpkin spice lattes, might provide a convenient mathematical abstraction, but risks trivializing the purpose of using differential privacy in the first place. We believe that this change could be relatively lightweight and considerably increase the ways in which the document would resonate with potential audiences. Indeed, several other NIST documents take this more "impact first" approach to discussing privacy technologies [16, 17]. Moreover, the document actually already contains several examples of differential deployment that are more emotionally evocative in Section 2.4: protecting browsing history, taxi ride data, and electricity use. While a slightly more complex approach, these examples could also be expanded into full case studies that could recur throughout the more technical discussions.

## Suggestion 4: Engage with social dimensions of utility

We appreciate that the draft guidance highlights the distinction between accuracy and utility, as there is a tendency among researchers to use these two words interchangeably. We feel, however, that there is an opportunity to further clarify the difference between accuracy and utility, particularly by expanding the discussion about utility.

Utility is currently described as "how useful a dataset or statistic is for a specific purpose." This simple and elegant definition risks obscuring the social dimensions of utility and the ways in which specific purposes may vary significantly based on the data user and context. For example, utility may refer to the usefulness of a statistic in learning a population-wide attribute (e.g., rate of a particular medical condition among a subpopulation), or it could refer to something more ambiguous, like the ways in which the statistic (or set of statistics) can help inform policy around allocating federal funds in an equitable manner [41]. Put another way, the reason that there is no general solution to measuring utility is not only because utility is a multifaceted concept, but because it is an inherently *social* concept which may not permit meaningful, mathematical measurement. Moreover, because a statistic may be used post-publication in several decision making pipelines (as the current text acknowledges), it may be impossible to quantify utility in advance—an argument that has been made in previous research [24]. In order to address the social nature of utility, we suggest reworking its definition to more holistically capture how it refers to the usefulness of a dataset or statistic for *various societally-beneficial purposes.* Under this more nuanced definition, the guidelines should also note that utility may not always be easy to cleanly measure and requires both qualitative and quantitative engagement with data users.

Another way in which the current text could be updated to engage with the social nature utility is to expand the discussion in the "Metrics for Utility: No General Solution" subsection to include qualitative approaches. Specifically, there may be a significant number of cases where qualitative analysis methods will be much better at establishing the utility of a differential privacy deployment (either proposed are already realized). For example, it may be useful to collect qualitative data from people impacted by differentially-private data releases to better understand the impacts of a data release on various communities, data users, and society at large. Using this qualitative epistemology will permit rigorous analysis in settings where quantitative metrics are fundamentally limited.

## Suggestion 5: Re-frame discussion on "Human Bias" towards "Human Factors"

### (Relevant to Q2 and Q5)

Section 3.3.2 introduces the idea that the structures *surrounding* the data can impact the data itself. This is a critical observation that is too often overlooked within communities that are primarily focused on technology. The existing text highlights two known examples of "human biases" that can impact the ways that data is received: (1) the intentional injection of randomness can undermine *trust* in the data release, and (2) structural requirements on the "shape" of the data may demand postprocessing. We are excited to see this observation included in the document and believe that there is an opportunity to expand and improve this subsection.

While we understand that the term "Human Bias," along with taxonomizing bias into Systemic, Human, and Statistical, is inherited from NIST's recent Special Publication on Bias and Artificial Intelligence [36], the phrasing may be misleading in this context because of the negative valance associated with the word bias. The phrase "human biases" implies that there are factors that prevent human beings from seeing the facts for what they are—that is, there is a normative understanding that it is possible to be better in a particular circumstance if it were not for the mistaken reasoning of humans. In this context, however, we note that the two examples outlined in the text are not inherently instances of "bias." For example, if post-processing is not applied to a data release it may make the data release itself *effectively without utility*, as it may not be interoperable with existing tools or pipelines (which may be too expensive to replace). Similarly, the noise injected in a dataset might be calibrated in a way that the data release no longer has utility for certain tasks, i.e. the privacy parameter is set to high to preserve utility.

Instead, we propose that the existing text should be reformulated to emphasize that the structural, socially-created restrictions on data are unavoidable and are something that should be reckoned with explicitly. Put another way, data creation and processing is always the product of social processes [18,32], and differentially private data releases are no exception. These constraints shape what is possible for a differential privacy release, assuming that maintaining the utility of the data release is required.

