Journal of Privacy and Confidentiality
Vol. 14 (3) 2024

Noisy measurement workshop 2022

Submitted Feb 2023
Published August 2024

PERSPECTIVE: BETTER PRIVACY THEORISTS FOR BETTER DATA STEWARDS

JEREMY SEEMAN

Department of Statistics, Penn State University, 326 Thomas Building, University Park, PA 16802
e-mail address: jhs5496@psu.edu

Abstract. The U.S. Census Bureau's use of differential privacy (DP) fundamentally changed how academic DP researchers perform outreach to official statistics stakeholders. In this perspectives piece, I propose ways for us in this community to improve those processes by being more receptive to the practical concerns raised by building DP systems. First, I discuss how academic DP work fundamentally differs from the policy decisions needed to implement DP systems, and why this distinction has political consequences. Through examples and discussions from workshops, I show how the DP community largely asked applied stakeholders to communicate on DP's theoretical terms, when such a request foreclosed important considerations relevant for the Census Bureau's policy problems. Second, I discuss how existing polarization between theoretical and empirical privacy researchers has unintentionally seeped into the ways we communicated about DP, pointing to why both perspectives are necessary, in different ways, for policy conversations. Finally, I conclude by discussing how these issues are not unique to data privacy work, but instead reflect structural problems in translating theoretical science into practice. These ideas are presented in service of a single goal: to ensure DP theory supports substantive, privacy-aware data processing and dissemination in practice for data curators.

In 2018, the U.S. Census Bureau announced its decision to switch from traditional statistical disclosure limitation (SDL) methods to formally private (FP) methods based on differential privacy (DP) in releasing their 2020 Decennial Census data products [Abowd, 2018]. Today, in 2023, we see how such a change unintentionally revealed intertwined political and epistemological fissures in official statistics as data infrastructure, prompting reflection among numerous stakeholder communities. For data users, DP makes transparent the need to address official statistics as measurements and not enumerations [Boyd and Sarathy, 2022]. For policymakers, DP makes tangible new privacy risks associated with large-scale public use data products [Nanayakkara and Hullman, 2023]. For statistical disclosure limitation (SDL) researchers studying empirical disclosure risk, DP makes clear the shortcomings of purely empirical approaches to data privacy [Abowd and Hawes, 2023]. All of these arguments demonstrate how introducing DP has prompted constructive, consequential, and long-overdue insights into how we address privacy problems for official statistics.

At the same time, when DP scholars champion the impact of DP on real-world data processing systems, we may be missing opportunities to reflect on what we can learn from other disciplines and how we can better support those who build and use DP systems. Privacy

Key words and phrases: differential privacy, public policy.

www.journalprivacyconfidentiality.org
DOI:10.29012/jpc.865

scholarship is inherently interdisciplinary, and successful interdisciplinary work demands enough humility to recognize the shortcomings of any one perspective and remain open to two-way communication across disciplinary boundaries. More so than ever, the academic DP community has done substantial outreach to a wide variety of Census stakeholders, but our communications have been largely one-directional.

In this perspective piece, I propose ways in which academic DP scholars could enable truly two-sided communications that better support public data stewards at the Census Bureau and beyond. First, I discuss how academic DP work fundamentally differs from the policy decisions needed to implement DP systems and why this distinction has political consequences. Through examples and discussions from workshops, I show how the DP community largely asked applied stakeholders to communicate on DP's theoretical terms, when such an ask foreclosed important considerations relevant for the Census Bureau's policy problems. Second, I discuss how existing misalignment between DP and SDL researchers unintentionally seeped into the ways we communicated about DP, pointing to why both perspectives are necessary, in different ways, for policy conversations. Finally, I conclude by discussing how these issues are not unique to data privacy work, but instead reflect structural problems in translating theoretical science into practice, prompting suggestions for more productive interdisciplinary engagement.

As an essential disclaimer, what follows is in no way intended to chastise the DP community, but only to propose how we can better collaborate with those concretely impacted by DP systems. Most readers are well aware of the many controversies surrounding the Census Bureau's decision, and we do not need to belabor these well-trodden debates. Still, the DP community is largely comprised of academic researchers (myself included) who benefit from the Census Bureau as a hallmark use case, one which will certainly be a touchstone when evaluating the broader impacts of publications, grants, and other scholarly output. As an early-career academic statistician developing formal data privacy technologies, I benefit from leveraging an instrumentalist view on technology; studying ethics in abstract, mathematical models of data processing allows us to treat generalizable mathematical mechanisms as "neutral" tools whose ethics stem from how they are deployed in practice. Sociological research, in contrast, focuses on how technologies and social values shape one other, helping to invite some missing reflexivity that challenges this neutrality. To move forward, academic DP researchers must acknowledge how this non-neutrality affects our external engagements, so that we remain sensitive to the most pertinent issues raised by new data privacy perspectives.

1. Miscommunications Between DP Theory and Privacy Policy

At the center of the Census debates lie two simple questions: what does DP afford public data stewards, and do these affordances support the substantive policy goals of privacy-aware data publishing? In academic spaces, we naturally gravitate towards answering these questions as generically and quantitatively as possible. Abstraction is the language of computer science, and academic DP papers rely on neat, self-contained mathematical abstractions to encapsulate the most pertinent properties of privacy-enhancing technologies like DP. Outside DP research, however, privacy scholars have identified numerous potential harms from data processing [Solove, 2006]; in response, position pieces advocating for broader DP adoption are careful to delineate which harms can or cannot be addressed by DP (or with formal reasoning at all) [Dwork and Naor, 2010, Dwork et al., 2017, Nissim and Wood,

2018]. Technical precision and transparency regarding DP's affordances and limitations are no doubt important, but it is tempting to view these as a sufficient prerequisite for engaging in policy debates. Unfortunately, this *de facto* mode of communication presumes a seamless hand-off between technical scaffolding and policy evaluation, an assumption thoroughly challenged by the Census use case. In this section, I discuss how friction in this process affected our engagements with the applied communities that leaves much room for improvement.

DP methods do not themselves produce policy evidence, but instead dictate the terms on which evidence is produced and legitimized. DP's definition of privacy risk is inherent to how its mechanisms operate, making choices about how DP gets implemented the locus of negotiation for academic research. Evidence that may be irrelevant for academic DP research, though, could be relevant for policy, precisely because DP abstracts away or brackets some privacy harms such as broad-based surveillance or fine-grained statistical inferences [Seeman and Susser, 2022, Seeman, 2023]. Failing to reintroduce these nuances leaves us vulnerable to mistaking policy arguments for flawed methodological logic. For example, studies on inferring voter race and ethnicity from DP outputs yielded evidence about the kinds of privacy harms that could persist even after DP's application [Kenny et al., 2021]. Such evidence prompted a back-and-forth rejoinder in which the original analysis was dismissed as "privacy irrelevant" from academic DP's perspective [Bun et al., 2021]. In hindsight, this could have been an opportunity to provide additional context on broader privacy harms inherent to many, if not all, possible Census policy alternatives, situating them as part of a more nuanced and comprehensive policy discussion. Instead, our community tended to interpret these arguments as referendums on DP's efficacy, and in response, we shut down contestation on technical grounds alone. Interactions such as these caused proponents and opponents of the new methodology to talk past each other, instead of trying to address *both* the scientific and policy issues at stake.

Similar tensions played out within DP's technical scaffolding, as the 2020 Decennial Census use case exposed places where the Census Bureau's data product needs strained what was possible to achieve on DP's terms alone. The Census Bureau's logistical and operational constraints were prescribed by existing laws and policies, leaving substantial gaps between what DP theoretically offered and what the resulting disclosure avoidance system delivered. On the privacy side, some data publishing requirements necessarily strayed from the way in which DP models privacy loss. Certain "invariant" statistics needed to be released without privacy-preserving noise, and certain quality assurance processes needed access to confidential Census records to ensure compliance with their internal procedures [Abowd et al., 2022]. On the utility side, the opportunity to correct downstream inferences through transparent specification of the release mechanism proved pragmatically difficult; technically precise methods were computationally infeasible at the granularity many researchers needed, and the highly data-dependent errors introduced by DP could not be easily empirically disentangled.

These gaps proved to be a conundrum for academic DP researchers, as they gestured at an inconvenient open secret: DP's assumptions are sociologically fragile, and implementing DP systems in practice strains those assumptions. Should we deviate even slightly from the models posed by DP, its desirable properties no longer necessarily hold [Dwork and Naor, 2010, Dwork et al., 2017]. Our community has taken this as an imperative to make real-world data processing systems as amenable to DP as possible, even when doing so was demonstrably impossible for the Census Bureau to accomplish. These challenges, though, could have been

an opportunity to co-construct formal privacy with Census stakeholders by reflecting on how FP methods could better accommodate real-world deviations from DP's requirements and how to make methodological transparency more actionable by design for end users. Doing so would require grappling with multiple kinds of data governance, some within DP and some at its boundaries where DP's presence would not suffice to solve the policy problems at hand. Instead, our communications with data users and stakeholders often either trivialized these gaps or denied their existence altogether: we frequently communicated about DP's theoretical properties as if they were synonymous with the realized system outcomes, which proved to be counterproductive.

Take, for example, the Rutgers DIMACS (the Center for Discrete Mathematics and Theoretical Computer Science) Workshop on the Analysis of Census Noisy Measurement Files and Differential Privacy [Dwork et al., 2022]. The goal of the workshop was, as one point, to "articulate a technical research agenda for statistical inference on the differentially private Census noisy measurement files, to support social research and policy decisions". At the time of the workshop, though, the namesake files were unavailable, athough the situation is actively being rectified. Still, the way the academic DP community engaged with data users displayed how much distance we have placed between academic work and real-world systems. Many Census data users articulated precisely the kinds of problems they wanted to solve with the Census's data products, but in response we often proposed new DP algorithms presuming modes of access to data that typical users did not possess. In Session 2 (Applied Demography), for example, demographers expressed concerns about working with the new DP data products for small area estimation, and academic DP researchers proposed data publishing algorithms that incorporated small area estimation models into the Census's data publishing methods. Such a strategy is theoretically sound, but it fails to take seriously constraints from working with existing data products.

Similarly, when stakeholders provided feedback about data quality concerns, we critiqued their lack of comparison to other policy alternatives (such as past disclosure avoidance methods) or other sources of errors (such as those from imputing missing responses) without considering whether such comparisons were feasible. In Session 6 (Bridging the Use Cases and the Methodology Research Communities), data users discussed how they traditionally lacked privileged access to information on past disclosure avoidance methods and other sources of errors necessary to model how DP errors compared holistically to others. Moreover, even if these analyses were technically feasible, many data users are contractually bound to use Census data as prescribed for specific policy purposes, most frequently in ways which fail to account for uncertainty in Census estimates. None of this devalues the immediate need to encourage quantitative reasoning about policies with uncertain data, which has clear academic and scientific value. Rather, the problem lies in assuming such reasoning is immediately possible and automatically yield better evidence-based policy-making.

Unfortunately, what happened at Rutgers was not an isolated incident: from CNSTAT workshops to statistics conferences, similar interactions highlighted the same tensions between theory and practice. The reality is that the Census Bureau's use case simply operates on too many different scales and resolutions to provide both reasonable data privacy guarantees and reasonable data quality for every use case [Abowd et al., 2021, Abowd and Hawes, 2023]. Just as it is foolish to pretend that DP will destroy all federal data as we know them, it is equally foolish to assume all data users are merely crying wolf. When we dismiss criticisms of the Census Bureau's methodology because data users cannot express them in terms amenable to DP, we are implicitly signaling that the only stakeholders that deserve a seat at the table

are those capable and willing to do so. Yet many of these concerns were raised precisely because the kinds of governance DP could have enabled was not materializing as intended. So long as we continue to leverage DP's real-world use cases to support theoretical research, at a minimum we should not evangelize the technology by acting as if it encapsulates the relevant governance questions.

2. Breaking Binaries in Stakeholder Communication

Many concerns about DP at the Census Bureau stemmed from the particulars of how the Census Bureau established its system. However, long before the Census Bureau announced its decision back in 2018, data privacy researchers already tended to self-organize into two distinct camps: SDL emerged from applied statistics in official and administrative contexts [Hundepool et al., 2012], and FP emerged from theoretical cryptography in computer science [Dwork et al., 2006, 2014]. SDL researchers generally take a more empirical approach by analyzing particular statistical outputs and absolute measures of disclosure risk (i.e., what can concretely be learned about particular outputs) under specific adversarial assumptions. By contrast, FP research generally takes a more theoretical approach by analyzing release mechanisms acting on entire database schemas with their associated relative risks (i.e., how disclosure risk changes between different data publishing scenarios). What makes data privacy inherently challenging is that negative results, such as the database reconstruction theorem [Dinur and Nissim, 2003] and the "no free lunch" principle [Kifer and Machanavajjhala, 2011], make these conceptual concessions inescapable. There is no magic set of metrics that captures the most general possible privacy and utility guarantees simultaneously. Therefore, all data privacy methodologists must choose which assumptions and metrics they are willing to tolerate [Slavković and Seeman, 2023].

To some degree, all methodological changes for official statistics are contentious because of their consequential relationship to material political outcomes, such as democratic representation [Cohen et al., 2022] and allocations for social services [Brummet et al., 2022]. So what happened with the Census Bureau is historically emblematic of official statistics as a locus of power and knowledge production [Bouk, 2017, 2020]. Despite all this, the change at the Census Bureau from SDL to FP impacted how we communicate about differences between these two, making this one choice yet another battleground in a long-standing methodological dispute between advocates of FP and SDL. What started as a methodological difference evolved into claims of authority over the bounds of legitimate and illegitimate data privacy research, and the Census's decision unintentionally further entrenched this division.

As a result, these dynamics reduced complex, multidimensional topics to false binaries. When we communicate about methods being "private or non-private," or data being "useful or not useful," or methods being "transparent or not transparent,", or "objective or *ad hoc*," we are collapsing sociotechnical complexities in ways that implicitly dismiss all non-DP methods as illegitimate. This has made attempts to bridge the gap between DP and non-DP approaches to data privacy nearly impossible. When we claim a method is "not private," what we typically mean is "this data processing system does not have self-evident or provable privacy guarantees." Yet provable privacy guarantees come at the sociological cost of abstracting away social and technical details that must subsequently be reintroduced to solve policy problems. Binary language dismisses empirical evidence solely because it is irrelevant for DP's academic purposes, when that evidence could be highly relevant for policy

decisions. If the DP community wants more substantive engagement with other privacy scholars and data stakeholders, we ought to be more sensitive to how these tendencies affect external communications, especially when others have legitimate policy interests different from our own.

To make this conversation more concrete, consider one potential harm from Census data: re-identifying trans youth. Under the Census Bureau's former methodology, reconstructed records yield a tangible risk for reidentifying trans youth [Keyes and Flaxman, 2022]; moreover, reidentified records members of this sub-population who are out, making them a target for discrimination and healthcare access disparities [Gordon, 2022]. One policy goal is then to decide how to place upper bounds on disclosure risks. The privacy guarantees provided by DP describe how much the Census Bureau might exacerbate existing reidentification harms by releasing these data, capturing relative differences in disclosure risk based on whether a particular individual contributed to the data. At the same time, others may care more about absolute measures of disclosure risk, for example, how many trans youth could be reidentified with a given accuracy from Census data; this is solely because they want to ensure the realized results do not enable a baseline level of harm for trans youth. Here, the core tension at the heart of DP, the fundamental law of information recovery, comes into play: no method can provably limit the number of trans youth whose records could be reconstructed from Census records [Dinur and Nissim, 2003]. This limitation, however, does not, *a priori*, disqualify empirical evidence about the effectiveness of real attacks, as they are tied to material privacy harms.

The policy conversations here highlight a central tension: how should we hold individual data processors accountable for collective privacy harms, i.e., harms that emerge from multiple data processors publishing and sharing data? DP's abstraction choice relies on relative disclosure risks as the necessary concession to isolate one data processor's contributions to privacy risk. The policy question, though, reorients decision-making around this balance between individual and collective accountability: how much should the Bureau change the plausible deniability of trans youth reidentification, and with what accuracy should certain baseline attacks prevent this reidentification? No one mathematical yardstick solves this accountability problem, and delegitimizing evidence that suggests otherwise shuts down essential contestation and forecloses on deeper questions about the politics inherent in choosing how to measure privacy.

The DP community's relationship to empirical evidence ought to evolve when we leave academic spaces and move these conversations into policy spaces. The distinction between FP and SDL guarantees is acceptable in academic settings when we evaluate technologies for their scientific merits, but can sometimes be counterproductive in policy settings. Moreover, it betrays the intertwined histories of SDL and FP research, particularly for official statistics [McKenna et al., 2018] . When we dismiss empirical methods as "*ad hoc*" or "not meaningfully private," we are dismissing over half a century of disclosure limitation scholarship that uses many of the same fundamental principles and technological interventions. None of this denies the clear and immediate value in generalizing and rigorously theorizing these issues to make such guarantees provable and future-proof, the hallmark contributions of FP research. Instead, because the fundamental interventions of careful query selection and noise injection are not fundamentally new, SDL and FP can easily complement each other, rather than compete against one another in policy conversations. Most SDL researchers are more than happy to use DP methods, and we should not unfairly demonize this community because of a few outspoken and unrepresentative critics. Neither the DP nor the SDL community can

solve privacy policy problems alone and in isolation, and we owe it to data users to embrace more pluralism.

Such an ask can be difficult in practice, however, because academic institutional practices reward more homogeneous viewpoints within research sub-communities [Klein and Stern, 2009]. Moreover, the kinds of pluralism that persist in some research fields run counter to the technical values that incentivize narrower, mathematically productive abstractions, especially those driving the explosion of academic DP research in the past decade. Regardless, other privacy scholarship communities have taken steps to make their fields more pluralistic, and we could learn much from how they are organized. Qualitative privacy scholars in particular tend to take a more expansive approach to research, embracing a more charitable view of alternative frameworks and disciplines [Marwick, 2022], while understanding how contestation is essential to substantive privacy interventions [Mulligan et al., 2016]. Inviting contestation into DP scholarship does not make it "less rigorous," but instead acknowledges a politics that was always there in how the way we frame problems implies particular solutions [Akrich, 1992]. Technical tools that imbue data processing tasks with social values have social utility not in spite of, but *because*, they make explicit which values are being (or not being) articulated. We must then decide if these formal guarantees match our substantive political goals [Green and Viljoen, 2020]. Engaging in this process can make DP scholarship more substantive when applied to policy problems.

3. Toward Accountability Between Theory and Practice

The tensions discussed above are partially the growing pains of DP evolving from a theoretical mathematical framework to one enmeshed with a collection of sociotechnical practices. Similar issues between theory and practice have long been observed, for instance, by philosophers of science [Soler et al., 2014], management scholars [Van de Ven and Johnson, 2006], and political theorists [Kitcher, 2003]. Even within theoretical computer science, we are not the only field exhibiting awkward tensions between theoreticians and practitioners: in sequential decision-making, there is a growing gap about how theoretical work is motivated by application, but the (lack of) obligation for theorists to attend to these motivating applications seems to imply a missing social contract between theoretical and applied scholars [Hullman, 2023]. We see the same dynamics in DP: many major DP papers cite official statistics as "motivating examples," yet only a small proportion of them directly address end-to-end applications in official statistics or other public good settings, especially when it comes to the real-world problems official statistics producers face [Machanavajjhala et al., 2017, Miklau, 2022].

To conclude, I discuss some ways the DP community could improve these dynamics.

First, our community should more attentive to stakeholders, particularly when hand-offs between theory and practice implicate differences in power, labor, liability, or responsibility. DP scholars often justify the soundness of their theory by deferring these messier ethical questions to practitioners; this in turn makes privacy-enhancing technologists "on the ground" akin to operations managers, making essentially political choices in the way they configure access rights and the human arrangements needed to enforce those rights [Bamberger and Mulligan, 2015, Rogaway, 2015]. At the same time, the priorities of academic researchers often can be misaligned with those for practitioners. Academic DP researchers depend on funding sources that historically prioritize innovative foundational research at the expense of translation and implementation research [Vinsel and Russell, 2020]; consequently, there

is little incentive to engage with implementations when they cannot simultaneously help develop theoretical foundations. When we train future DP systems engineers using the dominant ideological principles of today's academic DP literature, we are not preparing them for the realities of building DP systems that leverage essential implementation work. Our community depends on interconnected networks of practitioners' labor to solve the problems we choose to ignore, bolstering the real-world impact of DP; we should more carefully examine how we can improve this social contract, especially when working with data infrastructures intended to serve the public good.

It is, unfortunately, quite easy for us to dismiss these criticisms, claiming that it is the job of others to make DP more interoperable with real-world data processing systems. The infrastructural issues brought to light by DP at Census suggest multiple sweeping changes to Census and social science writ large, namely making law and policy more compatible with probabilistic reasoning [Cooper et al., 2022], and shifting cultural norms surrounding how data users conceive privacy and utility [Nanayakkara and Hullman, 2023]. As part of a long-term strategy to improve DP's viability, these are entirely sensible goals, but there are two important wrinkles that remain relevant for today's academics. First, by presuming this particular division of labor, we take for granted that the "gap" between theory and practice can and should be closed by others to accommodate DP's technical logic. Second, by focusing on novel technological developments, we are neglect current problems where the sociotechnical prerequisites to implement these technologies may not be met. As DP pivots from cryptographic theory to a high-impact real-world technology, we should encourage substantive co-construction between academics and practitioners to avoid these trappings and constitute a more productive social contract between theoreticians and practitioners.

Second, our community needs to more seriously engage with how privacy gets intertwined with other (intangible) ethical values in data processing, both formally and substantively. This proves difficult because we are accustomed to placing additional requirements on stakeholders in service of intangible benefits, including but not limited to privacy. Take for example, measurement error corrections for private inferences: we are asking data users to exert more methodological effort for less precise inferences in the service of reproducible social science, as reflected by the numerous feedback workshops organized by the Census Bureau and CNSTAT. Even though moving towards more reproducible social science ought to be a universal good, we may fail to convince stakeholders of this if we dismiss how achieving this future goal affects today's data use practices and material political outcomes. I argue that the DP community could be more attentive to ways in which the technology unintentionally strengthens or weakens how stakeholders interact with and perform these values, especially when advertising DP as a tool to enable scientific values beyond privacy itself. This would also let us invite needed reflexivity into how we study privacy itself. For example, when we ask privacy practitioners to analyze holistically the privacy guarantees of end-to-end implementations, including understanding how privacy loss budgets were set, we are asking them to exert more methodological effort to capture the privacy loss attributable to actions the data curator actually took. Practices such as these help ensure DP and other privacy-enhancing technologies are wielded with a careful eye towards accountability.

To reiterate, the Census use case can teach our community how to better engage with policymakers in more cooperative, pluralistic means. I present these arguments with hope that the tools privacy researchers develop will ultimately be usable by practitioners. Our commitment to privacy as a social value ought to bind our community together more strongly

than theoretical and methodological disputes divide us, and that commitment starts with making our community more open to the realities of policymaking.

References

J. Abowd, R. Ashmead, R. Cumings-Menon, S. Garfinkel, D. Kifer, P. Leclerc, W. Sexton, A. Simpson, C. Task, and P. Zhuravlev. An uncertainty principle is a price of privacy-preserving microdata. *Advances in Neural Information Processing Systems*, 34:11883–11895, 2021.

J. M. Abowd. The U.S. Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2867–2867, 2018. doi: 10.1145/3219819.3226070.

J. M. Abowd and M. B. Hawes. Confidentiality protection in the 2020 US Census of Population and Housing. *Annual Review of Statistics and Its Application*, 10:119–144, 2023.

J. M. Abowd, R. Ashmead, R. Cumings-Menon, S. Garfinkel, M. Heineck, C. Heiss, R. Johns, D. Kifer, P. Leclerc, A. Machanavajjhala, et al. The 2020 Census Disclosure Avoidance System topdown algorithm. *Harvard Data Science Review*, 2022.

M. Akrich. The de-scription of technical objects. In J. B. Wiebe and T. Pinch, editors, *Shaping Technology / Building Society: Studies in Sociotechnical Change*. MIT press, Cambridge, MA, 1992.

K. A. Bamberger and D. K. Mulligan. *Privacy on the Ground: Driving Corporate Behavior in the United States and Europe*. MIT Press, 2015.

D. Bouk. The history and political economy of personal data over the last two centuries in three acts. *Osiris*, 32(1):85–106, 2017.

D. Bouk. Error, uncertainty, and the shifting ground of Census data. *Harvard Data Science Review*, 2(2):2–9, 2020.

D. Boyd and J. Sarathy. Differential perspectives: Epistemic disconnects surrounding the US Census Bureau's use of differential privacy. *Harvard Data Science Review*, 2022.

Q. Brummet, E. Mulrow, and K. Wolter. The effect of differentially private noise injection on sampling efficiency and funding allocations: Evidence from the 1940 Census. *Harvard Data Science Review*, 2022(Special Issue 2), June 2024 2022. https://hdsr.mitpress.mit.edu/pub/ft9fhmku.

M. Bun, D. Desfontaines, C. Dwork, M. Naor, K. Nissim, A. Roth, A. Smith, T. Steinke, J. Ullman, and S. Vadhan. Statistical inference is not a privacy violation. *Posted June*, 3: 2021, 2021.

A. Cohen, M. Duchin, J. Matthews, and B. Suwal. Private numbers in public policy: Census, differential privacy, and redistricting. *Harvard Data Science Review*, 2022(Special Issue 2), June 2024 2022. https://hdsr.mitpress.mit.edu/pub/954ycugm.

A. F. Cooper, J. Frankle, and C. De Sa. Non-determinism and the lawlessness of machine learning code. In *Proceedings of the 2022 Symposium on Computer Science and Law*, pages 1–8, 2022.

I. Dinur and K. Nissim. Revealing information while preserving privacy. In *Proceedings of the Twenty-second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 202–210, 2003.

C. Dwork and M. Naor. On the difficulties of disclosure prevention in statistical databases or the case for differential privacy. *Journal of Privacy and Confidentiality*, 2(1), 2010.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference*, pages 265–284. Springer, 2006.

C. Dwork, A. Roth, and Others. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3-4):211–407, 2014.

C. Dwork, A. Smith, T. Steinke, and J. Ullman. Exposed! A survey of attacks on private data. *Annual Review of Statistics and Its Application*, 4:61–84, 2017.

C. Dwork, R. Gong, W. Su, and L. Zhang. Workshop on the Analysis of Census Noisy Measurement Files and Differential Privacy, 2022.

C. M. Gordon. Caught in the middle: the care of transgender youth in Texas. *Pediatrics*, 149(6), 2022.

B. Green and S. Viljoen. Algorithmic realism: expanding the boundaries of algorithmic thought. In *Proceedings of the 2020 conference on Fairness, Accountability, and Transparency*, pages 19–31, 2020.

J. Hullman. What is the obligation of the theorist to the "motivating application?", 2023.

A. Hundepool, J. Domingo-Ferrer, L. Franconi, S. Giessing, E. S. Nordholt, K. Spicer, and P.-P. De Wolf. *Statistical Disclosure Control*. John Wiley & Sons, 2012.

C. T. Kenny, S. Kuriwaki, C. McCartan, E. T. Rosenman, T. Simko, and K. Imai. The use of differential privacy for Census data and its impact on redistricting: the case of the 2020 US Census. *Science Advances*, 7(41):eabk3283, 2021.

O. Keyes and A. Flaxman. How Census data put trans children at risk. *Scientific American*, 2022.

D. Kifer and A. Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*, pages 193–204, 2011.

P. Kitcher. *Science, Truth, and Democracy*. Oxford University Press, 2003.

D. B. Klein and C. Stern. Groupthink in academia: Majoritarian departmental politics and the professional pyramid. *The Independent Review*, 13(4):585–600, 2009.

A. Machanavajjhala, X. He, and M. Hay. Differential privacy in the wild: A tutorial on current practices and open challenges. In *Proceedings of the 2017 ACM International Conference on Management of Data*, pages 1727–1730, 2017.

A. Marwick. Privacy without power: what privacy research can learn from surveillance studies. *Surveillance & Society*, 20(4):397–405, 2022.

L. McKenna et al. Disclosure avoidance techniques used for the 1970 through 2010 Decennial Censuses of Population and Housing. Technical report, U. S. Census Bureau, 2018.

G. Miklau. Negotiating privacy/utility trade-offs under differential privacy. *USENIX*, 2022.

D. K. Mulligan, C. Koopman, and N. Doty. Privacy is an essentially contested concept: a multi-dimensional analytic for mapping privacy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083):20160118, 2016.

P. Nanayakkara and J. Hullman. What's driving conflicts around differential privacy for the U.S. Census. *IEEE Security & Privacy*, 21:33–42, sep 2023. ISSN 1558-4046. doi: 10.1109/MSEC.2022.3202793.

K. Nissim and A. Wood. Is privacy privacy? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2128):20170358, 2018.

P. Rogaway. The moral character of cryptographic work. *IACR Cryptol. ePrint Arch.*, 2015: 1162, 2015.

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

J. Seeman and D. Susser. Between privacy and utility: on differential privacy in theory and practice. *Privacy Law Scholar's Conference, under revision at ACM Journal of Responsible Computing*, 2022. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4283836.

A. Slavković and J. Seeman. Statistical data privacy: a song of privacy and utility. *Annual Review of Statistics and Its Application*, 10(1):null, 2023. doi: 10.1146/annurev-statistics-033121-112921. URL https://doi.org/10.1146/annurev-statistics-033121-112921.

L. Soler, S. Zwart, M. Lynch, and V. Israel-Jost. *Science after the Practice Turn in the Philosophy, History, and Social Studies of Science*. Routledge, 2014.

D. J. Solove. A taxonomy of privacy. *University of Pennsylvania Law Review*, pages 477–564, 2006.

A. H. Van de Ven and P. E. Johnson. Knowledge for theory and practice. *Academy of Management Review*, 31(4):802–821, 2006.

L. Vinsel and A. L. Russell. *The Innovation Delusion: How our Obsession with the New has Disrupted the Work that Matters Most*. Currency, 2020.