# From algorithmic to institutional logics: the politics of differential privacy

JAYSHREE SARATHY

Harvard University

Over the past two decades, we have come to see that traditional de-anonymization techniques fail to protect the privacy of individuals in sensitive datasets. To address this problem, computer scientists introduced differential privacy, a strong definition of privacy that bounds the amount of information a statistical release leaks about any individual. Differential privacy has become a gold standard for privacy protection: organizations from Google to the U.S. Census Bureau have adopted differentially private methods, and the MIT Technology Review named it as one of the top ten technologies expected to have "widespread consequences for human life." Yet, while differential privacy offers rigorous statistical guarantees, we must also examine how these guarantees interact with social and contextual factors. In this paper, I investigate the political dimensions of differential privacy. What does the adoption of this standard reveal or obscure about the privacy practices within sociotechnical systems? And how might a reliance on this standard impact our progress towards broader notions of privacy? Drawing on scholarship from sociology, law, computer science, and science and technology studies, I describe the entanglements between algorithmic privacy and institutional logics, highlighting disempowering practices that may emerge despite, or in response to, the adoption of differential privacy. The goal of this work is not to discourage the use of differential privacy, which I argue is necessary and beneficial in a wide range of settings, but to examine where it may have unintended consequences. I conclude with recommendations on how the privacy community can continue to develop formal privacy standards while elevating broader visions of privacy.

KEYWORDS • differential privacy • politics of technology • science & technology studies

## 1 INTRODUCTION

Over the last two decades, a proliferation of data and computational techniques has given rise to a new capability for privacy harms. Previously, data collection agencies attempted to protect against these violations through heuristic approaches, but numerous attacks have demonstrated that traditional anonymization techniques do not preclude re-identification of individuals in datasets [48,51] or reconstruction of their sensitive attributes [19]. To address this problem, computer scientists introduced differential privacy [22], a framework to measure and control how much information a statistical release reveals about any individual. Differential privacy has gained legitimacy and traction in the last few years: organizations from Google to the

U.S. Census Bureau have adopted differentially private methods [3,26], and in 2020, the MIT Technology Review named differential privacy as one of the top ten technologies expected to have "widespread consequences for human life" [49].

While differential privacy offers rigorous guarantees for statistical disclosure limitation, its algorithmic formalisms [35] do not account for social and contextual factors that impact the amount of privacy actually achieved in the real world. Recent works analyze the differential privacy deployments within statistical agencies [9–11,20], design and governance of privacy-preserving data analysis tools [4,12,18,38,79], and gaps among various notions of privacy [5,6,56,78]. Concurrent work also considers the politics of the axioms of formal privacy, highlighting how the implicit goals and assumptions of formal privacy render certain privacy risks (il)legitimate and privilege data curators and platforms over data subjects [65]. All of these works highlight the importance of examining not just the technical details, but also the values and assumptions embedded within differential privacy. But to truly interrogate the social impacts of this definition of privacy, we must also critically examine the entanglements between algorithmic protections and institutional logics[1] that govern privacy, particularly in the context of industry deployments. In this paper, I ask: How does the adoption of differential privacy reify or subvert existing power dynamics and privacy practices? What does the rapid stabilization of differential privacy reveal or obscure about our sociotechnical systems? And how might a reliance on this standard impact our progress towards a privacy-centric digital ecosystem? Drawing on scholarship from sociology, law, computer science, and science and technology studies, as well as on my experience as a differential privacy researcher, I highlight tradeoffs that may emerge in response to the adoption of differential privacy.

First, reducing privacy to a simple, quantifiable guarantee makes it possible for institutions to performatively adhere to this standard through technological means, without undergoing restructurings and engaging in democratic processes that are fundamental to the project of 'privacy as power' [72]. Second, elevating one definition of privacy, especially one that encourages the use of technical language to discuss risks and harms, can foreclose productive contestations around definitions of privacy and devalue critiques that are not easily expressed in mathematical terms. Third, privacy technologies that offer protection for some harms may encourage privacy violations in other realms, and can be used to legitimate inherently privacy-violating operations. And finally, even when differential privacy is designed for decentralization, its reliance on technical expertise and its modes of adoption in practice may reinforce centralized power.

Overall, I find that a privacy discourse dominated by any single – and mathematically powerful – standard risks ascribing the role of privacy protection solely to technological artifacts, rather than to the social, political, and economic orders that are co-produced [41] along with these technologies. Indeed, this paper argues that reducing the broad and multi-faceted nature of privacy to a narrow yet alluringly elegant technical definition is part of the motivation for Big Tech companies to adopt differential privacy; it allows these institutions to achieve closure [59] of the privacy problem without changing their underlying values and practices. We must be wary of closure as a means to shift attention away from core logics of these institutions, as this may restrict

[1] Institutional logics are defined as "systems of cultural elements (values, beliefs, and normative expectations) by which people, groups, and organizations make sense of and evaluate their everyday activities, and organize those activities in time and space" [39].

our broader visions of privacy and hinder our ability to collectively work towards such visions. Therefore, while differential privacy is a valuable standard, we must remain critical of its adoption by institutions whose logics are fundamentally misaligned with preserving privacy.

The goal of this article is not to discourage the use of differential privacy, which can be powerful and valuable in a wide range of settings, but to examine where it may have unintended consequences. In bringing these scenarios to light, I hope to encourage more nuanced development and deployment of privacy technologies. I conclude with recommendations on how the privacy community can continue to develop formal algorithmic standards while elevating broader visions of privacy, including: recognizing the deeply contextual nature of privacy tools; embracing a plurality of privacy definitions, especially those that center holistic analyses of power and politics; designing participatory processes in tandem with decentralized privacy technologies; and translating technical results towards meaningful political action.

## 2 NOTIONS OF INFORMATIONAL PRIVACY

Privacy law in the United States[2] is often traced back to an essay by Warren and Brandeis in 1890 [75]. Concerned with developments in photography and print media that threatened to invade the "sacred precincts of private and domestic life," Warren and Brandeis argued that individuals have the "right to be let alone." Their seminal essay gave rise to the four privacy torts (taxonomized in Prosser's 1960 article, "Privacy" [62]) that continue to guide courts' interpretations of information privacy interests today.

In the late twentieth century, scholars began to analyze privacy through critical perspectives of power. Most famously, Foucault [27] understood technological surveillance as not just a violation of individual rights, but also a mode of discipline and societal control. His work argues that as tools of surveillance, judgment, and examination attain the capability to target individuals more precisely, they perpetuate a "descending individualization" that constitutes societal domination.

Recently, in light of rapid advances in digital technology and social media, scholars in law, sociology, and computer science have once again aimed to clarify definitions of informational privacy. Daniel Solove argues that privacy is too vague a term to have meaningful legal force and proposes a taxonomy of sixteen privacy violations, which fall under four categories: information collection, information processing, information dissemination, and invasion [68]. Helen Nissenbaum offers a more fluid approach, proposing an account of privacy based on the idea of a plurality of overlapping spheres of life, or contexts, which are constituted by a distinct set of norms. Her framework of contextual integrity [55] demands that information gathering and dissemination follow the norms of distribution within the given context. Computer scientists, meanwhile, have aimed to mathematically formalize societal concepts of informational privacy, developing tools to quantify and mitigate privacy loss online. Recent works have accounted for the information leakage that occurs through analyzing user behavior on social networks [46], linking anonymized datasets [51,52,69], and releasing aggregate statistics [19].

[2] In Griswold v. Connecticut, the Supreme Court recognized that although a right to privacy is not explicitly enshrined in the U.S. constitution, it exists in a penumbra of the First, Third, Fourth, Fifth, and Ninth Amendments.

Scholars are continuously working to bridge the gaps between differentiating notions of privacy across the many fields in which it is studied [16,56], an endeavor crucial for aligning technical, legal, and societal standards. But recognizing the continual contestation around definitions of informational privacy, Deirdre Mulligan, Colin Koopman, and Nick Doty [50] argue that privacy is, in the words of Gallie, an essentially contested concept: similar to concepts such as democracy, art, and freedom, "disputes about the concept's 'essence or meaning' are both paramount and central to the concept itself" [28]. Mulligan et al. call for analytic methods that embrace the multi-faceted, shifting, and situated notions of privacy.

In this paper, I build on Mulligan et al.'s characterization of privacy as an essentially contested concept. By exploring the entanglements between differential privacy and institutional logics, I show how reifying a single notion of privacy can cause us to overlook, and even exacerbate, deeper harms in sociotechnical systems.

## 2.1 The emergence of differential privacy

Although the scope of privacy problems is broad, one specific set of privacy violations — attacks against sensitive datasets or released statistics — has gained significant attention over the last two decades. As computational power and availability of data sources have increased, scholars have outlined three main attacks that can now be carried out against seemingly anonymized datasets or statistical releases [24]. The first is a data linkage attack, where an adversary cross-references the anonymized dataset with auxiliary data sources to identify a specific individual's information. The second is a membership inference attack, where an attacker uses her auxiliary knowledge of a target individual's data to determine whether or not the individual is included in a sensitive dataset. The third is a reconstruction attack, where the adversary approximately infers the sensitive attribute for all individuals in a dataset.

Numerous studies over the last two decades have demonstrated that heuristic anonymization techniques, such as removing personally identifiable information (PII), swapping cells, suppressing outlier values, or only releasing aggregate statistics, cannot hold up against these increasingly powerful attacks [19,51]. The main reason for this is that PII-based protections are premised on a binary notion of privacy risk, where information is either identifying or not, which is inconsistent with a modern understanding of how privacy loss degrades with successive releases on sensitive data and how arbitrary information can be used to infer an individual's sensitive attributes [52].

In light of this growing asymmetry between traditional protections and modern attacks, computer scientists introduced *differential privacy* [22], a definition that quantifies how much information a statistical release reveals about any one individual in the dataset. The definition is parametrized by a quantity that denotes the *privacy loss* incurred by running a given set of analyses on the data. In order to satisfy a guarantee of small privacy loss, the analyses must add carefully calibrated noise to any computation based on the data. The mathematical formalization of differential privacy is both simple and powerful; it accounts for current and future attacks, remains robust to arbitrary auxiliary information, and measures compositions of privacy loss over multiple data releases [77]. Differential privacy also offers an unprecedented amount of transparency; unlike heuristic approaches that rely on security by obscurity,[3] differential privacy enables third-party scrutiny

[3] Security through obscurity is defined as "the reliance in security engineering on design or implementation secrecy as the main method of providing security to a system or component" [81]. This view has been rejected by security experts for more than a century, and there is widespread consensus that obscurity should not be relied upon to provide security or privacy.

of the algorithm, which makes it possible to generate confidence intervals and other uncertainty measures for the statistical releases [77].

Today, more than a decade after it was first proposed, differential privacy has become a gold standard for measuring and controlling privacy loss for statistical releases on sensitive datasets. As publics are sounding the alarm on data privacy [8], and as data holders are coming to understand the deficiencies of heuristic approaches [3,6], institutions are rushing to meet the standard of differential privacy. Disclosure methods that satisfy this standard have been adopted by a variety of data-collection agencies and institutions, including Google, Apple, Uber, Meta (Facebook), Microsoft, and the U.S. Census Bureau [3,71]. As a brief summary, Google uses differential privacy to analyze data about the browser settings of Chrome users [26], Apple uses differential privacy to collect emoji and word usage data on iOS 10 and macOS 10.2 [36], Uber applies differential privacy to understand trip trends [43], Meta offers privacy-protected data for research on social media's role in elections [53], and Microsoft has partnered with Harvard's OpenDP Initiative to launch an open-source differential privacy platform [44]. Google also used differential privacy to produce COVID-19 mobility reports, "provid[ing] insights into what [movement trends have] changed in response to policies aimed at combating COVID-19" [31].

The largest-scale deployment of differential privacy to date, however, is by the U.S. Census Bureau in its 2020 Decennial Census [3]. The Bureau is mandated by Title 13, Sections 8 and 9 of the United States Code to protect the confidentiality of individuals who are represented in census data. This mandate is seriously enforced. Officers who engage in violations can be fined up to $250,000 and imprisoned for five years. Beyond a legal obligation, however, the Bureau is keenly aware that "the quality and accuracy of [its] censuses and surveys depend on [its] ability to keep the public's trust," and therefore, researchers within the Bureau have emphasized that a commitment to privacy is "a core component of [the Bureau's] institutional culture" [40].

The Bureau has traditionally used statistical disclosure limitation (SDL) techniques such as top-and bottom-coding, suppression, rounding, binning, noise-infusion, and sampling to prevent privacy attacks on census data [3]. In 2018, however, scientists at the Bureau found that they were able to use the published data tables from the 2010 decennial census to accurately reconstruct the census block, sex, age, race, and ethnicity for 42% of the population (71% when allowing for age to vary by +/- one year). Then, by linking the reconstructed data table with commercially available datasets, they were able to correctly assign a name and address to the reconstructed record for at least 17% of the population, or 52 million individuals [3]. Based on the demonstrated potential for reconstruction and re-identification attacks, and the scientific consensus that there is only one known approach to protect against such attacks [14], the Bureau's leadership announced that for its 2020 Decennial Census, it would transition to a modern and mathematically formal method of disclosure limitation: namely, differential privacy.

This announcement was reassuring to those concerned about census data confidentiality, but it was met with uproar by social science researchers and policymakers who rely heavily on census data [10,11,64,82]. Decennial Census data is used for apportionment of the House of Representatives, allocation of federal tax dollars, and redistricting within states, as well as for research, evidence-based policy-making, and business and investment decision-making. danah boyd explains that the decision to modernize has been controversial for a variety of reasons. For instance, due to the new understanding of privacy risks, the Bureau will not be releasing some data products that were previously available for public use. In addition, the Bureau has used differential privacy to offer "maximal transparency about technical process," which has unfortunately drawn

attention to previous approaches to data privacy and may be seeding new forms of distrust in census data [10]. Indeed, sources of error have historically not been made explicit, and while differential privacy fixes this failing, it will take time and concerted rebuilding of the "statistical imaginary" around census data for data users to become accustomed to taking error—which has always existed—into account [11] .

It is interesting to contrast the reactions to the Census Bureau's rollout of differential privacy with reactions to deployments by tech companies that have far less stringent norms around data privacy. This latter set of deployments has not been met with nearly as much skepticism or controversy (except when the technology has been deployed incorrectly [37]), and the media has largely applauded the adoption of a rigorous, transparent metric to quantify privacy loss in these settings [73]. Given the maelstrom around the Bureau's use of differential privacy, it is striking that the tech companies' deployments have not been debated or critically analyzed by scholars to nearly the same degree. For this reason, this paper focuses on the use of differential privacy in industry deployments, leaving analysis and critique of the Census Bureau's deployment to other work [9–11]. While differential privacy is seen as the way to address privacy harms across all of these different institutions, we must ask: What does the adoption of differentially private methods reveal or obscure about the privacy practices within these various sociotechnical systems? In reifying the standard of differential privacy, what else becomes stabilized? And how might a reliance on this technical standard shape our "sociotechnical imaginaries" [42] of a digital ecosystem that respects and protects privacy?

## 3 CRITICAL PERSPECTIVES OF DIFFERENTIAL PRIVACY

In this section, I will bring together characterizations of privacy described above with perspectives from science and technology studies. As a methodological note, these insights are not derived from ethnographic data, but rather from my personal experience as a computer science researcher involved in the theoretical and practical development of differential privacy,[4] coupled with my training in science and technology studies. In particular, I will draw on the concept of co-production [41], which calls attention to how science, technology, and society are mutually constitutive, as well as on themes of visibility, discourses of reason, politics of technology, and moral dimensions of privacy research. On the one hand, these perspectives demonstrate the various ways in which formal models such as differential privacy are a much-needed improvement over prior, heuristic approaches to data privacy. At the same time, they highlight how a narrow reliance on differential privacy may inadvertently reify privacy-violating institutions and stabilize un-democratic modes of governance.

### 3.1 Performative privacy

Like many technologies that fall within the umbrella of 'ethical AI,' differential privacy can be used by institutions to perform privacy to the public. One mode of performativity is when differential privacy is deployed with such poor parameters that its protections become negligible. For example, in a presentation in

[4] I am a team member of OpenDP (https://opendp.org/), which is an open-source software project and community organization based at Harvard University that aims to build general-purpose, trustworthy, and usable tools for differential privacy.

2016, Apple's CEO Tim Cook dramatically unveiled the company's deployment of differential privacy for iOS and macOS data collection (see Figure 1). He claimed that unlike Facebook and Google, Apple "does not traffic in your personal data." After six months of reverse engineering, however, researchers found that the privacy protections were so weak that they were relatively pointless [70]. A WIRED article noted that "MacOS uploads significantly more specific data than the typical differential privacy researcher might consider private. iOS 10 uploads even more" [36]. And according to Frank McSherry, one of the inventors of differential privacy, "Apple has put some kind of handcuffs on in how they interact with your data. It just turns out those handcuffs are made out of tissue paper" [36]. Clearly, this first iteration of differential privacy at Apple was more for show than for real protection.

The second mode of performativity is 'privacy theater,' where privacy protections are applied only sparingly across data products, or used on data that is not too valuable to the institution. For example, Google applies differentially privacy for Google Fi location data and Google Maps business meters [54], but given that Google collects vast amounts of personal data, it is likely that Google continues to rely on the sort of


Figure 1: Tim Cook at the Apple Worldwide Web Developer Conference, 2016. Image from theverge.com

hyper-targeted data collection that is incompatible with differential privacy for the majority of its data products. This would not be a surprise; Google's business model "hinges on knowing as much about its users as possible" [54], and as Shoshana Zuboff has argued, the corporation as we know it cannot meaningfully coexist with any substantive protections for privacy [80].

This begs the question: why do institutions adopt privacy technologies incorrectly, or such that they are rendered impotent by the governing institutional logics of data extraction and targeted inference? One answer is that differential privacy's characterization of privacy loss as a mathematically provable, quantitative metric lends authority and projects "organizational legitimacy and perceptions of control and accountability" [60]. As cryptographer Phillip Rogaway puts it, differential privacy appears to policymakers and the public like "cryptomagic [that] protects people from data misuse" [63], although as the examples from Apple and Google indicate, this is not always the case. Adopting differential privacy allows companies to signal a broader

commitment to privacy—a commitment that seems to be enforced by the technology itself—without sacrificing their bottom lines.

In my experience as a differential privacy researcher, I find that the research community encourages such adoption, but for the opposite reason: computer scientists believe that mathematical formalization sets the stage for progress. Quantification of privacy risk seems to offer objectivity and transparency, to encourage the "subordination of interests and prejudices to a public standard" [60]. In many ways, this is true. These early efforts by corporations, even if they are performative, may pave the way for wider rollouts of differential privacy. The deployments so far have already made data providers more comfortable with differential privacy, and they have led to improvements in algorithms and software tools [43,46,58]. In addition, by leveraging the transparency afforded by differential privacy as outlined by Dwork, Kohli and Mulligan [21], it may be possible to incentivize firms to compete with one another and apply increasingly stronger privacy parameters. At the same time, if these firms have dominant market power and no real incentives to change their business models, this strategy may simply provide these firms with a "veneer of objectivity" [60], solidifying power in the very entities whose interests run counter to comprehensive privacy reform and democratic data governance.

As a community, we must be aware of how performative differential privacy, in certain settings, may stabilize unjust hierarchies within our data ecosystem. One place to start is to critically inquire whether institutions allow for participation and contestation by publics. In particular, we might ask: What are the legal, political, and economic obligations of these institutions? Who holds the institutions accountable? Are the authorities in question beholden to a duty of care to the public? For example, even though the U.S. Census Bureau's transition to differential privacy has been controversial, the Bureau in recent years has maintained a strong commitment to privacy and democracy: census officials are sworn under oath to protect confidentiality of responses [40], public forums offer opportunities for participatory deliberation [13], and the Bureau operates within a democratic government.[5]With many data-collecting corporations, however, the picture is less favorable. Privacy technologies may not align with the commitments of the company, and the corporation may have fiduciary obligations to its stakeholders, rather than to the public. The challenge for the privacy community, then, is to advocate for transparent uses of differential privacy that will lead to more robust deployments, without allowing differential privacy to lend legitimacy to institutions that continue to operate under privacy-violating logics.

### 3.2 Technical rhetoric and productive contestations

As we have seen, the strengths of differential privacy render it a valuable marketing strategy to engender favorable public perceptions. But a more benign motivation for its adoption is that technical formalizations of privacy seem to offer tractable definitions around which engineers can more easily build privacy-preserving products. While privacy researchers advocate for using differential privacy as one component within a larger

[5] Of course, each of these commitments have been, and should continue to be, critically analyzed. We should also be careful not to paint an overly rosy picture of the Census Bureau's historical adherence to privacy and ethical uses of data. William Seltzer and Margo Anderson have written extensively about how census data was used to facilitate internment of Japanese-Americans during WWII [66,67]. More recently, special tabulations on Arab-Americans based on the 2000 census were found to have been shared with the Department of Homeland Security [25], which prompted a public debate about the Census Bureau's moral and ethical commitments.

system of protections [6,78], institutions may perceive differential privacy as a clearly defined set of hoops to jump through in order to satisfy legal standards, avoid liability, and improve efficiency [73].

Even though policymakers are wary of blanket reliance on technical solutions [7], ease of compliance remains an important consideration within privacy law. For example, the HIPAA Privacy Rule safe harbor method of de-identification [57], in which the removal of identifiers from a list of 18 identifiers is deemed sufficient to protect privacy in accordance with the law, was settled on as the standard, in part, due to the low cost of compliance. It is possible that a similar situation might occur with differential privacy. The alignment of legal standards and technical methods is important for uptake of any privacy protections, but it is important to consider what is lost in this approach. In particular, if differential privacy becomes the main practice of privacy compliance, an engineer designing an app that collects sensitive data may spend less time thinking about whether some data is too sensitive to be collected, to liaise with the privacy and legal teams within her company, and to conduct user studies to understand attitudes around providing a particular type of sensitive data. A set of algorithmic techniques may simply overshadow the "tangled, ambiguous and essentially contested terrain of privacy" [50].

Furthermore, contestations of privacy that cannot be articulated in technical language, and that cannot be implemented through direct actions of the engineer, may be deemed out of scope [17]. For example, imagine that the mobile application in question collects sensitive fertility data, and suppose that the data is used to derive population-level health information, provide personalized fertility tracking, and train models to target ads to persons seeking to become pregnant. There may be a wide range of concerns about this application, including the privacy of individuals' sensitive attributes, the ethics of targeted advertising within such an intimate setting, and the possibilities for collective governance over the data pipeline and the aggregate insights. These rich, multi-dimensional concerns may, however, become reduced to a debate over the specific differential privacy definition that is used, the privacy parameters applied, and the data collection model that is adopted, as these are the most straightforward for the engineer to explain or improve on her own. The privacy technologies become the legitimate, or de facto [47], way to "settle the matter" [76] on privacy in sociotechnical systems; meanwhile, the ethics and governance questions may be pushed out of the conversation.

Finally, companies may seek to hide behind technical language when privacy violations are exposed to the public. We can look to examples of this prior to the adoption of differential privacy. Three years after the launch of Google Street View (a feature that provides panoramic images of locations around the world), data users found that the cars that photographed neighborhoods for Street View were also scraping "emails, passwords and other private information from Wi-Fi networks in more than thirty countries" [15]. Google initially attempted to deflect blame by characterizing the privacy violation as merely a technical mistake, where experimental code from a different project had accidentally been ported to Street View software. Later, however, investigators found that this code was embedded to "intentionally intercept the data" [15]. This situation highlights the "boundary work" [29] that is performed to separate technologies from their contexts, and in doing so, to sidestep social and political critiques.

Indeed, such boundary work often manifests in companies funding technical research in 'ethical AI' to address the purported shortcomings (usually described narrowly as privacy, fairness, or transparency) of the socio-technical systems in question. Privacy developers and policymakers who participated in Agrawal et al.'s user study on privacy-preserving computation alluded to this very situation, describing how "societal problems

of data processing technologies — such as the ways they create distinctions and hierarchies that reinforce power, shape politics, or facilitate abuse — are sidelined, redefined, or collapsed under the banner of 'privacy', so that privacy-preserving computation techniques can be positioned as the solution" [4] These computational research questions are taken up enthusiastically by computer scientists, as they seem like a tangible way to use the tools of computer science for social good [32]. But a closer look into the history of the funding of ethical AI initiatives reveals a strategy of prioritizing incremental computational goals [1] in order to distract from critiques of core institutional logics.

As companies and institutions adopt the standard of differential privacy, and as privacy law is rewritten to conform with this standard, we must remain aware of how elevating one definition of privacy can foreclose essential contestations around privacy and power. If we believe that the problems of a system are not merely computational ones, we must keep the broader critiques in focus even as we work to strengthen technical infrastructures.

### 3.3 Narrow metrics and other harms

Privacy technologies are designed for use in bounded, computational environments. They capture low-dimensional distillations of the multi-modal, fluid, and essentially contested concept of privacy. By necessity of the formalism of these technologies, each can only address a narrow sliver of privacy harms. Differential privacy, for example, is tailored to prevent reconstruction of individuals' data from statistical analyses. It does not protect against all sixteen privacy harms delineated in Solove's taxonomy [68], nor can it single-handedly account for Nissenbaum's shifting informational norms that characterize contextual spheres of communication [55]. Indeed, it is entirely possible for an institution to adhere to the standards of differential privacy while simultaneously engaging in a multitude of privacy-violating practices.

When we narrow our focus to a single privacy definition and small set of harms, we may miss tradeoffs that are made—both avoidable and otherwise—to accommodate a certain technology. In particular, institutions may strengthen privacy protections in one realm by doubling down on privacy violations in another. For example, differentially private analysis offers "more privacy" on larger datasets. This is because an individual's information is more easily masked within larger amounts of data (e.g., the privacy loss incurred by releasing a noisy mean with a fixed scale of noise decreases as the dataset size increases). This inverse relationship between dataset size and privacy leakage may encourage institutions to expand data collection in order to achieve certain privacy loss guarantees under differential privacy. However, the privacy implications of this additional data collection will not automatically be accounted for within the differential privacy metric.

The narrowness of the differential privacy metric also obscures the sensitivity of the data collected. Privacy regulation that relies on this metric, then, may unintentionally encourage collection of data that was previously deemed too sensitive. Consider Google's COVID-19 mobility reports, which rely on hyper-sensitive location data from users. Google encourages this opt-in data collection by claiming that it is protected with "world-class anonymization technology… including differential privacy" and by highlighting the possibility of providing "helpful" yet "stringent[ly] protected" data products for public use [31]. A modern privacy technology, combined with a compelling public health need, make way for Google to collect data that otherwise may be deemed much too sensitive to track.

People have disagreed on whether or not Google's data collection is acceptable in light of the public health crisis and the differential privacy protections, but as Goldenfein et al. [30] argue, this debate misses the bigger picture:

> "In the debate of "privacy vs. X", where X is a broader social concern like health or security, an individualistic framing of privacy built around dignity and autonomy always loses. What we need is not individual control over data about us, but collective determination over the infrastructures and institutions that process data and that determine how it will be used. This requires moving beyond privacy entailing the choice to opt-in or opt-out of public-private coronavirus surveillance infrastructures and towards developing democratic mechanisms to shape the structure, applications, and agendas of technological architectures."

The authors remind us that in addition to concerns about individual data protection, we must recognize the societal harms of allowing Google to cement new data collection pipelines in place that may be misused in the future. Differential privacy does not create these insidious societal harms, but it does not fully account for them, and it can certainly be used to distract from them. We must remain critical of data pipelines that threaten our collective agency, whether or not these pipelines are packaged with the protections of differential privacy.

The potential disconnect between adoption of privacy technologies and actual privacy protections echoes Ari Waldman's account of symbolic audits in the context of privacy law. He writes: "Toothless trainings, audits, and paper trails, among other symbols, are being confused for actual adherence to privacy law, which has the effect of undermining the promise of greater privacy protection for consumers" [74]. To be sure, differential privacy, by nature of its mathematical formalism and quantitative risk assessment, can sometimes provide more teeth than, say, a subjective privacy audit, which asks the auditor to assess whether the information in question is sensitive or at a high risk of disclosure. But even still, differential privacy may function as a symbolic audit if it is used on its own to measure an *instance* rather than an *underlying logic* of privacy violations—for example, measuring the risk of third-party re-identification of sensitive data rather than the harms of unrestrained data extraction in the first place. When implemented within institutions that are misaligned from the goals of privacy, differential privacy may be reduced to a mere "ritual of inspection" [61], decoupled from broader conceptions of privacy and power.

Researchers have attempted to combat over-optimistic uses of privacy technologies by emphasizing their limits; yet, these results, too, can be exploited.[6] For example, consider the Fundamental Law of Information Recovery [19], which informally says that "overly accurate answers to too many questions will destroy privacy in a spectacular way" [23]. This impossibility result is powerful: it shows that there is 'no free lunch' when it comes to privacy, and that algorithmic protections can only mitigate harm for so long. The result legitimizes political action against data systems that rely on unmitigated data collection, and that are therefore fundamentally misaligned with privacy goals. Yet, we must be aware of the ways in which a result such as this can itself be intentionally misinterpreted. It can be *ignored*: institutions may claim to use differential privacy, while actually placing no restrictions on the number of differentially private computations performed on the sensitive dataset in question. It can be used to *justify poor privacy practices*: institutions may use a weak privacy parameter or increase the amount of data collected, citing the result as evidence that it is hard to

[6] See Ben Green's earlier arguments on the limits of algorithmic fairness [33,34].

generate useful data otherwise. This misinterpretation conflates the limits of differential privacy with the inherent privacy implications of collecting targeted data and performing excessive inferences. Finally, the result can be *twisted to abandon privacy protections entirely*: institutions may use the result to delegitimize any use of disclosure limitation, thereby evading responsibility for implementing any privacy protections at all. It is therefore not sufficient to state theoretical limits of privacy technologies; these limits must also be translated and transformed into action. The limitations of privacy protections should be used to create not only 'epistemic reform' [33], as Ben Green advocates for, but also structural reform around privacy, justice, and collective power in our sociotechnical systems.

### 3.4 Centralized power

Social studies of technology have underscored the ways in which technological artifacts can be "strongly compatible" [76] with certain kinds of political relationships. Privacy technologies are no exception. Even when they are designed to remain decentralized, they can still reinforce centralized power and monopolistic market structures.

One reason for this is that the expertise required for differentially private data analysis lends itself towards centralization. Designing and deploying differentially private algorithms require careful oversight by statisticians, cryptographers, and data scientists [4]. As differential privacy becomes a legal or institutional standard, smaller companies without access to in-house expertise may face greater barriers to analyzing data compared to larger institutions. Therefore, it is critical to continue the ongoing efforts to make differential privacy usable by non-experts—through offering open-source libraries, user-friendly software, and practical education materials [58].

Of course, decentralization and distributed trust architectures in another sense have been a long-standing focus of the cryptography and privacy communities. Differential privacy researchers, too, have paid special attention to decentralized frameworks, and have developed the concept of 'local differential privacy' [45], a model that requires privatization to be applied by individuals before their data is collected by an untrusted party. It is one thing to propose a decentralized model, however, and another to convince institutions to adopt it. In 2015, cryptographer Phillip Rogaway voiced his skepticism that differential privacy would be decentralized in practice, noting that "considerations of efficiency, familiarity, and economics—not to mention authorities' fundamental desire to have and to hold the data—make it easy to predict what will happen: almost always, a centralized design will emerge" [63].

It turns out that Rogaway was overly pessimistic on one account: Google, Apple, and others have actually opted to use the decentralized local model in some of their deployments, in part to shield themselves from subpoenas to turn over user data. Even so, the goal of allowing users to control their own privacy remains mostly unrealized in practice. In these deployments, local differential privacy is implemented at the device-level—not by individual users, but through code designed by Google and Apple to run on the devices. Actions and decisions that were meant to remain in the hands of individuals are made by the centralized 'untrusted party' itself, which to some extent, undermines the goal of local differential privacy to resist institutional power. One might counter that companies such as Google have made their differential privacy code open-source and available for inspection [26], but such transparency does little to shift data collection and decision-making power away from these companies. In institutions that lack participatory deliberation and collective

governance, the centralized deployment of local differential privacy does not allow for crucial contestations about implementation details—what privacy level is chosen, which algorithms are used, and which data falls under the privacy protection—as well as higher-level concerns—who has access to the privatized data, how is the data aggregated and used, and who profits from the insights.

## 4 RESISTING CLOSURE OF THE PRIVACY PROBLEM

In the previous section, STS perspectives about reputation, discourse, metrics, and technological politics allowed us to highlight four disempowering practices that may emerge along with the adoption of differential privacy. Given these four ways in which differential privacy is important yet on its own, insufficient, for addressing the privacy problem, and understanding now how differential privacy can easily be misused, we are led to ask: Why did differential privacy achieve such rapid stabilization across the Big Tech companies? What does the adoption of differential privacy in these contexts reveal about these sociotechnical systems?

To analyze these questions, we can start with the concept of closure, which refers to the stabilization of a technological artifact and the disappearance of problems. Closure does not mean the problem at hand is solved, but rather that the relevant social groups perceive the problem to be solved. Bijker and Pinch [59] identify two closure mechanisms: *rhetorical closure*, which refers to results, proofs, or arguments that close the debate, and *closure by problem redefinition*, which refers to artifacts that are posed as solutions to a re-interpretation of the problem. I add to this list *market closure*, which refers to infrastructural elements that pose economic barriers to the emergence of alternative technologies. I argue that each of the four unintended consequences listed above enables some mechanism of closure. Then, I hypothesize that due to its propensity for closure, differential privacy is appealing for corporations to 'solve the privacy problem' without changing their underlying logics.

Following the arguments from the previous sections, we can outline how the four political qualities of differential privacy facilitate mechanisms of closure. First, performances of differential privacy are in service of the first two mechanisms of closure: applying the technology incorrectly shifts the focus from the privacy problems as a whole to the deficiency in the technical implementations (i.e. closure by problem redefinition), while deploying differential privacy where it may not make a difference shifts public pressure and opinion away from corporations to adequately act on the problem of privacy (i.e. rhetorical closure). Second, the mathematical formalism of differential privacy is compatible with rhetorical closure, as numbers, theorems and proofs have discursive force in closing down debates [60]; criticisms that are not able to be articulated in the language of mathematics or statistics may be deemed out of scope. Third, the narrowness of the differential privacy metric allows for closure by problem redefinition, as the problem of privacy becomes reduced to the problems addressed by differential privacy and broader concerns are more easily sidelined, ignored, or exploited. And finally, the centralizing tendencies of differential privacy facilitate market closure, as the resulting economies of scale pose barriers to entry for smaller corporations and alternative technologies. Therefore, the political qualities of differential privacy are correlated to its rapid stabilization within Big Tech companies over the past decade. Yet, as discussed in the previous section, these qualities also offer reasons to be cautious about blanket adoption of this technology.

Closure of the privacy problem flies directly against the nature of privacy as an essentially contested concept [50]. As Mulligan et al. caution, our debates around privacy must remain open, inter-disciplinary, and ever-

evolving for the concept to remain meaningful. It is undeniably important to update the technical protections of our data systems, yet it is also critical to recognize that closure of privacy as a concept will be detrimental in the long run. This is because closure feeds into the creation of socio-technical imaginaries, which as Jasanoff and Kim define, are "collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology" [42]. Closure of the privacy problem around a single framework such as differential privacy threatens to condense and restrict our socio-technical imaginary of privacy from an expansive, layered, multi-dimensional understanding of privacy to a narrow, technical one.

The rapid move from theory to practice of differential privacy should, therefore, not be uncritically celebrated by technologists and privacy advocates. We must not take as a foregone conclusion that privacy-violating systems will inevitably persist, and that our choice lies between data-extractive pipelines that use differential privacy and ones that don't; the nature and existence of the pipelines themselves must be called into question. Social studies of technology have long highlighted how technological development—and the construction of sociotechnical imaginaries—relies on social acceptance [59], and how researchers have agency in discouraging practices from persisting. Therefore, we must be careful not to reduce a potentially broad critique of the privacy of a sociotechnical system to a narrow technical problem to be overcome, for that shifts responsibility and agency away from the institutions and humans who support these systems. Instead, we must recognize that if technology is political, it is time for us to become involved in the politicking. In the words of Abebe et al. [2], it is the obligation of computer scientists to "acknowledge the political valence of their technical work, [and] make it more difficult for others to leverage computing—both practically and rhetorically— for political ends." Recognizing the moral, technical, and financial value that differential privacy brings to institutions, I believe that developers of differential privacy techniques have more power than they realize in endorsing or criticizing deployments of this technology. I suggest that computer scientists take an active role in negotiating the contexts in which differential privacy is deployed, and by extension, in shaping our broader visions of privacy.

## 5 CONCLUSION

As increases in computational power and availability of data have given rise to new privacy threats, institutions ranging from Google to the U.S. Census Bureau have adopted differential privacy to safeguard individuals' sensitive data. Privacy advocates have celebrated these deployments, and with good reason: compared to prior heuristic approaches, differential privacy offers a provable guarantee and rich suite of algorithmic techniques for formally measuring and controlling privacy loss. But while it is important to strengthen privacy technologies to meet current threats, we must also recognize that the success of privacy practices cannot be reduced to any single technology or metric. Technology is only as privacy-preserving as the social, political, and economic contexts it lives in—and vice versa. If we condense the multi-dimensional landscape of privacy into a single notion, we will miss this bigger picture.

As I have shown in this paper, differential privacy provides a necessary safeguard against certain modern privacy attacks: it acts as a counterweight to structural forces that threaten privacy, provides teeth to privacy audits, and offers a rigorous technical metric for measuring privacy risk. At the same time, a narrow focus on

differential privacy can encourage performativity, foreclose productive contestations, distract from or be used to justify other privacy harms, and reinforce centralized power in sociotechnical systems.

How can we complicate our understandings of privacy, and in doing so, broaden our sociotechnical imaginaries of privacy-preserving futures? We can begin by embracing a plurality of privacy definitions. Privacy is never settled; its definitions are contingent on ever-evolving networks of actors, contexts, and cultural values. It is important to clarify and formalize rigorous definitions of privacy to make them tractable in computational systems, but at the same time, we must make space for productive debates around privacy.

Second, when considering deployments of privacy technologies, we must critically inquire whether institutions allow for participation and contestation by publics. In particular, we might ask: What are the legal, political, and economic obligations of these institutions? Who holds the institutions accountable? Are the authorities in question beholden to a duty of care to the public? And crucially, how might differential privacy interact with these institutional logics? At the same time that we celebrate developments in privacy technologies, we must remain wary of institutions that adopt these technologies as a performative distraction to their privacy-violating logics.

Finally, we must remain critical of deployments that play up the possibilities afforded by privacy technologies without acknowledging their limits. Theoretical work showing that algorithmic privacy cannot be achieved within certain frameworks should be translated towards condemnation, rather than legitimization, of these structures. Only when armed with richer visions of privacy that foreground power, agency, and democracy will we avoid re-creating the futures we wish to resist.

## ACKNOWLEDGEMENTS

Many thanks to Sheila Jasanoff for introducing me to many of the ideas behind this paper, to the Bridging Privacy working group at the Berkman Klein Center and the participants of the 2021 Privacy Law Scholars Conference for generative discussions, and to Salil Vadhan, Alexandra Wood, and Jonathan Zittrain for valuable feedback.

## REFERENCES

[1] Mohamed Abdalla and Moustafa Abdalla. 2021. The Grey Hoodie Project: Big Tobacco, Big Tech, and the threat on academic integrity. *arXiv:2009.13676 [cs]* (2021). DOI:https://doi.org/10.1145/3461702.3462563

[2] Rediet Abebe, Solon Barocas, Jon Kleinberg, Karen Levy, Manish Raghavan, and David G. Robinson. 2020. Roles for Computing in Social Change. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (January 2020), 252–260. DOI:https://doi.org/10.1145/3351095.3372871

[3] John Abowd. 2018. The US Census Bureau adopts differential privacy.

[4] Nitin Agrawal, Reuben Binns, Max Kleek, Kim Laine, and Nigel Shadbolt. 2021. Exploring Design and Governance Challenges in the Development of Privacy-Preserving Computation. DOI:https://doi.org/10.1145/3411764.3445677

[5] Micah Altman, Aloni Cohen, Kobbi Nissim, and Alexandra Wood. 2020. What a Hybrid Legal-Technical Analysis Teaches Us About Privacy Regulation: The Case of Singling Out. *SSRN Journal* (2020). DOI:https://doi.org/10.2139/ssrn.3681729

[6] Micah Altman, Alexandra Wood, David O'Brien, Salil Vadhan, and Urs Gasser. 2016. Towards a Modern Approach to Privacy-Aware Government Data Releases. *Berkeley Technology Law Journal* 30, 3 (2016). Retrieved January 1, 2021 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2779266
[7] Article 29 Data Protection Working Party. 2014. wp216_en.pdf. Retrieved January 19, 2022 from https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf
[8] Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. 2019. Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information. *Pew Research Center: Internet, Science & Tech*. Retrieved from https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/
[9] Dan Bouk and danah boyd. 2021. Democracy's Data Infrastructure. *knightcolumbia.org*. Retrieved from https://knightcolumbia.org/content/democracys-data-infrastructure
[10] danah boyd. 2019. Differential Privacy in the 2020 Decennial Census and the Implications for Available Data Products. *SSRN Electronic Journal* (2019). DOI:https://doi.org/10.2139/ssrn.3416572
[11] danah boyd and Jayshree Sarathy. 2022. *Differential Perspectives: Epistemic Disconnects Surrounding the US Census Bureau's Use of Differential Privacy*. Social Science Research Network. Retrieved April 7, 2022 from https://papers.ssrn.com/abstract=4077426
[12] Brooke Bullek, Stephanie Garboski, Darakhshan J. Mir, and Evan M. Peck. 2017. Towards understanding differential privacy: When do people trust randomized response technique? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3833–3837.
[13] US Census Bureau. Census Advisory Committees (CAC). *Census.gov*. Retrieved January 21, 2022 from https://www.census.gov/cac
[14] Ryan Calo, Ran Canetti, Aloni Cohen, Cynthia Dwork, Roxana Geambasu, Somesh Jha, Nitin Kohli, Aleksandra Korolova, Jing Lei, Katrina Ligett, Deirdre K Mulligan, Omer Reingold, Aaron Roth, and Guy N Rothblum. AMICUS BRIEF OF DATA PRIVACY EXPERTS. 27.
[15] CBSN Bay Area. 2019. Google Agrees To Pay $13 Million In Street View Privacy Case. *CBSN Bay Area*. Retrieved January 1, 2020 from https://sanfrancisco.cbslocal.com/2019/07/22/google-street-view-privacy-lawsuit-settlement/
[16] Aloni Cohen and Kobbi Nissim. 2020. Towards formalizing the GDPR's notion of singling out. *Proceedings of the National Academy of Sciences* 117, 15 (2020), 8344–8352. DOI:https://doi.org/10.1073/pnas.1914598117
[17] Carol Cohn. 1987. Slick'Ems, Glick 'Ems, Christmas Trees, and Cutters: Nuclear Language and how we learned to pat the bomb. *Bulletin of the Atomic Scientists* 43, 5 (1987), 17–24. DOI:https://doi.org/10.1080/00963402.1987.11459533
[18] Rachel Cummings, Gabriel Kaptchuk, and Elissa M. Redmiles. 2021. " I need a better description": An Investigation Into User Expectations For Differential Privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, 3037–3052.
[19] Irit Dinur and Kobbi Nissim. 2003. Revealing information while preserving privacy.
[20] Joerg Drechsler. 2021. Differential Privacy for Government Agencies -- Are We There Yet? *arXiv:2102.08847 [cs, stat]* (February 2021). Retrieved January 22, 2022 from http://arxiv.org/abs/2102.08847
[21] Cynthia Dwork, Nitin Kohli, and Deirdre Mulligan. 2019. Differential Privacy in Practice: Expose your Epsilons! *Journal of Privacy and Confidentiality* 9, 2 (2019). DOI:https://doi.org/10.29012/jpc.689
[22] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis.
[23] Cynthia Dwork and Aaron Roth. 2014. *The algorithmic foundations of differential privacy*. Now Publ.
[24] Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman. 2017. Exposed! A Survey of Attacks on Private Data. *Annual Review of Statistics and Its Application* 4, 1 (2017), 61–84. DOI:https://doi.org/10.1146/annurev-statistics-060116-054123
[25] Samia El-Badry and David A. Swanson. 2005. Controversy over providing special census tabulations to government security agencies in the United States: the case of Arab-Americans. Tours, France. Retrieved July 24, 2021 from https://iussp2005.princeton.edu/papers/51890
[26] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. 2014. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response.
[27] Michel Foucault. 1979. *Discipline and Punish*. Vintage.
[28] W B Gallie. 1956. *Essentially Contested Concepts*. Aristotelian Society.

[29] Thomas F Gieryn. 1983. BoundaryWork and the Demarcation of Science from NonScience: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review* 48, 6 (1983), 781–795. DOI:https://doi.org/10.2307/2095325
[30] Jake Goldenfein, Ben Green, and Salomé Vilijoen. 2020. Privacy Versus Health Is a False Trade-Off. *jacobinmag.com*. Retrieved January 1, 2021 from https://jacobinmag.com/2020/4/privacy-health-surveillance-coronavirus-pandemic-technology
[31] Google. 2020. COVID-19 Community Mobility Report. *COVID-19 Community Mobility Report*. Retrieved from https://www.google.com/covid19/mobility/
[32] Ben Green. 2019. "Good" isn't good enough. Retrieved January 1, 2021 from https://www.benzevgreen.com/wp-content/uploads/2019/11/19-ai4sg.pdf
[33] Ben Green. 2020. The false promise of risk assessments: epistemic reform and the limits of fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ACM, Barcelona Spain, 594–606. DOI:https://doi.org/10.1145/3351095.3372869
[34] Ben Green. 2022. *Escaping the Impossibility of Fairness: From Formal to Substantive Algorithmic Fairness*. Social Science Research Network, Rochester, NY. DOI:https://doi.org/10.2139/ssrn.3883649
[35] Ben Green and Salomé Viljoen. 2020. Algorithmic realism: expanding the boundaries of algorithmic thought. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ACM, Barcelona Spain, 19–31. DOI:https://doi.org/10.1145/3351095.3372840
[36] Andy Greenberg. 2016. Apple's 'Differential Privacy' Is About Collecting Your Data---But Not Your Data. *WIRED*. Retrieved from https://www.wired.com/2016/06/apples-differential-privacy-collecting-data/
[37] Andy Greenberg. 2017. How One of Apple's Key Privacy Safeguards Falls Short. *Wired*. Retrieved January 1, 2020 from https://www.wired.com/story/apple-differential-privacy-shortcomings/
[38] Seda Gurses and Jose M. del Alamo. 2016. Privacy Engineering: Shaping an Emerging Field of Research and Practice. *IEEE Secur. Privacy* 14, 2 (March 2016), 40–46. DOI:https://doi.org/10.1109/MSP.2016.37
[39] Heather A. Haveman and Gillian Gualtieri. 2017. Institutional Logics. *Oxford Research Encyclopedia of Business and Management*. DOI:https://doi.org/10.1093/acrefore/9780190224851.013.137
[40] Michael Hawes. 2019. *Title 13, Differential Privacy, and the 2020 Decennial Census*. Retrieved January 1, 2020 from https://www2.census.gov/about/policies/2019-11-paper-differential-privacy.pdf
[41] Sheila Jasanoff. 2004. *States of knowledge : the co-production of science and social order*. London Routledge.
[42] Sheila Jasanoff and Sang-Hyun Kim (Eds.). 2015. *Dreamscapes of Modernity*. The University of Chicago Press, Chicago. Retrieved July 29, 2021 from https://press.uchicago.edu/ucp/books/book/chicago/D/bo20836025.html
[43] Noah Johnson, Joseph P. Near, and Dawn Song. 2018. Towards practical differential privacy for SQL queries. *Proceedings of the VLDB Endowment* 11, 5 (2018), 526–539. DOI:https://doi.org/10.1145/3187009.3177733
[44] John Kahan. 2020. How differential privacy enhances Microsoft's privacy and security tools: SmartNoise Early Adopter Acceleration Program Launched. *Microsoft On the Issues*. Retrieved January 1, 2021 from https://blogs.microsoft.com/on-the-issues/2020/12/10/differential-privacy-smartnoise-early-adopter-acceleration-program/
[45] Shiva Prasad Kasiviswanathan, Homin K. Lee, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. 2011. What Can We Learn Privately? *SIAM Journal on Computing* 40, 3 (2011), 793–826. DOI:https://doi.org/10.1137/090756090
[46] Aleksandra Korolova. 2012. Protecting Privacy When Mining and Sharing User Data.
[47] Lawrence Lessig. 2000. Code Is Law. *Harvard Magazine*. Retrieved from https://www.harvardmagazine.com/2000/01/code-is-law-html
[48] Steve Lohr. 2010. Netflix Cancels Contest After Concerns Are Raised About Privacy (Published 2010). *The New York Times*. Retrieved January 1, 2020 from https://www.nytimes.com/2010/03/13/technology/13netflix.html
[49] Bernard Marr. 2020. The Top 10 Breakthrough Technologies For 2020. *Forbes*. Retrieved January 1, 2020 from https://www.forbes.com/sites/bernardmarr/2020/02/26/mit-names-top-10-breakthrough-technologies-for-2020/?sh=2267ee68d482
[50] Deirdre K. Mulligan, Colin Koopman, and Nick Doty. 2016. 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 (2016), 20160118. DOI:https://doi.org/10.1098/rsta.2016.0118
[51] Arvind Narayanan and Vitaly Shmatikov. 2006. How To Break Anonymity of the Netflix Prize Dataset. *arXiv:cs/0610105* (2006). Retrieved January 1, 2021 from https://arxiv.org/abs/cs/0610105v1

[52] Arvind Narayanan and Vitaly Shmatikov. 2010. Myths and fallacies of "personally identifiable information." *Communications of the ACM* 53, 6 (2010), 24. DOI:https://doi.org/10.1145/1743546.1743558
[53] Chaya Nayak. 2020. New privacy-protected Facebook data for independent research on social media's impact on democracy. *Facebook Research*. Retrieved January 1, 2021 from https://research.fb.com/blog/2020/02/new-privacy-protected-facebook-data-for-independent-research-on-social-medias-impact-on-democracy/
[54] Lily Hay Newman. 2019. Google Wants to Help Tech Companies Know Less About You. *Wired*. Retrieved from https://www.wired.com/story/google-differential-privacy-open-source/
[55] Helen Nissenbaum. 2011. A Contextual Approach to Privacy Online. *Daedalus* 140, 4 (2011), 32–48. DOI:https://doi.org/10.1162/daed_a_00113
[56] Kobbi Nissim and Alexandra Wood. 2018. Is privacy privacy? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, 2128 (2018), 20170358. DOI:https://doi.org/10.1098/rsta.2017.0358
[57] Office for Civil Rights (OCR. 2012. Methods for De-identification of PHI. *HHS.gov*. Retrieved from https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html
[58] OpenDP. 2020. OpenDP. *projects.iq.harvard.edu*. Retrieved January 1, 2020 from https://projects.iq.harvard.edu/opendp
[59] Trevor J. Pinch and Wiebe E. Bijker. 1984. The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Soc Stud Sci* 14, 3 (August 1984), 399–441. DOI:https://doi.org/10.1177/030631284014003004
[60] Theodore M Porter. 1995. *Trust in numbers : the pursuit of objectivity in science and public life*. Princeton University Press.
[61] Michael Power. 1997. *The Audit Society*. Oxford University Press.
[62] William L. Prosser (Ed.). 1960. Privacy. *Calif. L. Rev.* (1960). DOI:https://doi.org/10.15779/Z383J3C
[63] Phillip Rogaway. 2015. *The Moral Character of Cryptographic Work*. Retrieved January 1, 2021 from https://web.cs.ucdavis.edu/~rogaway/papers/moral-fn.pdf
[64] Steven Ruggles, Catherine Fitch, Diana Magnuson, and Jonathan Schroeder. 2019. Differential Privacy and Census Data: Implications for Social and Economic Research. *AEA Papers and Proceedings* 109, (May 2019), 403–408. DOI:https://doi.org/10.1257/pandp.20191107
[65] Jeremy Seeman. The Politics of Formal Privacy's Axioms.
[66] William Seltzer and Margo J. Anderson. 2000. After Pearl Harbor: The Proper Role of Population Data Systems in Time of War. Los Angeles, CA. Retrieved July 24, 2021 from https://margoanderson.org/govstat/newpaa.pdf
[67] William Seltzer and Margo J. Anderson. 2007. Census Confidentiality under the Second War Powers Act (1942-1947). In *Confidentiality, Privacy, and Ethical Issues in Demographic Data*, New York, NY. Retrieved July 24, 2021 from https://paa2007.princeton.edu/abstracts/70299
[68] Daniel J. Solove. 2006. A Taxonomy of Privacy. *University of Pennsylvania Law Review* 154, 3 (2006), 477. DOI:https://doi.org/10.2307/40041279
[69] Latanya Sweeney. 2002. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 05 (2002), 557–570. DOI:https://doi.org/10.1142/s0218488502001648
[70] 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*. Retrieved January 1, 2020 from http://theory.stanford.edu/~korolova/Privacy_Loss_in_Apple%27s_Implementation_of_Differential_Privacy.pdf
[71] U.S. Census Bureau. OnTheMap. *onthemap.ces.census.gov*. Retrieved from http://onthemap.ces.census.gov
[72] Carissa Véliz. 2019. Privacy matters because it empowers us all – Carissa Véliz | Aeon Essays. *Aeon*. Retrieved January 1, 2020 from https://aeon.co/essays/privacy-matters-because-it-empowers-us-all
[73] James Vincent. 2019. Google is making it easier for AI developers to keep users' data private. *The Verge*. Retrieved January 21, 2022 from https://www.theverge.com/2019/3/6/18253002/google-ai-data-privacy-tensorflow-differential-module-code
[74] Ari Ezra Waldman. 2019. Privacy Law's False Promise. *Washington University Law Review* 97, 2 (2019). DOI:https://doi.org/10.2139/ssrn.3339372
[75] Samuel D. Warren and Louis D. Brandeis. 1890. The Right to Privacy. *Harvard Law Review* 4, 5 (1890), 193. DOI:https://doi.org/10.2307/1321160
[76] Langdon Winner. 1980. *Do artifacts have politics?* University of Chicago Press.

[77] Alexandra Wood, Micah Altman, Aaron Bembenek, Mark Bun, Marco Gaboardi, James Honaker, Kobbi Nissim, David O'Brien, Thomas Steinke, and Salil Vadhan. 2018. Differential Privacy: A Primer for a Non-Technical Audience. *Vanderbilt Journal of Entertainment & Technology Law* (2018). Retrieved January 1, 2021 from https://dash.harvard.edu/handle/1/38323292

[78] Alexandra Wood, Micah Altman, Kobbi Nissim, and Salil Vadhan. 2021. Designing Access with Differential Privacy. *admindatahandbook.mit.edu* (2021). Retrieved January 1, 2021 from https://admindatahandbook.mit.edu/book/v1.0/diffpriv.html

[79] Aiping Xiong, Tianhao Wang, Ninghui Li, and Somesh Jha. 2020. Towards Effective Differential Privacy Communication for Users' Data Sharing Decision and Comprehension. In *2020 IEEE Symposium on Security and Privacy (SP)*, 392–410. DOI:https://doi.org/10.1109/SP40000.2020.00088

[80] Shoshana Zuboff. 2019. *The age of surveillance capitalism the fight for the future at the new frontier of power*. London Profile Books.

[81] 2022. Security through obscurity. *Wikipedia*. Retrieved January 22, 2022 from https://en.wikipedia.org/w/index.php?title=Security_through_obscurity&oldid=1066450161

[82] Alabama v. United States Department of Commerce | Brennan Center for Justice. Retrieved April 8, 2022 from https://www.brennancenter.org/our-work/court-cases/alabama-v-united-states-department-commerce