Embracing this reformulation also provides the opportunity to discuss other "human factors" that might drive decisions when developing deployments of differential privacy. For example, the document could discuss the ways in which data has *affordances* [30]—an acknowledgment that the ways in which data is created, presented, described, and processed naturally lends the data to use in different ways [42]. This is true both in terms of the computation that can be done with the data (e.g. presenting data a particular file format, rounding numbers to be "reasonable" by typical data consumer standards, or choosing to generate synthetic data instead of tabular summary data can push data consumers into using it in particular ways) and in terms of the arguments that can be made with the data (e.g. choosing to explicitly include or exclude error metrics).

### Including a Case Study

Another useful tool for exploring the impact of structural and human factors on differential private data releases would be to include a case study. The existing text appears to be referencing the impact of these factors on the U.S. Census Bureau's 2020 data release and the complex effects on accuracy due to post-processing. The reference to this complex data release is apt for readers with a deep understanding of this existing case study, as significant amounts have been written about it from both sociological and technical perspectives [7, 8, 15, 19–21, 31, 41]. However, there is a risk that an uninitiated reader will not understand the immense impact of these factors, the unavoidable nature of these constraints, and the risk associated with ignoring them. While the details of the 2020 Census disclosure avoidance system are likely too complicated (and, perhaps, too controversial) for this document, we suggest that a “toy” example could be synthesized.

## Suggestion 6: Expand Discussion of Sociotechnical Considerations

We suggest expanding the discussion around the sociotechnical considerations of making a DP deployment. We realize that due to limited visibility of DP deployments in the private sector, it may be difficult at this time to fully understand the space of considerations that should be recommended. That said, we suggest using examples with more documentation (e.g., Census, Wikimedia) to generate a set of questions that organizations deploying DP might want to consider prior to the deployment. In particular, these questions would help facilitate more productive conversations around specific deployment decisions, how the data should be used moving forward considering added noise was added, etc. Some of these questions might include:

- *How does DP re-orient the data release pipeline, as experienced by data subjects and data analysts?* In some cases DP enables data to be released that previously were not available, while in others, DP introduces a new source of noise that previously was not present in the data. Data users and other stakeholders may have different concerns and responses to DP based on the answer to this question.

- *What are the threats that DP is intended to thwart in the specific use case, and to what extent do those align with data subjects' concerns?* We suggest that clear communication around the intended purpose of DP, beyond just ‘protecting privacy.’ Additionally, some discussion of how this intended purpose aligns with data subjects' concerns will make it easier to figure out whether DP is the right tool for the specific use case, and when it is, ensure smoother communication.

- *How might DP impact data release timelines and the costs associated with data release?* As we witnessed in the Census use case, delays in releasing data that resulted from hammering out the differential privacy machinery was cause for concern for some groups of data users. Hence, organizations should consider timelines in advance, and communicate those clearly to preempt downstream communication issues. Additionally, using differential privacy is not without costs, as external expertise must sometimes be acquired to construct or validate software.

## Conclusion

As researchers in the differential privacy community, we are excited to see NIST's guidelines around evaluating this technology and its claims around privacy and utility in real-world settings. In this comment, we offer suggestions to improve the accuracy and depth of these guidelines. In particular, we recommend (1) expanding the discussion around privacy-loss parameters, (2) clarifying the uses of the privacy pyramid, (3) using more compelling and realistic examples, (4) engaging with the social dimensions of utility, (5) reorienting the discussion on human bias, and (6) highlighting the importance of communication. With these improvements, we believe the draft will be able to provide more thorough guidance for parties seeking to evaluate differential privacy deployments.

# References

[1] ACMCCS 2024. Diversity and inclusion. `https://www.sigsac.org/ccs/CCS2024/diversity-and-inclusion.html`. Accessed on 28 Feb, 2024.

[2] Daniel Alabi, Audra McMillan, Jayshree Sarathy, Adam Smith, and Salil Vadhan. Differentially private simple linear regression. *Proceedings on Privacy Enhancing Technologies*, 2022(2):184–204, 2022.

[3] Borja Balle, Giovanni Cherubin, and Jamie Hayes. Reconstructing training data with informed adversaries. In *2022 IEEE Symposium on Security and Privacy*, pages 1138–1156, San Francisco, CA, USA, May 22–26, 2022. IEEE Computer Society Press.

[4] Sebastian Benthall and Rachel Cummings. Integrating differential privacy and contextual integrity. *arXiv preprint arXiv:2401.15774*, 2024.

[5] C Bowker Geoffrey. Data flakes: an afterword to raw data is an oxymoron. *Raw Data Is an Oxymoron*, pages 167–171, 2013.

[6] danah boyd and Kate Crawford. Six provocations for big data. In *A decade in internet time: Symposium on the dynamics of the internet and society*, 2011.

[7] danah boyd and Jayshree Sarathy. Differential perspectives: Epistemic disconnects surrounding the us census bureau's use of differential privacy. In *Harvard Data Science Review*, 2022.

[8] Aloni Cohen, Moon Duchin, JN Matthews, and Bhushan Suwal. Private numbers in public policy: Census, differential privacy, and redistricting. *Harvard Data Science Review*, (Special Issue 2), 2022.

[9] Aloni Cohen and Kobbi Nissim. Towards formalizing the gdpr's notion of singling out. *Proc. Natl. Acad. Sci. USA*, 117(15):8344–8352, 2020.

[10] Rachel Cummings, Gabriel Kaptchuk, and Elissa M Redmiles. "i need a better description": An investigation into user expectations for differential privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, pages 3037–3052, 2021.

[11] Rachel Cummings, Gabriel Kaptchuk, and Elissa M. Redmiles. "I need a better description": An investigation into user expectations for differential privacy. In Giovanni Vigna and Elaine Shi, editors, *ACM CCS 2021: 28th Conference on Computer and Communications Security*, pages 3037–3052, Virtual Event, Republic of Korea, November 15–19, 2021. ACM Press.

[12] Rachel Cummings, Sara Krehbiel, Kevin A Lai, and Uthaipon Tantipongpipat. Differential privacy for growing databases. *Advances in Neural Information Processing Systems*, 31, 2018.

[13] Tim Davies and Mark Frank. 'there's no such thing as raw data' exploring the socio-technical life of a government dataset. In *Proceedings of the 5th annual ACM web science conference*, pages 75–78, 2013.

[14] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In Shai Halevi and Tal Rabin, editors, *TCC 2006: 3rd Theory of Cryptography Conference*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284, New York, NY, USA, March 4–7, 2006. Springer, Heidelberg, Germany.

[15] John L Eltinge. Disclosure protection in the context of statistical agency operations: Data quality and related constraints. *Harvard Data Science Review. Published online June*, 24, 2022.

[16] National Institute for Standards and Technology. Nist privacy engingeering program: Collabroative research cycle. `https://pages.nist.gov/privacy_collaborative_research_cycle/index.html`. Accesssed on 28 Feb, 2024.

[17] National Institute for Standards and Technology. Nist privacy engingeering program: Prize challenges. `https://www.nist.gov/itl/applied-cybersecurity/privacy-engineering/collaboration-space/challenges/prize-challenges`. Accesssed on 28 Feb, 2024.

[18] Lisa Gitelman. *Raw data is an oxymoron*. MIT press, 2013.

[19] Ruobin Gong. Transparent privacy is principled privacy. *Harvard Data Science Review*, (Special Issue 2), 2022.

[20] Ruobin Gong, Erica L Groshen, and Salil Vadhan. Harnessing the known unknowns: Differential privacy and the 2020 census. *Harvard Data Science Review*, 6, 2022.

[21] Erica L Groshen and Daniel Goroff. Disclosure avoidance and the 2020 census: What do researchers need to know. *Harvard Data Science Review*, 2022.

[22] Ian Hacking. *The taming of chance*. Number 17. Cambridge University Press, 1990.

[23] Shlomi Hod. Designing the pilot release of israel's national registry of live births: Reconciling privacy with accuracy and usability. 2023.

[24] V Joseph Hotz, Christopher R Bollinger, Tatiana Komarova, Charles F Manski, Robert A Moffitt, Denis Nekipelov, Aaron Sojourner, and Bruce D Spencer. Balancing data privacy and usability in the federal statistical system. *Proceedings of the National Academy of Sciences*, 119(31):e2104906119, 2022.

[25] F Houssiau, J Jordon, SN Cohen, O Daniel, A Elliott, J Geddes, C Mole, C Rangel-Smith, and L Szpruch. Tapas: a toolbox for adversarial privacy auditing of synthetic data. 2022.

[26] Matthew Jagielski, Jonathan R. Ullman, and Alina Oprea. Auditing differentially private machine learning: How private is private sgd? In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin, editors, *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020.

[27] Priyanka Nanayakkara, Johes Bater, Xi He, Jessica Hullman, and Jennie Rogers. Visualizing privacy-utility trade-offs in differentially private data releases. *Proceedings on Privacy Enhancing Technologies*, 2:601–618, 2022.

[28] Priyanka Nanayakkara, Mary Anne Smart, Rachel Cummings, Gabriel Kaptchuk, and Elissa M. Redmiles. What are the chances? explaining the epsilon parameter in differential privacy. In Joseph A. Calandrino and Carmela Troncoso, editors, *32nd USENIX Security Symposium, USENIX Security 2023, Anaheim, CA, USA, August 9-11, 2023*, pages 1613–1630. USENIX Association, 2023.

[29] Milad Nasr, Shuang Song, Abhradeep Thakurta, Nicolas Papernot, and Nicholas Carlini. Adversary instantiation: Lower bounds for differentially private machine learning. In *2021 IEEE Symposium on Security and Privacy*, pages 866–882, San Francisco, CA, USA, May 24–27, 2021. IEEE Computer Society Press.

[30] Don Norman. *The design of everyday things: Revised and expanded edition.* Basic books, 2013.

[31] Daniel L Oberski and Frauke Kreuter. Differential privacy and social science: An urgent puzzle. *Harvard Data Science Review*, 2(1):1–21, 2020.

[32] Theodore M Porter. *Trust in numbers: The pursuit of objectivity in science and public life.* Princeton University Press, 1996.

[33] Phillip Rogaway. The moral character of cryptographic work. Cryptology ePrint Archive, Report 2015/1162, 2015. `https://eprint.iacr.org/2015/1162`.

[34] Ahmed Salem, Giovanni Cherubin, David Evans, Boris Köpf, Andrew Paverd, Anshuman Suri, Shruti Tople, and Santiago Zanella-Béguelin. SoK: Let the privacy games begin! A unified treatment of data inference privacy in machine learning. In *2023 IEEE Symposium on Security and Privacy*, pages 327–345, San Francisco, CA, USA, May 21–25, 2023. IEEE Computer Society Press.

[35] Jayshree Sarathy, Sophia Song, Audrey Haque, Tania Schlatter, and Salil Vadhan. Don't look at the data! how differential privacy reconfigures the practices of data science. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2023.

[36] Reva Schwartz, Apostol Vassilev, Kristen Greene, Lori Perine, Andrew Burt, and Patrick Hall. Nist special publication 1270: Towards a standard for identifying and managing bias in artificial intelligence. `https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf`, 2022. Accessed on 28 Feb, 2024.

[37] Jeremy Seeman. Framing effects in the operationalization of differential privacy systems as code-driven law. In *International Conference on Computer Ethics*, volume 1, 2023.

[38] Jeremy Seeman and Daniel Susser. Between privacy and utility: On differential privacy in theory and practice. *ACM Journal on Responsible Computing*, 2023.

[39] Aleksandra Slavković and Jeremy Seeman. Statistical data privacy: A song of privacy and utility. *Annual Review of Statistics and Its Application*, 10:189–218, 2023.

[40] Luke Stark. The emotional context of information privacy. *The Information Society*, 32(1):14–27, 2016.

[41] Ryan Steed, Terrance Liu, Zhiwei Steven Wu, and Alessandro Acquisti. Policy impacts of statistical uncertainty and privacy. *Science*, 377(6609):928–931, 2022.

[42] Langdon Winner. Do artifacts have politics? In *Computer ethics*, pages 177–192. Routledge, 2017.

[43] Wanrong Zhang, Olga Ohrimenko, and Rachel Cummings. Attribute privacy: Framework and mechanisms. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pages 757–766, 2022.