# Framing Effects in the Operationalization of Differential Privacy Systems as Code-Driven Law

*Jeremy Seeman, Pennsylvania State University (United States)*

*International Conference on Computer Ethics: Philosophical Enquiry (CEPE) 2023, Chicago, IL*

**Keywords:** privacy-enhancing technologies, differential privacy, science and technology studies, critical code studies, political implications of AI use democratic data

## 1 Introduction

Modern data privacy regulations, like the General Data Protection Regulation (GDPR) [16], aim to protect data subjects from personal data processing harms. However, implementing laws like these requires interpreting ambiguous and contested statutes about what constitutes breaches of personal data privacy, made more complicated by the fact that removing unique identifiers from data sets no longer sufficiently anonymizes data [33]. Such debates have been especially important for GDPR's notion of "singling out," one such attempt at locating the bounds of the "personal" in personal data [18]. Computer scientists from theoretical cryptography have recently proposed borrowing tools from the mathematical framework of differential privacy (DP) [12, 14] as an avenue to help formalize compliance with these text-driven legal notions [7, 3, 2]. These proposals suggest DP could facilitate a form of code-driven law [23] in two different ways. First, the kinds of privacy harms quantified by DP could serve as a technical benchmark to audit whether privacy violations have occurred. Second, systems built to process data with DP protections could satisfy a code-driven in terpretation of compliance with privacy law. Both proposals could move legal definitions of anonymization significantly closer to being an automatic byproduct of DP outputs, an ar gument which I'll refer to throughout this paper as "DP-legal reasoning" (DLR). Naturally, this emerging intersection of privacy-enhancing technologies and privacy law poses many un resolved questions about the operational logistics of using DP systems to verify compliance with text-driven law.

From the perspective of critical code studies and science and technology studies, DP sys tems are technological artifacts with their own politics [50] which imply certain "scripts of action" for their users [1]. In particular, mathematical frameworks which attempt to imbue data processing systems with socially desirable values are subject to "framing effects" of how those values are abstracted in their systems [36]. In this paper I investigate how DP systems are influenced by DP's particular mathematical framing of privacy quantification, more specifically how this framing might affect DLR in practice. I argue that treating DP systems as legal apparatuses can both strengthen some substantive privacy protections while obfuscating some potential harms from data sharing. Therefore, for better and worse, DP modulates the terms of the natural contestations that occur when different parties negotiate between the real harms of privacy violations and the social benefits of data sharing. In order to properly assess in what capacity DP systems could act as legal apparatuses, or to deter mine which aspects of this process require different interventions, we must precisely articulate these framing effects. Our past research investigates these framing effects in the differences between DP as formal mathematics versus data processing software as sociotechnical sys tems [35]. This paper extends and applies previous work to omnibus privacy regulations, investigating framing effects in using DLR to inform legal definitions of anonymization.

I briefly summarize the main arguments here. First, I describe precisely how this emergent DP-legal reasoning could act as a form of legislative code-driven law, one that could end up making compliance with privacy law operationalizable by the design principles inherent to DP data processing systems. Just as there are gaps between DP's mathematical formalism and DP as realized in sociotechnical systems, so too are there gaps between DP's mathematical formalism and DP-legal reasoning. Next, I discuss how the way DP frames data privacy is intimately tied to notice and consent. While DP can offer data subjects some inherent insights into their privacy risks, it could exacerbate issues with privacy self-management at large. As applied to DLR, I argue that new notice practices could produce hermeneutical injustices, in which data subjects are deprived of knowledge necessary to understand their role in data processing systems. Next, I discuss how DP's framing enables a particular kind of technical auditing that helps isolate precisely how data processors contribute to individual privacy harms. As applied to DLR, DP's technical requirements and operational burdens for verification may inadvertently allow data processors greater self-regulatory flexibility, which has historically worsened issues with legal endogeneity in privacy law. Lastly, I conclude with a discussion on how DLR can continue to help inform privacy law while addressing these framing effects.

As a disclaimer, none of what follows detracts from the need for research at the intersec tion of emerging technologies and the law, nor the contributions of those who advocate for DLR. There are self-evident benefits both in improving privacy-enhancing technologies and making privacy law more easily interpretable. My goal with this paper is, instead, to unpack the social and normative implications inherent in how these approaches are fundamentally driven by the technical properties of DP as formal mathematical system. Methods derived from formal reasoning and axiomatic logic are often at odds with their substantive ethical goals, not just with privacy but all computational approaches to social problems [21, 20]. As sessing DLR's fitness-for-use in regulatory practice requires careful reconsideration of these abstractions and their sociotechnical effects. DLR advocates add that the simplifications made by this technical reasoning "remain useful so long as they capture a portion of real human activity, bounded by conditions that can be reliably recognized from outside of the formal system" [2]. The purpose of this paper is precisely to address this "big-if" condition by characterizing the degrees to which DLR captures "real human activity" and conditions "reliability recognized from outside the system."

## 2 DLR as Code-Driven Law

To start, we briefly summarize the key technical properties of differential privacy; interested readers may turn to [51] for a more in-depth introduction. While privacy is a multifaceted and multidimensional topic [39], and DP is not a single mathematical definition [10], DP is a mathematical framework centered around one particular kind of privacy harm: individual information disclosure. DP architects aim to develop data processing systems which enable generalized statistical inferences from their outputs while limiting individual information dis closures attributable to the data processing tasks. These disclosures come in many different forms: published data sets with PII removed could be used to reidentify individuals within those data sets, large-scale AI models could memorize and leak personal information used to train them, etc. Yet quantifying these harms proves difficult due to what's known as the "fundamental law of information recovery," which states that any released output statistics inadvertently leak some information about the individuals who contributed to them, regard less of the presence or absence of PII [11]. In an evolving landscape where data about us continually amasses, isolating the effect of any one data processing task for its privacy harms can be difficult to nay impossible [13]. It is therefore impossible for data processors to build any non-trivial statistical model without compromising some data privacy [26]. These "im possibility theorems" bookend the unavoidable trade-off between individual disclosure risks and overall data utility, the central conundrum of all statistical data privacy techniques [38].

Reckoning with these impossibility theorems has meant addressing shortcomings in past anonymization techniques and proposing new design principles. This has led to the devel opment of DP through an axiomatic lens, where privacy properties are deduced from a few key design desiderata [25]. First, DP outputs should have "provable" privacy guarantees, wherein the description of their privacy protections is inherent to how the data is processed. Second, DP outputs should be "future-proof" in that the output's privacy properties should not degrade with further data processing. In other words, it should be impossible to "undo" a privacy-enhancing intervention. Finally, DP outputs should be "transparent," meaning the intervention taken to anonymize the data processing should be visible without increasing any privacy risks.

DP overcomes these challenges and achieves its design goals by choosing a clever unit of analysis: measuring disclosure risk in relative terms by comparing how outputs would change depending on one person's contribution to the output statistic. This explicitly frames DP as a harm-reduction project by isolating changes in disclosure risk due to one individual contribution for one particular data processing task. In other words, DP captures a relative measure of disclosure risk (i.e., a change in disclosure risk between two similar scenarios) as opposed to an absolute measure of disclosure risk (i.e., a direct measure of disclosure risk in one particular scenario). DP systems inject randomized noise into output statistics precisely so that, if the system were to process data on two datasets differing on one person's data, the outputs would be similar with high probability. As more noise is added, there's further plausible deniability about the presence or absence of any one's data on the data processor's part. The amount of noise is quantified by a privacy loss budget (PLB); larger PLB values correspond to less noise and greater disclosure risks, and smaller PLB values correspond to more noise and lesser disclosure risks. By focusing on this unit of analysis, DP maintains the properties discussed above in a provable manner. Protections offered by DP hold true regardless of whatever information an adversarial user might possess, even if they knew in formation about others in the dataset. Moreover, DP results are robust to post-processing, meaning DP protections are "future-proof" against further potential misuse. DP's PLBs allow for privacy loss accounting, wherein multiple data processing outputs can be assessed by combining their respective PLBs. Finally, DP's interventions are methodologically transparent, allowing for visibility into how data processing systems actively work to protect user privacy.

DP as a formal, mathematical system provides its guarantees to a broad suite of data processing tasks with many possible inputs and outputs, each of which consuming some privacy loss. Academic DP research largely focuses on optimal procedures for processing data under fixed DP constraints. Implementing this formal logic in practice requires deciding which inputs and outputs to consider, which methods would be used to process the data for each task, and how much privacy loss to allocate to each task [27, 30]. It's in the process of making these decisions where the gap between DP theory and practice emerges, as the mere presence of DP implemented at a fixed privacy loss for a particular task says little about the sociotechnical trade-offs between privacy and data utility captured by DP. Our past work discusses the effect of DP's framing on sociotechnical systems more broadly[35], but here, we turn to DP's influence on legal reasoning.

Note that for the purposes of this paper, I'll refer to "omnibus privacy laws" as regulations which attempt to holistically strengthen data subject privacy rights through broad definitions of "personal information" possessed by data processors. As many DLR articles focus on GDPR as an emblematic case study, this article, too, will draw heavily on examples from GDPR. Omnibus privacy laws are lengthy and complex, meaning no one set of shared legal principles are essential among them. The arguments made in this paper are no substitute for the critical legal legwork in enforcing GDPR and similar regulations. At the same time, many omnibus privacy laws in other locales, like the California Consumer Privacy Act (CCPA) and the Virginia Consumer Data Protection Act (VCDPA), operate on principles clearly inspired by GDPR. This will allow us to elucidate more general normative issues in defining anonymization using DLR.

While DP's conception of privacy risk and personal data is technical and narrow, mod ern privacy law's vision of personal data is more open-ended and nebulous. Current work in DLR has focused on recasting the properties of these technical definitions as desirable legal ones. For example, the text of the EU's GDPR defines "anonymized" data that is not subject to GDPR through three concepts: "singling out," in which an individual record could be reasonably isolated from others; "linkage," in which processed data could be asso ciated with external data through a matching process; and "inferences," in which processed data enables particular inferences about individual's sensitive data. While Working Article 29 has proposed multiple technologies which could hypothetically be used to successfully anonymize data [17], DLR broadly argues that a successful legal definition of anonymization ought to have a few key properties. First, anonymization should be a property of how data is processed, not specifically its output. Next, the ability to "single out" should be mea sured as a function of this process, namely how the process changes depending on what one individual contributes to the output. Finally, anonymized data should be "future-proof" in the sense that privacy guarantees should be maintained under further post-processing. DLR additionally provides a crucial negative result: because inferences about persons cannot be generically limited due to the fundamental law of information recovery, limiting inferences made about individuals ought not be codified in GDPR. This should all sound familiar to the reader, as all these proposals amount to conceptualizing anonymization in the same way that DP conceptualizes data privacy.

Just as there's a gap between DP as a formal mathematical system versus how DP 4 systems operate in practice, there are gaps between DLR as a set of legal principles and how DLR might play out as regulatory practice. These gaps dictate the scope of DLR on anonymization through the following unaddressed questions: which DP implementations satisfy the criteria for GDPR exemption? How does a data processor demonstrate GDPR exemption? How does a regulator audit a particular data processing task performed with DP? What role do data subjects play in having their data contributed to legally anonymized data processing systems? These are essential policy questions which must be addressed in the context of specific regulatory practices, yet each question is framed by the technical lens of DP in different ways that warrant our attention.

To be clear, the DLR literature largely and intentionally leaves these questions unan swered, and its architects are quick to acknowledge the limitations of proposing legal prin ciples derived from technical results. Few would endorse a worldview where DP's mere presence is the de facto standard for anonymization, just as few would advocate for DLR to technocratically determine the scope of privacy law. Despite these clearly good intentions, the stakes of interpreting DLR correctly are high, necessitating a clear understanding these framing effects and their consequences for effective privacy regulation. Should DLR prompt changes to GDPR and other regulations, data processors will have new incentives to use DP as a means to eschew regulatory interference. Such a scenario begs for critical attention to ensure DP is used properly in operationalizing and enforcing GDPR and other omnibus regulations.

## 3 DLR and Data Subjects

In this section, I discuss how DP's framing effects change the role of data subjects within data processing systems. Starting with the positive, DP improves upon previous anonymization techniques by making its characterization of privacy guarantees transparent and communi cable to end users. Such an approach sidesteps past methods colloquially known as "security by obscurity," where processors could not publicly disclose how data was anonymized. Ad ditionally, DP's measure of disclosure risks centers on the influence of individual user's data; by design, DP systems process data with randomized noise so that, with high probability, processed results with or without one individual's personal data are similar. This robustness to individual data contributions is a key selling point for DP: as advertised, a data subject's knowledge that a particular data collecting organization uses DP should assuage fears about the misuse of personal data. To make DP's particular measure of disclosure risk more salient to data subjects, DP offers numerous semantic interpretations which characterize aspects of a data subject's privacy risks (see [24] for an example based on government data prod ucts). Data subjects could use this information to distinguish which data processing tasks directly concern their personal information and which others only concern their contribution to some aggregate. If the way DP data processing systems are presented to data subjects helps them make more informed decisions, this would certainly be a net positive for data subjects. Whether this would happen in practice under DLR, however, is a much more difficult question to answer.

DP's underlying philosophy is closely tied to notice and consent, as its particular defini tion of privacy risk considers a counterfactual comparison between worlds where individuals either do or do not contribute their data to a data processing system. Moreover, the intervention by which data subjects choose to contribute to a data processing system is itself modeled as a causal intervention [45]. As a result, one could argue DP attempts to characterize privacy decisions as a collection of atomized "yes-no" decisions about individual's boundaries. Yet privacy scholars have long raised concerns about the efficacy of notice and consent, or any privacy self-management framework for that manner [40]. In our past work [35], we argued that DP exacerbates these existing problems as it requires data subjects to express their privacy preferences in the language of PLBs, which flattens the necessary contextual information needed to make informed decisions relative to their cultural and normative understanding of their privacy experiences. This should not be taken lightly, as social theories of privacy necessitate a shared understanding of social context [31, 32, 6]. Moreover, empirical research demonstrates that the way DP quantifies the aforementioned similarity between results with or without individual contributions is obtuse to the average person [9], as it pertains to a quantitative threshold that's difficult to generically express for any data processing system. So overall, DP can provide more information to end users about their privacy risks, but that information may be of limited use without additional context.

Introducing DLR into the picture could create new emergent effects in how data processors communicate with data subjects. First, and most importantly, DP could be counterintuitively applied for the purposes of hiding certain data processing tasks from data subjects. Most data processors use data subject contributions in multiple ways, only a subset of which accommodate DP. Under DLR, this could mean only a subset of data processing tasks would be subject to omnibus regulations that require data subject disclosure about data usage (for example, GDPR Article 29 and Recital 58). Although DP permits disclosing how data are processed, data processors may have no incentive to inform data subjects about these processing tasks if it benefits their ability to elicit consent from their subjects. This troublesome dynamic could serve as a form of "modulated consent" [8], wherein DLR becomes a way to obfuscate data processing operations that processors wish to hide from data subjects.

On top of this, the same personal data may be processed with or without DP in ways that cannot be sociologically siloed off from one another. Data subjects could easily approve of their personal data used for certain purposes and not others, but DLR could be used to prevent data subjects from making informed decisions about that distinction. In particular, data users may not consent to certain inferences being made about them, especially if those inferences are only enabled because data processors amass personal data [5]. Yet DLR, as motivated by impossibility results, could be wielded to say such claims are invalid. Many data processing tasks that raise broader inferential privacy concerns like personalization, recommendation systems, and record linkage can be performed using DP, but that alone does not address data subject's sociological privacy concerns. The purpose of these tasks is, ultimately, to place people into population-level "data relations" with one another [47], and it's precisely this relational aspect of data that cannot be captured by DP's individualistic notion of data privacy.

In addition to affecting what information data subjects see, DLR could also affect how such information is presented to data subjects. Communicating DP's privacy guarantees remains an active area of academic research; however, the privacy guarantees afforded by DP systems can be presented with wildly varying tones, affects, and other signifiers meant to influence data subjects' willingness to consent. These come into play because DP simultaneously uses broad assumptions about adversarial assumptions and narrow assumptions about how to measure disclosure risk. Emphasizing only the former might be too optimistic, bracketing how DP primarily limits how much data processors are liable for privacy harms. Conversely, emphasizing only the latter might be too pessimistic, as DP's ability to resist reconstruction should confer at least some substantive privacy guarantees to data subjects. Were there no legal constraints on how DP is communicated to data subjects under DLR, unbalanced language could be used to evoke consent.

All this suggests that DLR hinges on a particular kind of privacy self-management, one that could be strengthened or weakened based on how data processors communicate with data subjects. Nailing this communication requires a degree of neutrality that ought to be captured by third-party regulators in an effort to protect data subjects. Without such efforts, DLR could unintentionally exacerbate "hermeneutical injustices" present in compu tational law. First introduced by Fricker in [19], van den Hoven highlights how computational law could entrench injustice by"[depriving] the individual in question of (legally) significant knowledge" [46]. Individual hermeneutic injustices naturally arise when considering the role of the data subject in DP systems, as we've seen throughout this section; DLR as code-driven law dictates which information is legally significant from a regulatory perspective, but that same information may not be surfaced to data subjects. Even though DP outputs may be classified as beyond the scope of certain omnibus privacy laws, data subject willingness to participate still depends on comprehensive understanding of their data's use. Anonymized personal data occupies an uncomfortable middle ground in assessing the scope of privacy law, one that characterizes a processes that starts within scope and ends outside of scope. For better and worse, this process remains relevant to data subjects so long as omnibus privacy regulations rely on notice and consent.

## 4 DLR and Regulators

In this section, I discuss how DP systems change the regulatory calculus for agencies seeking to audit their data processing systems. First, I discuss the benefits of DP's particular unit of analysis from a regulatory perspective. The advent of large-scale data and computation has made addressing issues of accountability in data privacy exceedingly difficult. Vast networks of public and private data sources and mass markets connecting them make it difficult to trace the origin of any one privacy harm. When we take a broad definition of re-identifiability as suggested by the fundamental law of information recovery, DP's unit of analysis becomes a blessing. By understanding exactly what actions a data processor takes to protect data subject privacy and making that process transparent to regulators, DLR encourages a tech nical auditability that allows regulators to decide which data processing tasks warrant which tolerances for privacy loss. Ideally, this will help make data anonymization a more proactive process, in which data processors demonstrate ex ante protections for data subjects through DP's desirable properties.

To operationalize DLR as a regulatory practice, regulators will need to decide which applications of DP meet sufficient criteria to be exempt from classification as "personal data processing." Privacy loss budgets are the locus of negotiation for DP, therefore we first start by discussing the thresholding problem: how do we decide when a PLB is sufficiently small for satisfying a legal definition of anonymized data? Answering such a question is difficult to do generically, as a PLB budget in a vacuum tells us little about potential realized privacy harms. We already discussed the essential role of context in data subject's privacy loss understanding, and the same arguments apply to regulators aiming to determine these thresholds. Yet the question of "how much privacy loss is too much" is one DLR could inadvertently make harder to answer substantively.

First, DLR might impose new technical responsibilities and operational burden on regu lators who must interpret how DLR manifests in textual law through this threshold-setting process. Like with GDPR, many omnibus privacy laws require some process for setting what might be considered a "reasonable" threshold for disclosure risk. For example, regulators could establish "baseline" levels of disclosure risk and use these to translate DP's relative disclosure risks into absolute disclosure risks [34]. While this helps make PLBs more salient to regulators, it still requires establishing baseline levels of risk. DP's framing intentionally abstracts away adversarial modeling assumptions to avoid having to make such decisions, when in practice these decisions could be merely deferred to regulators when solving the thresholding problem. Regulatory agencies could likely lack the technical personnel neces sary to perform these precise quantitative analyses at scale, tailored to the numerous possible data processing contexts under scrutiny in these settings. This creates a vacuum of account ability for deciding who must perform the actuarial calculus of translating PLBs into shared understandings of privacy risk.

Even if this thresholding problem was hypothetically solved, regulators would still need to verify that certain data processing operations include DP protections. Verification of DP systems has been addressed in technical literature by developing different formal practices, such as DP data processing programming languages and black box verification systems [44]. While formal reasoning helps ensure the verification is correct, all these practices rely on the data processor exposing some part of their data processing system. For example, a data processor may need to establish a test system which allows regulators to see how hypothetical data might be processed in the future. In a more extreme case, data processors can make their code open source, allowing anyone to verify the correctness of their implementation. Yet many of the most prevalent data processing systems are highly proprietary, meaning such verification technologies may not be practical at the scale needed for omnibus legislation. Regardless of how data processors choose to demonstrate the presence of DP, it will exacerbate what's expected of regulators. We should not expect all those enforcing omnibus privacy laws to scour over the technical nitty-gritty of each data processing task. Therefore, DLR in practice will likely involve data processors relying on their own conceptions of the regulation to demonstrate their compliance.

Settings like these tee up an all-too-familiar dynamic in the U.S. tech industry: new technological tools can be best controlled and understood by their creators, motivating technical modes of self-regulation. When privacy regulations are guidelines-based or principles-based, data processors often take a surprisingly active role in determining the pragmatic aspects of privacy law compliance [4]. Such a process largely agrees with the culture of ethics in industry data settings [29], one that suggests that those closest to the technical processes should be the most qualified to ensure the self-regulation is working as anticipated. While this perspective does lessen the hypothetical operational burden placed on regulators, it allows DP data processors to demonstrate compliance on their own terms. Unfortunately, this performative, "box-checking" mentality has become a feature of U.S. privacy regulation. Edelman uses the term "legal endogeneity" to describe how data processors under regulatory scrutiny have undue influence over the regulatory process, largely by focusing on operational symbols of compliance instead of substantive privacy protections [15]. Through empirical research, Waldman demonstrates that legal endogeneity runs rampant in privacy law as a new class of compliance "professionals" largely dictate what privacy law means in substance through these performances [49]. Yet much like any privacy-enhancing technology, the quantification of disclosure risks could be set to the point where the mere presence of DP does little to nothing substantively to protect data subjects. For example, the privacy loss budgets set by Apple in their user analytics were large enough to raise serious concerns among DP scholars [43]. So while DLR improves the principles upon which privacy law rests, it could also exacerbate issues of legal endogeniety by unintentionally allowing DP data processors to dictate the terms on which their implementations satisfy legal requirements.

Should data processors gain more power in the process under DLR, we should expect similar communications issues with regulators as we saw with data subjects. As before, data processors can achieve their goals using combinations of DP and non-DP data processing, both of which are typically intertwined but only some of which may be subject to omnibus privacy regulations under DLR. Therefore the scope, goals, and methods used by DP systems may be relevant to assessing contextual risk but not available for regulators to understand. Suppose, as an example, a data processor used personal data to train a DP model which predicts sensitive attributes about individuals based on their non-DP contributions to the database. If regulators only have visibility into the non-DP component of the data processing task, then they may make more lax decisions about downstream privacy risks than otherwise, had they known what the DP data was used for in context. This could exacerbate enforcement issues by legitimizing or delegitimizing different kinds of data processing disclosures, a process described by van den Hoven as a "systemic" hermeneutical injustice in computational law [46].

It's important to recognize that ineffective regulatory enforcement itself is not specific to DP, nor any privacy-enhancing technology; the fact that DP could be weakly regulated is not itself a technological shortcoming for DP. Instead, the way DLR separates concerns in framing privacy risks could unintentionally enable the kinds of regulatory arbitrages described in this section. Characterizing privacy risks as inherent to a process helps to enable data processing at scale, but decoupling privacy and surveillance concerns leaves questions of power and justice wide open [28]. Operationalizing DLR by setting baseline risks brackets how privacy norms evolve over time, often in ways which merely normalize privacy harms [37, 42]. Focusing on relative changes in privacy loss helps isolate organizational impacts, but doing so brackets the compounding effects of individual actions on data privacy as a collective problem [22]. In short, DP's abstraction requires a mathematical "separation of concerns," that must be cast aside when addressing privacy as a complex, systemic issue whose facets cannot be siloed off from one another.

## 5 Conclusion

To summarize, treating DP as a legal apparatus shifts multiple loci of attention within privacy law. These shifts offer a mixture of potentials for more rigorous protections and op portunities for data curators to eschew substantive accountability. Examining these framings at the intersection of law and computer science allows us to delineate where new interven tions are needed for regulating DP systems while still harnessing the power of DP's privacy guarantees to protect data subjects from potential harms. To conclude, I'll make a few key recommendations.

First and foremost, the sociotechnical complexity of DP data processing outputs suggests that omnibus privacy laws should not treat these outputs as exclusively within or not within scope. Regulations like GDPR attempt to achieve multiple intertwined goals simultaneously, each of which has different degrees of substantive relevance for DP data processing systems. Greater flexibility in treating DP outputs would help leverage the heterogeneity and flexibil ity of text-driven law while more precisely articulating what DP systems can or cannot do to protect data subject privacy. As an example, data subjects can benefit from mandated notice about the use of DP systems, and such notice needs to be transparent about how DP reduces, but does not eliminate, privacy harm. Despite the many failings of privacy self management, effective notice remains a key ingredient of successful data subject governance [41]. Interventions like these could help privacy law capture more nuanced relationships between data collectors and data processors.

Second, just as we should not binarize DP outputs with respect to privacy law, we should give pause to entirely divorcing inferences from legal definitions of data privacy. Despite the fundamental law of information recovery, certain inferences may be enabled largely because of surveillance at scale. The way that DP decouples privacy and surveillance becomes problematic when surveillance enables inferences that data subjects may contest; some privacy scholars have argued that this ought to be codified as a "right to reasonable inferences" [48]. Such a limitation on inferences may not be formally quantifiable as inherent to how data is processed, but formal processes do not prevent empirical demonstrations of invasive inferences only made possible by amassing personal data. This topic in particular will be investigated in future work.

To reiterate, DLR makes important contributions to our understanding of privacy law through a technical lens, delimiting which aspects of privacy law can be quantitatively ad dressed and which cannot under the formal reasoning of DP. I've argued that this formaliza tion has many potential benefits and detriments depending on how the gap between DLR and regulatory practice closes, and co-production of technical and legal concepts remains useful for effective data subject protections.

## References

[1] Madeleine Akrich. The de-scription of technical objects, 1992.

[2] Micah Altman, Aloni Cohen, Evangelia Anna Markatou, Francesca Falzon, Kobbi Nis sim, Michel Jos´e Reymond, Sidhant Saraogi, and Alexandra Wood. A principled ap proach to defining anonymization as applied to EU data protection law. *Francesca and Nissim, Kobbi and Reymond, Michel Jose and Saraogi, Sidhant and Wood, Alexandra, A principled approach to defining anonymization as applied to EU data protection law (May 9, 2022), 2022.*

[3] Micah Altman, Aloni Cohen, Kobbi Nissim, and Alexandra Wood. What a hybrid legal technical

analysis teaches us about privacy regulation: The case of singling out. *BUJ Sci. & Tech. L.*, 27:1, 2021. Publisher: HeinOnline.

[4] Kenneth A Bamberger and Deirdre K Mulligan. *Privacy on the ground: driving corpo rate behavior in the United States and Europe*. MIT Press, 2015.

[5] Solon Barocas and Helen Nissenbaum. Big data's end run around anonymity and con sent. *Privacy, big data, and the public good: Frameworks for engagement*, 1:44–75, 2014. Publisher: Cambridge University Press, NY.

[6] Sebastian Benthall, Seda G¨urses, Helen Nissenbaum, and others. *Contextual integrity through the lens of computer science*. Now Publishers, 2017.

[7] Aloni Cohen and Kobbi Nissim. Towards formalizing the GDPR's notion of singling out. *Proceedings of the National Academy of Sciences*, 117(15):8344–8352, 2020. Publisher: National Acad Sciences.

[8] Julie E Cohen. What privacy is for. *Harv. L. Rev.*, 126:1904, 2012. Publisher: HeinOn line.

[9] Rachel Cummings, Gabriel Kaptchuk, and Elissa M Redmiles. " I need a better descrip tion": An Investigation Into User Expectations For Differential Privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, pages 3037–3052, 2021.

[10] Damien Desfontaines and Bal´azs Pej´o. Sok: differential privacies. *arXiv preprint arXiv:1906.01337*, 2019.

[11] Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *Proceed ings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 202–210, 2003.

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

[13] Cynthia Dwork and Moni Naor. On the Difficulties of Disclosure Prevention in Statistical Databases or The Case for Differential Privacy. *Journal of Privacy and Confidentiality*, 2(1):93–107, 2010.

[14] Cynthia Dwork, Aaron Roth, and others. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

[15] Lauren B Edelman. Working Law. In *Working Law*. University of Chicago Press, 2016. 11

[16] European Union. General Data Protection Regulation, 2016.

[17] European Union Working Party. On the Protection of Individuals With Regard to the Processing of Personal Data, 2014.

[18] Mich`ele Finck and Frank Pallas. They who must not be identified—distinguishing personal from non-personal data under the GDPR. *International Data Privacy Law*, 2020.

[19] Miranda Fricker. Epistemic justice as a condition of political freedom? *Synthese*, 190(7):1317–1332, 2013. Publisher: Springer.

[20] Ben Green. Escaping the" Impossibility of Fairness": From Formal to Substantive Algorithmic Fairness. *arXiv preprint arXiv:2107.04642*, 2021.

[21] Ben Green and Salom´e Viljoen. Algorithmic realism: expanding the boundaries of algorithmic thought. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 19–31, 2020.

[22] Deborah Hellman. Big data and compounding injustice. *Journal of Moral Philosophy, forthcoming, Virginia Public Law and Legal Theory Research Paper*, (2021-27), 2021.

[23] Mireille Hildebrandt. Code Driven Law. Scaling the Past and Freezing the Future. *Criti cal Perspectives on Law and Artificial Intelligence, eds. Markou, Deakin, Hart Publishers (2020 Forthcoming)*, 2020.

[24] Daniel Kifer, John M Abowd, Robert Ashmead, Ryan Cumings-Menon, Philip Leclerc, Ashwin Machanavajjhala, William Sexton, and Pavel Zhuravlev. Bayesian and Frequen tist Semantics for Common Variations of Differential Privacy: Applications to the 2020 Census. *arXiv preprint arXiv:2209.03310*, 2022.

[25] Daniel Kifer and Bing-Rong Lin. Towards an axiomatization of statistical privacy and utility. In *Proceedings of the twenty-ninth ACM SIGMOD-SIGACT-SIGART sympo sium on Principles of database systems*, pages 147–158, 2010.

[26] Daniel Kifer and Ashwin Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pages 193–204, 2011.

[27] Ashwin Machanavajjhala, Xi He, and Michael Hay. Differential privacy in the wild: A tutorial on current practices & open challenges. In *Proceedings of the 2017 ACM International Conference on Management of Data*, pages 1727–1730, 2017.

[28] Alice Marwick. Privacy without Power: What Privacy Research Can Learn from Surveil lance Studies. *Surveillance & Society*, 20(4):397–405, 2022.

[29] Jacob Metcalf, Emanuel Moss, and others. Owning ethics: Corporate logics, silicon val ley, and the institutionalization of ethics. *Social Research: An International Quarterly*, 86(2):449–476, 2019. Publisher: Johns Hopkins University Press.

[30] Gerome Miklau. Negotiating Privacy/Utility Trade-Offs under Differential Privacy. 2022.

[31] Helen Nissenbaum. *Privacy in context*. Stanford University Press, 2009.

[32] Helen Nissenbaum. Differential Privacy in Context: Conceptual and Ethical Consider ations. In *Four Facets of Differential Privacy Symposium, Princeton, NJ, USA*, 2016.

[33] Paul Ohm. Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA l. Rev.*, 57:1701, 2009. Publisher: HeinOnline.

[34] Nathan Reitinger and Amol Deshpande. E-Differential Privacy, and a Two Step. 2023.

[35] Jeremy Seeman and Daniel Susser. Between Privacy and Utility: On Differential Privacy in Theory and Practice. *Available at SSRN 4283836*, 2022.

[36] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 59–68, 2019.

[37] Evan Selinger and Hyo Joo Rhee. Normalizing surveillance. *Sats*, 22(1):49–74, 2021. Publisher: De Gruyter.

[38] Aleksandra Slavkovi´c and Jeremy Seeman. Statistical data privacy: A song of privacy and utility. *Annual Review of Statistics and Its Application*, 10, 2023. Publisher: Annual Reviews.

[39] Daniel J Solove. A taxonomy of privacy. *U. Pa. l. Rev.*, 154:477, 2005. Publisher: HeinOnline.

[40] Daniel J Solove. Introduction: Privacy self-management and the consent dilemma. *Harv. L. Rev.*, 126:1880, 2012. Publisher: HeinOnline.

[41] Daniel Susser. