Between Privacy and Utility: On Differential Privacy in Theory and Practice

JEREMY SEEMAN, University of Michigan, USA
DANIEL SUSSER, Cornell University, USA

Differential privacy (DP) aims to confer data processing systems with inherent privacy guarantees, offering strong protections for personal data. But DP's approach to privacy carries with it certain assumptions about how mathematical abstractions will be translated into real-world systems, which—if left unexamined and unrealized in practice—could function to shield data collectors from liability and criticism, rather than substantively protect data subjects from privacy harms. This article investigates these assumptions and discusses their implications for using DP to govern data-driven systems. In Parts 1 and 2, we introduce DP as, on one hand, a mathematical framework and, on the other hand, a kind of real-world sociotechnical system, using a hypothetical case study to illustrate how the two can diverge. In Parts 3 and 4, we discuss the way DP frames privacy loss, data processing interventions, and data subject participation, arguing it could exacerbate existing problems in privacy regulation. In part 5, we conclude with a discussion of DP's potential interactions with the endogeneity of privacy law, and we propose principles for best governing DP systems. In making such assumptions and their consequences explicit, we hope to help DP succeed at realizing its promise for better substantive privacy protections.

CCS Concepts: • **Security and privacy → Social aspects of security and privacy;**

Additional Key Words and Phrases: Differential privacy, science and technology studies, critical code studies

ACM Reference format:
Jeremy Seeman and Daniel Susser. 2024. Between Privacy and Utility: On Differential Privacy in Theory and Practice. *ACM J. Responsib. Comput.* 1, 1, Article 3 (March 2024), 18 pages.
https://doi.org/10.1145/3626494

1 INTRODUCTION

Formal privacy frameworks, such as **differential privacy (DP)** [21] are increasingly prominent, both as an approach to privacy engineering [23] and in discussions about privacy law [3]. In the academic literature, DP is usually depicted as a sophisticated new tool—an upgrade to anonymization techniques of yesteryear, which proved incapable of protecting personal data against database reconstruction attacks [18, 24, 50]. By injecting statistical noise into datasets, DP renders them robust to such attacks, making it difficult to infer information about the individuals they describe. Thus, as advertised, DP claims to enable more useful data analysis with less risk, making hard choices about how to balance privacy and utility supposedly easier to navigate.

Authors' addresses: J. Seeman, University of Michigan, 426 Thompson St, ISR 4065 Ann Arbor, MI 48104; e-mail: jhseeman@umich.edu; D. Susser, Cornell University, Gates Hall #225, 107 Hoy Rd. Ithaca, NY 14853; e-mail: susser@cornell.edu.

2832-0565/2024/03-ART3 $15.00
https://doi.org/10.1145/3626494

But technology studies scholars have long reminded us that every instrument frames the problem it sets out to solve in a particular way [2]. This is especially true for cryptographic tools, which explicitly configure power relations by allocating trust and access to different parties [51]. In this article, we explore how DP frames privacy questions and who benefits from that framing. In doing so, we do not mean to minimize the value of tools that enable data sharing while protecting against database reconstruction. We argue, however, that thinking about privacy through the lens of DP carries with it certain practical and organizational assumptions about the implementation of DP in actual systems, which—if left unexamined and unrealized in practice—could function to shield data collectors from liability and criticism, rather than substantively protect data subjects from privacy harms. As engineers use DP to architect and build data processing systems, and as policymakers turn to DP as a way of conceptualizing the normative demands of privacy and data protection, we ought to be attentive to the shifts in power and responsibility that come with it.

The rest of Section 1 offers a brief overview of DP, both as a mathematical privacy framework and in terms of implementation in real-world sociotechnical systems, and Section 2 sketches a hypothetical case study illustrating how DP's framing effects create gaps between these two perspectives. In Section 3, we show how DP's abstraction choices privilege certain forms of risk management that can be favorable to the interests of data collectors. First, we show that by focusing attention on data release mechanisms, DP orients our attention to privacy harms in a predominantly forward-looking direction, obscuring potential harms enacted in data collection and database construction. In doing so, DP insulates data collectors from criticism and creates barriers to effective auditing processes necessary for holding data collectors accountable. Second, we argue that the act of setting "**privacy loss budgets**" (**PLBs**, or DP parameters that quantify disclosure risks) flattens social contexts and renders the benefits of data analysis more salient while obfuscating the risks. Finally, we argue that the risks accounted for by PLBs center individual privacy harms at the expense of collective ones.

Section 4 discusses how DP exacerbates well-known problems with informed consent and privacy self-management [57]. While the nature and scope of disclosure risks often depend on specific properties of the relevant database, DP ascribes its disclosure risk properties exclusively to the release mechanism and PLB parameters—not information about the database itself. Moreover, even if (theoretically) this gap was closed, empirical research demonstrates that the average person cannot easily or naturally express their privacy preferences in the language of DP [14, 55]. Given the centrality of individual consent to both DP [66] and existing US privacy regulation, in practice DP might therefore function more as cover for data collectors than as protection for data subjects.

Finally, Section 5 argues that effective governance for DP systems requires grappling with how DP frames these relationships between data subjects, data curators, and data users. Left unaddressed, DP could become another occasion for data curators to engage in a kind of "privacy theater," performing compliance with privacy regulations instead of fulfilling their more substantive goals, mirroring systemic problems with privacy regulation as a whole [72]. Thus, to conclude, we propose DP governance principles aimed at addressing the framing effects of DP, highlighting possible interventions both within the work of implementing DP systems and beyond it.

Two notes before diving in. First, some of the issues we raise in what follows are common to other **statistical disclosure limitation (SDL)** techniques, of which DP is only one variant. We focus on DP, in particular, for several reasons: (1) to illustrate specific issues that arise when mathematical privacy methods are applied in the real world, it is necessary to home in on a particular suite of formal approaches rather than rely on generalities; (2) DP's strong guarantees—the scope and strength of its technical affordances beyond what traditional SDL methods offer—give it rhetorical power when justifying privacy design choices; (3) because DP is technically complex, its theories and methods are rapidly evolving, and it is utilized by a variety of actors (some more

and others less equipped to reason carefully about its uses and limitations), carefully unpacking its assumptions about how mathematical abstractions will be translated into real-world systems can help avoid confusion among stakeholders; and (4) DP's widespread adoption and increasing prominence in academic, industry, and policy-making circles make it especially important for people to engage with critically. The stakes of getting DP right are high. At the same time, it is very much worth asking the kinds of questions outlined here about other "privacy-enhancing technologies (PETs)" and related engineering approaches to privacy.

Second, the issues we describe below do not follow automatically or necessarily from the application of DP in practice. They flow, we argue, from the way DP frames privacy problems and solutions. Which is to say, unless these assumptions are made explicit, so data users and data subjects can contemplate and address them, DP might incline people to understand and approach privacy in certain (sometimes unhelpful) ways. Our aim in articulating these implicit framing devices and their normative assumptions and implications is—again—not to call into question DP's value and contributions to strengthening privacy overall, but rather to render concrete the real-world consequences of DP's mathematical abstractions, and in doing so, to help DP proponents deploy it carefully and effectively.

1.1 DP as a Mathematical Formalism versus DP as a Sociotechnical System

Before going further, it is important to understand that the term DP has come to describe both a set of mathematical techniques (what we will call "DP math") as well as the larger sociotechnical systems in which they are embedded (what we will call "DP systems"). As a mathematical framework, DP defines privacy as a set of mathematical properties of database queries. Which is to say, it offers tools for generating aggregate statistics about information contained in a database, without leaking information about the individuals whose data comprises it. A release mechanism (i.e., any function designed to answer a particular query) will satisfy DP's mathematical requirements if the result is robust to changes in one individual's record. In other words, if two databases differ with respect to one individual's record, then the query responses under DP math are nevertheless similar with high probability—DP obscures the contribution of any single person's data to the overall dataset. That obscurity is the "privacy" in "differential privacy."

For any query, this similarity is typically quantified by a numeric parameter called the "privacy loss." Smaller privacy losses ensure outputs that are closer together—more difficult to distinguish—creating stronger privacy guarantees. To satisfy DP math for a particular privacy loss, DP injects randomized noise into the results of database queries, with smaller privacy losses (i.e., more privacy) requiring more noise. Importantly, each query (and subsequent privacy loss) degrades the system's overall privacy guarantees. To keep track of these losses over a series of queries, DP composes them into a cumulative, global privacy loss budget ("PLB") or total allowable losses under a collection of multiple queries. The PLB dictates the privacy guarantees of multiple queries at once, each query consuming some finite amount of privacy loss.

Some important theoretical findings in computer science and statistics motivated this approach. Before DP was introduced, many methods for designing mathematical data privacy protections took individual database outputs as their unit of analysis [33]. As a result, privacy-enhancing "data sanitization" methods were directly tailored to each specific database of interest (e.g., removing personally identifying information or pseudo-identifying information). Despite their utility, these methods still enabled adversaries to extract personal information from so-called "sanitized" outputs. Record linkage attacks were able to successfully de-anonymize individual data subject records, particularly for high-dimensional and sparse datasets [47]. Moreover, the publication of large-scale statistics enabled adversaries to reconstruct hypothetical individual records as inputs to these record linkage attacks [18, 24]. These "database reconstruction attacks" motivated the

original conception of DP. As access to the data and computing power needed to execute reconstruction attacks has grown, such attacks have become more realistically achievable, rendering even summary statistics about databases a threat to individual privacy. For example, although the U.S. Census does not release any record-level data, Bureau statisticians were able to reconstruct individual level records from published tabular summaries, and they were able to link those records to commercially purchased data [30, 34].

In response to these challenges, DP (in its original formulation) introduced two important conceptual shifts. First, DP requires adding noise to every released result. This differs from previous disclosure control methodologies, which sometimes used privacy-preserving randomized noise but did not require it. Second, because any statistic (noisy or not) can enable reconstruction, DP focuses on minimizing relative disclosure risks rather than absolute risks. Which is to say, DP quantifies how disclosure risks change relative to what an adversary might know before seeing the query answer. As we discuss in what follows, DP is thus best understood as a "harm reduction" project, rather than as a tool for definitively anonymizing/de-identifying datasets—it makes disclosure less likely, not impossible.

These innovations give DP desirable mathematical properties not shared by previous approaches to quantitative data privacy: e.g., they enable methodological transparency (disclosing the way in which noise is injected into the statistics does not itself change the PLB or any resulting privacy guarantees), release mechanisms are robust to post-processing (one can use DP-generated statistics as inputs to further statistical analysis while retaining the privacy guarantees of the original release), and DP outputs compose (if two different releases satisfy DP math for two different PLBs, then releasing both results simultaneously satisfies DP-math for some new, typically larger PLB). For these reasons, DP techniques have been incorporated into a wide variety of data-driven tools—database systems engineered to respond to queries using one of these formalisms (what we refer to as "DP systems") [17].

Important for our purposes, DP math abstracts away from numerous concrete design choices required when implementing DP in real-world systems. Or, to put the same point the other way around: Engineering DP systems requires answering questions that DP's mathematical formalisms ignore. First, data curators (i.e., designers and operators of DP systems) must determine who their intended data users are and what requests these users could make to the system. Second, data curators must establish which database queries they will (and will not) respond to. Third, data curators must allocate components of their PLB to these different queries. Although all three of these tasks are essential to establishing a DP system, DP math typically presumes these questions are answered in advance. Yet, in reality, they are rarely established in clean-cut terms, which, as we will see, creates a number of underappreciated problems.

A central aim of what follows is to highlight the tensions created when attempting to engineer DP systems, exploring what happens when neat mathematical formalisms make contact with real-world choices, contexts, and institutions. One benefit of mathematical framings is that they offer closure—we can prove or disprove whether the mathematical properties of a release mechanism satisfy a particular privacy definition under a set of established assumptions. No such closure exists, however, when analyzing DP systems, since we cannot definitively establish whether the assumptions embedded in the mathematical abstractions map onto any system in use. Real-world applications are unavoidably messy. Thus, there may be tensions or misalignments between DP's theorists and its systems engineers. Our goal in this article is not to resolve these tensions but to investigate their normative implications. Since our interest is in the sociotechnical effects of DP systems, we focus less on the etymology or axiomatic interpretations of any one privacy definition or formalism and more on developing a bird's-eye view of the properties most common across the most publicly visible DP systems [15].

Finally, to stave off potential confusion, a brief comment about terminology: Privacy is a highly—perhaps "essentially"—contested term [46]. Debates over its precise meaning and value have raged for decades and continue in contemporary technical, philosophical, legal, and policy arenas. In this article, we mostly use the term "privacy" in the way technical scholars studying DP and other formal privacy methods do: as a property of data processing outputs that enables protection against unwanted disclosure of personal information, and not other security vulnerabilities [42]. We realize that this is a narrow understanding of a rich and multifaceted concept. One goal of ours is surfacing how, to make privacy amenable to mathematical precision, DP abstracts away from its social and normative complexity. Viewing DP math as a uniform improvement over prior techniques might lead one to believe these normative issues have been resolved, when in fact they have merely been bracketed in a way that has not attracted sufficient scrutiny.

2 "MINDING THE GAP" BETWEEN DP MATH AND DP SYSTEMS

To start, imagine the following:

> A large hospital system is studying an emerging genetic disease in a small population of patients (henceforth "data subjects"). Researchers there ("data users") partner with a healthcare data technology company (the "curators") to analyze patient data using DP. After negotiating about query selection and privacy loss budgets, the DP-system parameters are established. Patients are advised that the system will allow researchers to analyze data about them without revealing sensitive information about their disease status. The patients consent, and the researchers analyze and publish their findings.
>
> Then, things start to go wrong. A few months after the system is introduced, two problems emerge: First, a life insurance company (or "adversary"), combines the published results with consumer spending data to identify people with the rare genetic disease. Given their higher financial risks, the insurance company raises the data subjects' rates. The revelation creates a public relations fiasco with participants and observers divided about who is at fault.
>
> Data subjects protest that they were discriminated against based on their disease status, which they were promised would be kept private. Moreover, they argue, had they better understood the risks they would not have consented to participating in the study. On the other side, the data curator argues that the DP system worked as intended, and the inferences the adversary made would have been similar with or without incorporating the researchers' published statistics. The curator points out that the adversary's actions were based on predictions of people's insurance risk, not their risk as attributable to having the rare disease, and thus, people without the disease might have been equally affected by the decision to use the data in their risk model.
>
> The second problem is the researchers' findings turn out to be less statistically sound than they had hoped: A nationally sponsored clinical trial fails to reproduce the DP system's results, undercutting their preliminary work. This second controversy creates a different set of debates—this time inside the scientific community—about the utility of DP-protected statistics, i.e., about how to ensure such statistics are scientifically valid.
>
> Proponents of DP argue that because the release mechanism can be transparently disclosed, it is the responsibility of data users to adjust their downstream inferences to account for the errors introduced to preserve privacy. If the resulting inferences are too weak to be useful, then the solution is to increase the PLB (i.e., make it easier

to learn about individual patients) or to design a more optimal release mechanism. Privacy loss is justified, they argue, by the more robust science it enables.

For the researchers, it is a scandal that the PLB was not set to allow for appropriate statistical power in the first place. They argue that having only released parameter estimates from a few models, it would be computationally implausible (though possible) to use that information to reveal anything sensitive about individual patients. Meanwhile, DP detractors question whether scientists alone (who are not exactly disinterested parties) are entitled to make these kinds of decisions, weighing the need for statistically reliable scientific evidence against privacy and other ethical concerns.

In this hypothetical scenario, the choice to use DP was motivated by two goals: (1) limiting disclosure risks for research participants, while (2) preserving enough of the data's informational content to allow researchers to produce useful (i.e., statistically valid, reproducible) inferences. Despite good intentions, neither goal was met. But that in itself is not particularly noteworthy, as any SDL technique could fail in these ways. What is notable is how DP sets up these debates—the particular way DP frames contestation and negotiation over social values. In the remainder of this article, we unpack and explain this, paying special attention to considerations about privacy, utility, and related values.

First, we argue that DP math centers PLBs as the primary site where privacy is negotiated. Through this lens, it can appear as though both the privacy guarantees for data subjects and the statistical utility of the researchers' results flow entirely from this choice. But PLBs are only one element of a larger, more complicated story. Many decisions leading up to the setting of a PLB—including, importantly, the decision to use DP at all—determine how risk and utility are balanced. Specifically, DP's abstractions can disguise the curator's role in creating disclosure risks ("the responsibility problem", Section 3.1); the implications of choosing a particular privacy budget are often more visible to data users than they are to data subjects ("the allocation problem", Section 3.2); and DP's framing emphasizes individual harms, while the harms to data subjects (like the patients in the hypothetical) can also be, in an important sense, collective ("the network problem", Section 3.3).

Second, DP creates unique obstacles to "privacy self-management," the regulatory approach that makes each individual responsible for evaluating and negotiating the terms of data collection about them. Despite sustained and well-known criticisms, privacy self-management is still the dominant paradigm in privacy law and policy, at least in the US, requiring data collectors to inform people about their data practices and then leaving it up to individuals whether to consent to them [61]. Where DP systems are used, data subjects are forced to reason about the risks and benefits of data collection in DP's terms. Because these terms are often unintuitive, expressing subjective privacy preferences in the language of PLBs can be difficult ("the mathematical language barrier, Section 4.1). For technically sophisticated data subjects, knowledgeable enough to articulate their preferences in terms of PLBs, making an informed decision about whether to disclose information still requires that curators provide sufficient detail about the specifics of the relevant DP system, adding to the ever-expanding set of legal, practical, and technical details one must internalize and deliberate about to meaningfully control their personal data ("extending the consent dilemma, Section 4.2).

Finally, we explore the implications of these problems for privacy governance. The data curator in our hypothetical relied on DP to balance privacy and data utility, and the system's failures reveal important obstacles to accountability for these decisions. Formal privacy guarantees can provide moral cover for data curators. And when data utility problems emerge, the obvious solution for data users—increasing PLBs—implies further privacy losses for data subjects. In this way, DP can

create the appearance of private data processing without providing substantive protections ("the endogeneity problem, Section 5.1). To overcome such obstacles, data subjects should be better represented within DP negotiations ("governance within DP, Section 5.2), and data curators and users must be clear about the fact that DP protects against some—but not all—of the privacy risks data subjects face ("governance beyond DP, Section 5.3).

To be clear: DP is a major advance, addressing many shortcomings of other disclosure control methods. Our aim is not to minimize these technical achievements. Rather, we want to surface the frames, expectations, and values DP embeds in sociotechnical systems when its methods are implemented in practice.

3 PRIVACY THROUGH DP'S LENS

In this section, we describe how DP's theoretical approach to privacy protection structures thinking and decision-making about privacy in practice—the relationships it creates between parties, the choices it presents to them, the tradeoffs it requires, and the salience it confers on different interests and values. How framing an issue or problem impacts people's perceptions, evaluations, and reasoning about it has been studied extensively in psychology, economics, communication, and political science (e.g., References [12, 16, 49, 67]). In what follows, we use the terms "framing" and "framing effects" in the way that **science and technology studies (STS)** scholars frequently do—to indicate how the use of a particular tool (in this case, DP) to address a particular problem can implicitly carry with it value-laden assumptions about how the problem should be solved (see, e.g., References [2, 39, 54]). Some of these framing effects are unique to DP; some are true of many SDL techniques. But crucially, adopting DP does not *commit* one to navigating privacy challenges in the ways we describe—indeed, our purpose in drawing attention to these frames is to help practitioners consider them more carefully and thus to avoid their blind spots.

3.1 The Responsibility Problem: On Forward-looking Harms

The first thing to notice about the way DP frames privacy decisions is that it focuses attention in one direction: toward the future. In its most common form, DP assumes that data has already been collected and stored or that data curators have pre-specified exactly which data will be collected and stored. Then, the decision has been made to further release it in some form, perhaps publicly. The primary question is how to do all of this without exposing data subjects to too much risk of harm from those disclosures.

That is not a bad question. And the tools DP offers for addressing it represent a true advance in privacy-enhancing technology. But we should also note the questions this future-orientation draws our attention away from. For example: Should this information have been collected in the first place? Was it collected in the right way? Should it be stored in this database, subject to these parties' control? Who decides who gets access to the data and on what terms? Whose interests does the existence of this database serve? Focusing entirely on forward-looking harms can deflect attention away from past and present harms—such as illegitimate data collection, concentrating data in unacceptable ways, and so on—and it can displace responsibility for those harms from data collectors and curators onto data subjects and users. We refer to this as the "responsibility problem."[1]

[1]Importantly, a range of DP variants have been developed to account for a variety of trust models. The conventional "central" model of DP assumes a trustworthy data aggregator, but others—such as local DP [25, 37] and distributed DP via shuffling [11]—do not. The choice between these various DP implementations invites system designers to consider some of the issues we are raising. But the future-orientation of privacy harm remains regardless of where in the data collection process distrust is modeled. In most cases, data processors still make DP implementation decisions such as choosing queries, mechanisms, and PLBs, regardless of which trust model is applied; as a result, using these alternative trust models may not substantively affect data subject privacy experiences. See, for example, the local DP cases in Reference [15].

DP's future-orientation is not accidental or contingent; it is encoded in its definition of privacy. By conceptualizing and quantifying privacy risks not in absolute terms ("how much risk will data subjects face after this release?"), but relative ones ("how much *more* risk does this release create for data subjects, relative to what adversaries knew about them beforehand?"), DP orients privacy analysis in an additive, forward-looking direction. More, it focuses attention on the risks of releasing particular statistics in isolation, instead of on the bigger picture described above.

There are good reasons for framing privacy this way—it turns intractable technical problems into tractable ones, allowing privacy engineers to make strong, mathematically provable privacy guarantees. But it also creates important limitations. At its core, DP is a harm-reduction project: Since any statistic can lead to potential disclosures by database reconstruction theorems, all we can do (having already decided to release some statistics) is minimize the additional risks they pose. On one hand, this technically motivated concession enables DP to quantify disclosure risks for "arbitrarily" powerful adversaries, i.e., any possible starting point for knowledge about individuals in a database. On the other hand, this theoretical property requires that no information about said individuals is used in implementing the system, such as for choosing queries or setting PLBs. Yet, in practice, many organizations do just that—releasing results without DP protection and then subsequently developing DP-protective systems based on information about the same data subjects.

For example, if the hospital researchers in our hypothetical had previously published papers describing the patient attributes that made them more likely to have the rare genetic disease, then the insurance company could use that information in attempting to detect which insurees posed higher financial risk. One could view this first task (publishing statistical results about the sample) as a violation of DP if one viewed the statistical results published in the papers as the result of a non-DP query answered without additional noise. However, that raises the question: When does modeling data publishing as a DP query answering system begin? Unless an organization's data has been hermetically sealed from the start, the query system model leveraged by DP does not comprehensively describe the data's end-to-end history. DP's mathematical properties only have non-trivial meaning when considering the subset of published results with DP noise, encouraging limiting the data processor actions under consideration to those whose risks can be mathematically formalized. Curators could thus absolve themselves of responsibility for downstream effects of statistics they release that inform potential adversaries. By considering harms in a forward-looking direction, past actions are obfuscated and thereby insulated from criticism or protest.

By invoking "responsibility," our goal is to show how this future-orientation fails to capture certain conditions that make privacy threats realizable in the present, particularly when data curators themselves help to create those conditions. For example, when curators and data users operate within the same organization, they can collaborate to make design choices favorable to data users—and potentially unfavorable to data subjects—based on confidential data (e.g., prioritizing accuracy over privacy in the setting of PLBs) and then release the sanitized statistics to data users after ensuring they meet their needs. That is, they can set the balance of privacy vs. utility in a way that prioritizes the needs of data users (utility) over those of data subjects (privacy). And such conflicts can be shielded from public view by way of trade secrecy protections or other intellectual property claims.[2]

3.2 The Allocation Problem: On Setting PLBs

Next, we discuss how DP frames implementation decisions for data curators and users in a way that privileges certain means for setting privacy loss budgets. Setting PLBs requires significant

[2]Some argue that the voicing of public policy concerns deserves exemption from trade secrecy rules [44].

work translating mathematical formalisms into practical privacy commitments. As we explain, the way DP-math frames privacy loss through PLBs can render real disclosure risks abstract and difficult to interpret. By contrast, the effects of PLB settings on data utility are more concrete and easily perceived. This can implicitly privilege data utility as the driving force behind how PLBs are determined. We refer to this as "the allocation problem."

Privacy loss budgets are one of DP's most celebrated innovations: They allow for a level of mathematical precision when balancing privacy against other values that normally eludes policy discussions. At the same time, these neat formalisms do not convey everything one needs to know when interacting with DP systems. A complete mathematical description of a release mechanism and its associated PLB describes how statistical noise is added to the data. But this description is not specific to the actual database in question—it applies to the entire database schema (of which any particular, realized database is but one instance). Some actual databases in that schema may be more susceptible to disclosure than others, depending on the relative uniqueness of sensitive attributes among data subjects. Setting the PLB thus implicitly requires reasoning about an unknown worst-case database. If, for example, the data curators in our hypothetical wanted to take a privacy-first approach, then they would need to consider the degree of homogeneity or heterogeneity within the patient group. Centering the PLB risks ignoring this.

Furthermore, the relationship between the PLB's relative disclosure risks and measures of absolute disclosure risks depends on numerous database properties abstracted away from DP-math. PLB-based guarantees are typically agnostic to database size and schema complexity, but knowing these might change how one chooses to set the PLB. For example, one might think differently about the privacy risks associated with learning how many people in a room have the disease described in our hypothetical, versus learning how many people in the country are afflicted by it. Practically speaking, one can learn more about individuals in the former case than the latter case. But PLB analysis treats these queries the same way, despite the substantially different empirical disclosure risks they pose.

Similar problems emerge when considering which record attributes are the most disclosive and who in a database is most at risk. Data subjects may have different conceptions of "worst-case" risks, bound up in their personal, subjective experiences of informational boundary management [60], but different individuals in a database also bear different risks, simply based on their database contributions. In the hypothetical, suppose that younger patients are more likely to have the disease than older patients. If that were true, then older patients in the database would have different disclosure risks than younger ones, for whom there are more similar records. Although the PLB accounts for the worst-case scenario among these possible databases, it flattens the privacy decision space, both qualitatively and quantitatively.

At the same time, while focusing on PLBs obscures some of these more nuanced privacy worries, it makes the upshots of data collection and analysis—data's utility—more salient. As we have seen, DP allows for transparent disclosure of the release mechanism, meaning the exact process used to sanitize any statistic can be made public. This is what allows data users to account for the privacy-preserving statistical noise DP introduces when they make inferences from query results. Because the magnitude of privacy-preserving errors is a direct consequence of setting the PLB, a curator can observe an exact change in data utility associated with different PLBs. At one level, this is good: Earlier privacy-preserving data sanitization methods did not have this transparency property, and it enables more accurate, reproducible inferences. But the salience of utility relative to risk could also have unintended consequences.

We can view database-specific privacy and utility analyses as tools for setting upper and lower bounds on PLBs, respectively. DP's framing makes setting lower bounds easier than setting upper bounds, privileging negotiations about DP-systems led by utility concerns. For example, in many

high-profile DP-systems use cases, PLBs are set using "fitness-for-use" modeling in which a set of key data utility goals are defined in advance, with the end goal of finding the smallest PLB that satisfies these data utility purposes [74]. The U.S. Census Bureau chose its PLBs this way when implementing DP [1], and other organizations will likely follow suit. Because fitness-for-use approaches first determine which tasks are important enough to lower bound the PLB and then set the lower bound on that basis, privacy plays second-fiddle to data utility. And when data utility is deemed insufficient for certain tasks, the default response is simply less privacy.

3.3 The Network Problem: On Bracketing Networked Harms

As we have seen, DP aims to balance two competing goals: privacy and utility. Much of our discussion to this point has focused on unpacking the "privacy" side of that equation, but it is worth examining the "utility" side, too. Privacy, according to DP, is protection against individual disclosure—DP aims to minimize the risk that someone could make reliable inferences about any individual contributor to a dataset. At the same time, DP aims to facilitate data's utility, which it defines as enabling statistical inferences about populations. As Michael Kearns and Aaron Roth write, "At its core, differential privacy is meant to protect the secrets held in individual data records while allowing the computation of aggregate statistics" [38].

This solves a real problem: Researchers—especially in health fields, but also computational social scientists, digital humanities scholars, and others—are eager to use data analysis techniques to learn from all of the data generated about us. DP promises to unlock these insights—this "utility"—latent in aggregate, population-level data, while simultaneously protecting the individuals whose data is being mined.[3] By conceptualizing privacy and utility in this way, though, DP places outside its scope protection against the risks of aggregate statistics [10]. We refer to this as "the network problem."

Privacy theorists have long argued that our privacy interests are deeply intertwined—one person can suffer from another person's disclosures. Different dimensions of this problem have been variously theorized through the lens of "networked privacy" [43], "group privacy" [65], "relational privacy" [40, 69], and "privacy dependencies" [5]. As Solon Barocas and Karen Levy argue, one person can become implicated by another person's data—i.e., inferences about the first can be made based upon data about the second—for a variety of reasons: if they have social ties, if they are alike in salient respects, or if they are different in ways that make one stand out in relief. For example, American Express famously reduced a card member's line of credit because his shopping patterns mirrored those of people who failed to pay their bills on time. Thus, while aggregate statistics protected by DP might make individuals harder to identify, they do not—as Barocas and Helen Nissenbaum put it—make them any less difficult to "reach" [6].

DP's architects are sensitive to this problem. Cynthia Dwork, for instance, makes plain that "the things that statistical databases are designed to teach can, sometimes indirectly, cause damage to an individual, even if this individual is not in the database" [19]. In response, DP proponents argue—reasonably—that such harms are simply not what DP is designed to protect against [22]. We should recognize, however, that this kind of harm—incurred not from disclosing information about oneself, but rather by bearing some relation to others who have—is a central feature of digital societies [68]. From algorithmic social sorting [28, 29] to manipulative targeting advertising [63] to unfair risk assessments [62], DP leaves us unprotected from precisely the harms data ethics and policy are most concerned to prevent. It is tempting to divorce privacy concerns from concerns about ethical data use, but as Barocas and Nissenbaum argue, the two are inextricably linked [6]. Worse yet,

[3] Judicious strategies for allocating privacy loss could also ensure formal measures of fairness in these downstream analyses [52].

if—as we have argued—DP's other framing effects can serve to sanction more rather than less data collection and processing, then there is reason to worry that DP could be used, rhetorically, in a way that increases rather than decreases the prevalence and scope of these harms.

To prevent that from happening, we must be just as nuanced about data's "utility" as DP is about the kind of privacy protection it offers. When contemplating whether to allow the production of "useful" aggregate statistics, we should ask: useful for whom, and useful toward what ends? In situations like our hypothetical, where DP is used to enable population-level inferences that inform medical research and advance scientific understanding, there is a case to be made that society at large stands to benefit. By contrast, when private firms such as Google and Facebook use DP to justify collecting more data about us to facilitate more precisely targeted ads, the "utility" at issue does not have the same normative force. Fortunately, data governance scholars are developing new frameworks for navigating these individual and collective costs and benefits, and distinguishing not just between individual and population statistics, but drawing a line between acceptable and unacceptable population-level inferences too [70, 71]. Understanding how to productively utilize DP in that process requires clarifying its promises—both in terms of the harms it aims to protect against and the insights it aims to unlock.

4 DATA SUBJECT PARTICIPATION

In the previous section, we saw how DP's mathematical abstractions can shift attention away from privacy risks and towards data utility. While these considerations are important for curators and data users, data subjects can also participate in privacy negotiations as they engage with data collectors. DP math frames their choices too, as DP models a specific hypothetical: whether an individual contributes their personal information to a database. In this section, we unpack two key assumptions DP makes about data subjects: (1) that data subjects can express their privacy preferences in the language of PLBs and (2) that the transparency DP provides meaningfully informs the decisions data subjects make about whether to contribute data.

4.1 The Mathematical Language Barrier: On Expressing Privacy Values through PLBs

In practice, as we have seen, DP systems are generally engineered with PLBs established by data curators and users before data subjects ever engage with the data collection process. In such cases, individuals express their privacy preferences by deciding whether the system's PLB exceeds their personal risk tolerance, and (consequently) whether or not to contribute their data to it. But it is an open question whether data subjects are equipped to express their personal privacy preferences in this way. We refer to this as "the mathematical language barrier."

Empirical user research suggests that doing so is difficult for many data subjects, as they struggle with understanding trust models and privacy loss budgets [14, 55, 75]. There are deeper problems, though. Fundamentally, individual data subjects cannot model their personal informational boundaries in a sociological vacuum. As we have seen, privacy risks are networked—they exist relative to the uniqueness of an individual's data contribution—but PLBs only capture a worst-case risk, regardless of how close any one individual's contribution is to this upper bound. Data subjects with relatively unique sociodemographic characteristics are typically easier to reconstruct, as has been consistently shown mathematically by techniques in the SDL literature [33]. For individual data subjects to reason about their risks of disclosure they need to understand how they relate to others contained in the database—information data subjects often cannot access.

Furthermore, when data subjects are asked to express their privacy expectations through the language of DP, they have to imagine their individual tolerances for worst-case harms. But not all data subjects are equally equipped to make educated guesses about that threshold. Imagine that certain patients in our hypothetical opted out of contributing their data to the study, because they

knew about the risk of insurance companies using it to set prices. If that were the case, then the privacy harms in the hypothetical were distributed partially based on who had that knowledge and the ability to act on it, not necessarily whether PLBs captured individuals' thresholds for privacy loss. Thus, assessing privacy risk requires the difficult work of conceptualizing networked risks and worst-case adversaries, not equally accessible to, or expressible by, all database participants. And DP makes individual data subjects responsible for it.

4.2 Extending the Consent Dilemma: On Informed Consent and the Perils of Transparency

The dynamics described above exacerbate a more general problem in privacy law and policy, known as the "consent dilemma." US privacy regulation centers on what legal scholar Daniel Solove calls "privacy self-management" [57]. On this model, there are few universally impermissible data practices, ruled out in advance. Rather, data collectors are required to inform data subjects about the information they want to capture and the uses to which they intend to put it, and to ask for their consent. Individual data subjects are responsible for deliberating about the privacy terms being offered and deciding whether to accept them, vetting potential relationships with every data collector they encounter [61]. The "consent dilemma" arises because providing data subjects with only *some* details about how their personal information will be collected and analyzed does not adequately prepare them to make such decisions, but providing *all* the relevant details is overwhelming, and thus, unhelpful, too [57]. Adding DP to the mix makes these decisions even more complicated.

In theory, DP math's assumptions about arbitrarily powerful adversaries could make discussing privacy risks easier: Because results under DP math are robust to post-processing, we might hope that the problem of anticipating harmful use cases is, to some degree, abstracted away. Indeed, robustness to post-processing is a hallmark feature of DP and an essential reason for its success. Yet, for the reasons discussed above, simply knowing a mechanism and its associated PLB tells us little about the *actual* risks of releasing statistics in a particular context. By construction, DP math creates a gap between the information needed to understand DP's theoretical guarantees and its guarantees in the context of a real-world, realized database.

One might ask how this is different from other statistical disclosure limitation procedures, many of which require methodological secrecy. Surely revealing what other approaches keep secret is a net positive? While transparency is often thought of as an unalloyed good in discussions about ethics and governance, an emerging critical literature has demonstrated its potential perils [13, 41]. Merely *performing* transparency ("transparency theater") is not enough; to facilitate real accountability, powerful actors and organizations must provide meaningful and actionable insights about their inner-workings. Thus, DP's transparency can be a tool for curators to establish unearned trust and garner soft institutional power. As Claire Birchall writes, transparency confers "cultural, political, and moral authenticity...an identity as much as a mechanism" [13].

Consider the data curator in our hypothetical, who had every incentive to appear committed to the goals of both data subjects and data users. By telling patients how their data would be processed and describing the protections afforded by DP, the data curator aimed to assuage patient fears about the risks of disclosure. Likewise, by telling data users exactly how the DP system transformed patient information, the curator provided the information necessary to enable the best possible inferences from the DP system's outputs. When the curator's promises to both parties fell short, their transparency served as a kind of shield, displacing the DP system's risks onto data subjects and data users. ("We were transparent about the risks! You should have incorporated them into your own decision-making!")

As we have seen, knowledge of the DP release mechanism and PLB alone omits crucial details about how queries were selected and how the PLB was chosen. When the interests of curators and

data users align, but are orthogonal to the interests of data subjects (e.g., when a for-profit company collects data for internal use but makes public promises about privacy through DP), DP's transparency can be used to maintain the appearance of neutrality, regardless of whether the PLB was chosen in a way that substantially protects individuals' data. This problem is not hypothetical: Consider that while Apple proudly advertises its commitment to user privacy, technical DP researchers have criticized the company's use of DP, pointing to exceedingly large privacy budgets and insufficient meaningful, actionable transparency with data subjects and users [64].

Alternatively, when the interests of curators and data users diverge, transparency can take on a different character. In the case of the U.S. Census Bureau's implementation of DP, some data users—such as social scientists and other researchers—have demanded more granular and accurate data, while the Bureau (the curator) is legally obligated to preserve data subject privacy in a way that bars them from providing it. The Bureau argues that DP math's methodological transparency provides data users all they need to adjust their calculations to account for the errors it introduces. But many data users (especially demographers working with fine-grained data) are concerned that this transparency is not sufficient for assessing the effects of errors on downstream inferences [48]. In this case, then, rather than create cover, transparency is the terrain of debate.

Of course, transparency *can* be helpful. It is useful for understanding and reproducing data analyses, and it helps DP resolve flaws in traditional statistical disclosure limitation methods, which required a degree of secrecy about their methodologies. But, we should be attentive to transparency's other effects as well, especially the way it can mediate and modulate discussions about PLB allocation and responsibility for privacy harms and data utility. By noticing these dynamics, we can better appreciate the consequences of choices often made out of view.

5 IMPLICATIONS FOR DP GOVERNANCE

5.1 "The Endogeneity Problem": On the Stakes of DP governance

DP frames data privacy as a harm-reduction project: Because all data processing operations create some risks, DP quantifies changes in those risks that are directly attributable to the data curator's actions. In purely technical terms, this is a consequence of the database reconstruction theorem. Yet, with any governance problem based on harm, solving the underlying political issue requires determining appropriate contexts, conditions, and thresholds for comparing policy alternatives, i.e., determining when a privacy harm requires legal intervention [73]. In this section, we describe how DP could entrench the power of technical decision-makers in implementing DP systems, despite the many apparent conflicts of interest we have seen so far. This bears a close connection to "legal endogeneity" in privacy law, a term that describes how managerial approaches to privacy regulation can engender bureaucratic box-checking over substantive compliance at the expense of data subject protections [72]. To that end, we refer to this as the "endogeneity problem."

Privacy regulation involves a multitude of actors, ranging from government agencies and lawmakers to privacy engineers. While privacy law is often formulated from the perspective of regulators, technical actors inside companies are integral to realizing its promises "on the ground"—in some cases, by using privacy-enhancing technologies like DP [4]. Should DP become a more salient tool for regulators, the role and power of technologists will likely increase, given the specialized technical knowledge required to make DP work. On its own, this is not a bad thing: Technical experts familiar with data privacy and security could help make concrete some of the difficult, abstract tradeoffs discussed throughout this article. Our concern is about the broader structures within which such decisions are made.

For many data processors, particularly those in for-profit technology firms, compliance tasks are overseen by ethics "owners" whose efforts are shaped and constrained by the nature of their

roles as mediators between the firm's "internal" goals and "external" pressure to comply with the law and social norms around privacy and related values. Empirical research has shown (perhaps unsurprisingly) that ethical interventions in industry settings are typically done in a way that avoids altering core business functions. And when change is necessary, there is a tendency toward technological "solutionism"—the assumption that there is a technological fix for any technological problem [45].

The way DP frames privacy, discussed throughout this article, offers ethics owners an attractive technical solution to privacy problems, requiring little change on the part of firms. By focusing on a narrow definition of privacy that centers disclosure risks attributable to data processing, DP brackets privacy harms associated with data collection. By making privacy harms abstract and data utility concrete, decision-making is oriented toward the latter. While true that any privacy-enhancing technology could be implemented in a purely utility-driven manner, DP encourages it, because privacy loss budgets are easier to lower-bound than upper-bound. In this way, DP offers a win-win for everyone—except data subjects—by giving data curators the veneer of ethical compliance, without seriously disrupting how organizations engage with personal information. While DP scholars have acknowledged the risks of privacy theater from setting too-high PLBs [20], emerging evidence supports a broader set of concerns: "on-the-ground" research from interviews with DP implementers indicates that DP does little to disrupt day-to-day operations for many data processors [59] or privacy expectations for data subjects [55]. Thus, without care and attention to how DP is governed, its impacts may be less substantial than its proponents imagine.

In what remains, we outline two forms such care and attention might take. The first works within DP (and related formal privacy approaches), focusing on how best to implement these technologies. The second explores DP's limits.

5.2 Governance within DP

Where DP successfully captures the risks of data processing, governance ought to contend with how to make tradeoffs between privacy and utility more responsive to the rights of data subjects. This necessarily entails leveraging new DP technical developments and alternative trust models that admit more nuanced expressions of privacy risk, including (but not limited to) risks that vary across data subjects [35] or concern certain aggregate attributes [36, 76]. To that end, we propose two directions for improving DP systems, with an eye toward substantive privacy guarantees.

5.2.1 Concretizing the Harms Attributable to DP. Privacy scholars are used to hollow conceptions of privacy harms. In the words of Ann Bartow, such conceptions often suffer from "too much doctrine, and not enough dead bodies" [7]. This turns out to be a systemic problem in privacy scholarship, in that many often divorce theories of privacy harms from the affective dimensions of lived privacy experiences [58]. DP's conception of privacy loss operates similarly and in doing so makes DP a more powerful tool for shielding data curators from liability than for protecting data subjects from harm. Privacy regulation that relies on DP should find ways to recenter data subjects, clearly articulating how DP systems function in practice and making potential risks concrete and visible.

To achieve this goal, data curators need to communicate about DP to data subjects in ways that make the potential harms of data sharing as tangible as the benefits. In part, this is a language issue: They must ensure that the DP "sales pitch" is not, functionally, a mechanism for extracting consent for data collection. It is easy to colloquially describe DP as a tool that "protects data subjects from arbitrary adversaries," but this wording masks the inevitable leakage that comes from data processing, misrepresenting a harm reduction project as a harm elimination project. Other approaches involve demonstrating the kinds of protections afforded (or not afforded) by DP data

processing. If data processing enables certain potentially risky inferences, then they ought to be publicly disclosed. For example, when the U.S. Census Bureau wanted to demonstrate the efficacy of their privacy-enhancing technology, they executed multiple reconstruction attacks giving tangible evidence for the harms reduced while acknowledging that not all risks were eliminated. While there are many such avenues for making harms more concrete, the end goal remains to ensure privacy risks are adequately communicated to data subjects.

5.2.2 Balancing Data Subject Participation with Expertise. In most DP systems, consent is the only mechanism through which data subjects negotiate their relationship with the data curator. There are rarely opportunities, for example, for data subjects to specify the PLB that matches their individual risk tolerance. (And as we have seen, most data subjects would be ill-prepared to do so if they were given the chance.) Given the importance of query choice and PLBs in determining the real meaning of DP's protections, the concerns of data subjects should be incorporated into privacy decision-making through collaborative, participatory governance that includes both experts and lay stakeholders. A participatory process would give data subjects some say in upper-bounding privacy loss budgets for carefully selected queries, where currently there is little forcing such constraints.

As with all participatory approaches to data governance, such an undertaking is far easier said than done. Participatory design approaches often struggle to capture nuanced values like accountability when intertwined with competing interests [27]. Similarly, skepticism towards technical expertise can degrade the legitimate value privacy engineering perspectives bring to the table [26]. These effects have played out in the Decennial Census, where both data subjects and data users fought for control over how to set DP parameters [9]. Despite these challenges, privacy and political contestation are inescapable, and we ought to embrace such contestation by ensuring data subject representation in the process.

5.3 Governance beyond DP

In the previous section, we outlined modes of DP governance appropriate in cases where DP accurately captures relevant risks. But, as we have seen, DP systems are engineered around a singular conception of privacy risk: unintentional disclosure of personal information attributable to statistical releases. While this focus is important, it captures only one of many possible privacy harms that can result from data processing [56]. By narrowing the scope of privacy protections to individual disclosure risks, DP systems ignore other privacy risks data governance ought to manage.

Again, as we have seen, each individual data subject's disclosure risks are bound up with—relative to—others in the database. Data governance needs to consider these relational and collective dimensions of privacy, and critical legal scholars have begun to develop frameworks for engaging substantively with these issues. For example, Salome Viljoen argues that individualized notions of data privacy "miss the point of data production in a digital economy: to put people into population-based relations with one another" [70]. On its own, DP cannot account for these dynamics.

There are both technical and social avenues for addressing misalignments between DP's formalisms and these broad-based privacy harms. On the technical side, new research in formal privacy methods could help address relational effects by quantifying different kinds of harms simultaneously [8, 53, 76]. Practically, when DP or any formal privacy technique fails to capture relevant privacy harms, regulation should ensure these techniques are not the only tools at work. Such an approach aligns with the politics of "algorithmic realism," which distinguishes mathematically formal approaches to data processing ethics problems from substantive ethical advances [31, 32].

Regardless of approach, understanding the limitations of DP ensures that its precise, mathematical description of certain privacy risks does not obscure other, equally important privacy harms.

6 CONCLUSION

In this article, we have described normative issues that emerge when DP's theoretical properties are translated into real-world data processing systems. To be clear, our goal has not been to diminish the social importance and technical contributions of DP and associated privacy-enhancing technologies. On the contrary, our aim in drawing attention to the way DP implicitly frames privacy protection is to encourage more careful reasoning, discussion, and negotiation about it—to help DP succeed at realizing its promise.

REFERENCES

[1] John M. Abowd, Robert Ashmead, Ryan Cumings-Menon, Simson Garfinkel, Micah Heineck, Christine Heiss, Robert Johns, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, Brett Moran, William Sexton, Matthew Spence, and Pavel Zhuravlev. 2022. The 2020 Census Disclosure Avoidance System TopDown Algorithm. *arxiv preprint arxiv:2204.08986* (2022).

[2] Madeleine Akrich. 1992. The de-scription of technical objects. *Shaping Technology Building Society: Studies in Sociotechnical Change* (1992), 205–224.

[3] Micah Altman, Aloni Cohen, Kobbi Nissim, and Alexandra Wood. 2021. What a hybrid legal-technical analysis teaches us about privacy regulation: The case of singling out. *BUJ Sci. Tech. L.* 27 (2021), 1.

[4] Kenneth A. Bamberger and Deirdre K. Mulligan. 2015. *Privacy on the Ground: Driving Corporate Behavior in the United States and Europe.* MIT Press.

[5] Solon Barocas and Karen Levy. 2020. Privacy dependencies. *Wash. L. Rev.* 95 (2020), 555.

[6] Solon Barocas and Helen Nissenbaum. 2014. Big data's end run around anonymity and consent. *Privac. Big Data, Pub. Good: Framew. Engag.* 1 (2014), 44–75.

[7] Ann Bartow. 2006. A feeling of unease about privacy law. *U. Pa. L. Rev. PENNumbra* 155 (2006), 52.

[8] Sebastian Benthall, Seda Gürses, and Helen Nissenbaum. 2017. *Contextual Integrity through the Lens of Computer Science.* Now Publishers.

[9] Danah Boyd and Jayshree Sarathy. 2022. Differential perspectives: Epistemic disconnects surrounding the US Census Bureau's use of differential privacy. *Harv. Data Sci. Rev.* 2 (2022).

[10] Mark Bun, Damien Desfontaines, Cynthia Dwork, Moni Naor, Kobbi Nissim, Aaron Roth, Adam Smith, Thomas Steinke, Jonathan Ullman, and Salil Vadhan. 2021. Statistical inference is not a privacy violation. *Posted June* 3 (2021), 2021. https://differentialprivacy.org/inference-is-not-a-privacy-violation/

[11] Albert Cheu, Adam Smith, Jonathan Ullman, David Zeber, and Maxim Zhilyaev. 2019. Distributed differential privacy via shuffling. In *38th Annual International Conference on the Theory and Applications of Cryptographic Techniques: Advances in Cryptology (EUROCRYPT'19).* Springer, 375–403.

[12] Dennis Chong and James N. Druckman. 2007. Framing theory. *Ann. Rev. Polit. Sci.* 10, 1 (June 2007), 103–126. DOI: https://doi.org/10.1146/annurev.polisci.10.072805.103054

[13] Clare Birchall. 2021. *Radical Secrecy: The Ends of Transparency in Datafied America.* University of Minnesota Press.

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

[15] Damien Desfontaines. 2021. A list of real-world uses of differential privacy. Retrieved from https://desfontain.es/privacy/real-world-differential-privacy.html

[16] Claes H. de Vreese and Sophie Lecheler. 2016. *Framing Theory* (1st ed.). Wiley, 1–10. DOI: https://doi.org/10.1002/9781118541555.wbiepc121

[17] Damien Desfontaines and Balázs Pejó. 2019. Sok: differential privacies. *arXiv preprint arXiv:1906.01337* (2019).

[18] Irit Dinur and Kobbi Nissim. 2003. Revealing information while preserving privacy. In *22nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems.* 202–210.

[19] Cynthia Dwork. 2011. A firm foundation for private data analysis. *Commun. ACM* 54, 1 (2011), 86–95.

[20] Cynthia Dwork, Nitin Kohli, and Deirdre Mulligan. 2019. Differential privacy in practice: Expose your epsilons! *J. Privac. Confident.* 9, 2 (2019).

[21] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference.* Springer, 265–284.

[22] Cynthia Dwork and Deirdre K. Mulligan. 2013. It's not privacy, and it's not fair. *Stan. L. Rev. Online* 66 (2013), 35.

[23] Cynthia Dwork and Aaron Roth. 2014. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.* 9, 3-4 (2014), 211–407.
[24] Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman. 2017. Exposed! A survey of attacks on private data. *Ann. Rev. Stat. Applic.* 4 (2017), 61–84.
[25] Alexandre Evfimievski, Johannes Gehrke, and Ramakrishnan Srikant. 2003. Limiting privacy breaches in privacy preserving data mining. In *22nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*. 211–222.
[26] Gil Eyal. 2019. *The Crisis of Expertise*. John Wiley & Sons.
[27] Christopher Frauenberger, Judith Good, Geraldine Fitzpatrick, and Ole Sejer Iversen. 2015. In pursuit of rigour and accountability in participatory design. *Int. J. Hum.-Comput. Stud.* 74 (2015), 93–106.
[28] Oscar H. Gandy and Oscar H. Gandy Jr. 2021. *The Panoptic Sort: A Political Economy of Personal Information*. Oxford University Press.
[29] Oscar H. Gandy Jr. 1996. Coming to terms with the panoptic sort. *Comput. Surveill. Privac.* (1996), 132.
[30] Simson Garfinkel, John M. Abowd, and Christian Martindale. 2019. Understanding database reconstruction attacks on public data. *Commun. ACM* 62, 3 (2019), 46–53.
[31] Ben Green. 2021. Escaping the "Impossibility of Fairness": From Formal to Substantive Algorithmic Fairness. *arXiv preprint arXiv:2107.04642* (2021).
[32] Ben Green and Salomé Viljoen. 2020. Algorithmic realism: Expanding the boundaries of algorithmic thought. In *Conference on Fairness, Accountability, and Transparency*. 19–31.
[33] Anco Hundepool, Josep Domingo-Ferrer, Luisa Franconi, Sarah Giessing, Eric Schulte Nordholt, Keith Spicer, and Peter-Paul De Wolf. 2012. *Statistical Disclosure Control*. Vol. 2. Wiley New York.
[34] John M. Abowd. 2021. Third Declaration of John M. Abowd in "Fair Lines America Foundation, Inc. v. United States Department of Commerce and United States Bureau of the Census." Testimony. https://www2.census.gov/about/policies/foia/records/disclosure-avoidance/17-1-abowd-decl-3.pdf
[35] Zach Jorgensen, Ting Yu, and Graham Cormode. 2015. Conservative or liberal? Personalized differential privacy. In *IEEE 31st International Conference on Data Engineering*. IEEE, 1023–1034.
[36] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. 2015. The composition theorem for differential privacy. In *International Conference on Machine Learning*. PMLR, 1376–1385.
[37] Shiva Prasad Kasiviswanathan, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. 2011. What can we learn privately? *SIAM J. Comput.* 40, 3 (2011), 793–826.
[38] Michael Kearns and Aaron Roth. 2019. *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*. Oxford University Press.
[39] John Law. 2017. *STS as Method* (4th ed.). MIT Press, 31.
[40] Karen E. C. Levy. 2013. Relational big data. *Stan. L. Rev. Online* 66 (2013), 73.
[41] Karen E. C. Levy and David Merritt Johns. 2016. When open data is a Trojan Horse: The weaponization of transparency in science and governance. *Big Data Soc.* 3, 1 (2016), 2053951715621568.
[42] Yehida Lindell. 2005. Secure multiparty computation for privacy preserving data mining. In *Encyclopedia of Data Warehousing and Mining*. IGI Global, 1005–1009.
[43] Alice E. Marwick and Danah Boyd. 2014. Networked privacy: How teenagers negotiate context in social media. *New Media Soc.* 16, 7 (2014), 1051–1067.
[44] Peter S. Menell. 2017. Tailoring a public policy exception to trade secret protection. *Calif. L. Rev.* 105 (2017), 1.
[45] Jacob Metcalf, Emanuel Moss, and Danah Boyd. 2019. Owning ethics: Corporate logics, Silicon Valley, and the institutionalization of ethics. *Soc. Res.: Int. Quart.* 86, 2 (2019), 449–476.
[46] Deirdre K. Mulligan, Colin Koopman, and Nick Doty. 2016. Privacy is an essentially contested concept: A multi-dimensional analytic for mapping privacy. *Philos. Trans. R. Soc. A: Math., Phys. Eng. Sci.* 374, 2083 (2016), 20160118.
[47] Arvind Narayanan and Vitaly Shmatikov. 2008. Robust de-anonymization of large sparse datasets. In *IEEE Symposium on Security and Privacy (SP'08)*. IEEE, 111–125.
[48] National Academies of Sciences Engineering and Medicine. 2020. 2020 Census Data Products: Data Needs and Privacy Considerations: Proceedings of a Workshop. Workshop proceedings. https://nap.nationalacademies.org/catalog/25978/2020-census-data-products-data-needs-and-privacy-considerations-proceedings
[49] Thomas E. Nelson, Zoe M. Oxley, and Rosalee A. Clawson. 1997. Toward a psychology of framing effects. *Polit. Behav.* 19 (1997), 221–246.
[50] Paul Ohm. 2009. Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA L. Rev.* 57 (2009), 1701.
[51] Phillip Rogaway. 2015. The moral character of cryptographic work. Cryptology ePrint Archive. https://eprint.iacr.org/2015/1162
[52] Lucas Rosenblatt, Joshua Allen, and Julia Stoyanovich. 2022. Spending Privacy Budget Fairly and Wisely. *arXiv preprint arXiv:2204.12903* (2022).

[53] Jeremy Seeman, Aleksandra Slavkovic, and Matthew Reimherr. 2022. A Formal Privacy Framework for Partially Private Data. *arXiv preprint arXiv:2204.01102* (2022).
[54] Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. In *Conference on Fairness, Accountability, and Transparency*. 59–68.
[55] Mary Anne Smart, Dhruv Sood, and Kristen Vaccaro. 2022. Understanding risks of privacy theater with differential privacy. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2 (2022), 1–24.
[56] Daniel J. Solove. 2005. A taxonomy of privacy. *U. Pa. L. Rev.* 154 (2005), 477.
[57] Daniel J. Solove. 2012. Introduction: Privacy self-management and the consent dilemma. *Harv. L. Rev.* 126 (2012), 1880.
[58] Luke Stark. 2016. The emotional context of information privacy. *Inf. Soc.* 32, 1 (2016), 14–27.
[59] Ryan Steed and Alessandro Acquisti. 2023. Privacy-preserving analytics on the ground. *Privacy Law Scholars Conference*.
[60] Daniel Susser. 2016. Information privacy and social self-authorship. *Techné: Res. Philos. Technol.* 20, 3 (2016), 216–239.
[61] Daniel Susser. 2019. Notice after notice-and-consent: Why privacy disclosures are valuable even if consent frameworks aren't. *J. Inf. Polic.* 9, 1 (2019), 148–173.
[62] Daniel Susser. 2021. Predictive policing and the ethics of preemption. *The Ethics Policing: New Perspect. L. Enforc.* (2021), 268.
[63] Daniel Susser, Beate Roessler, and Helen Nissenbaum. 2019. Online manipulation: Hidden influences in a digital world. *Geo. L. Tech. Rev.* 4 (2019), 1.
[64] Jun Tang, Aleksandra Korolova, Xiaolong Bai, Xueqiang Wang, and Xiaofeng Wang. 2017. Privacy loss in Apple's implementation of differential privacy on Macos 10.12. *arXiv preprint arXiv:1709.02753* (2017).
[65] Linnet Taylor, Luciano Floridi, and Bart Van der Sloot. 2016. *Group Privacy: New Challenges of Data Technologies*. Vol. 126. Springer.
[66] Michael Carl Tschantz, Shayak Sen, and Anupam Datta. 2020. SoK: Differential privacy as a causal property. In *IEEE Symposium on Security and Privacy (SP'20)*. IEEE, 354–371.
[67] Amos Tversky and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211, 4481 (Jan. 1981), 453–458. DOI: https://doi.org/10.1126/science.7455683
[68] Anton Vedder. 1999. KDD: The challenge to individualism. *Ethics Inf. Technol.* 1, 4 (1999), 275–281.
[69] Salomé Viljoen. 2020. Democratic data: A relational theory for data governance. Retrieved from SSRN 3727562 (2020).
[70] Salome Viljoen. 2021. A relational theory of data governance. *Yale L. J.* 131 (2021), 573.
[71] Sandra Wachter and Brent Mittelstadt. 2019. A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. *Colum. Bus. L. Rev.* (2019), 494.
[72] Ari Ezra Waldman. 2019. Privacy law's false promise. *Wash. U. L. Rev.* 97 (2019), 773.
[73] Felix T. Wu. 2013. Defining privacy and utility in data sets. *U. Colo. L. Rev.* 84 (2013), 1117.
[74] Yingtai Xiao, Zeyu Ding, Yuxin Wang, Danfeng Zhang, and Daniel Kifer. 2021. Optimizing fitness-for-use of differentially private linear queries. *Proc. VLDB Endow.* 14, 10 (2021), 1730–1742.
[75] Aiping Xiong, Chuhao Wu, Tianhao Wang, Robert W. Proctor, Jeremiah Blocki, Ninghui Li, and Somesh Jha. 2022. Using Illustrations to Communicate Differential Privacy Trust Models: An Investigation of Users' Comprehension, Perception, and Data Sharing Decision. *arXiv preprint arXiv:2202.10014* (2022).
[76] Wanrong Zhang, Olga Ohrimenko, and Rachel Cummings. 2022. Attribute privacy: Framework and mechanisms. In *ACM Conference on Fairness, Accountability, and Transparency*. 757–766.

Received 18 December 2022; revised 9 august 2023; accepted 20 December 2022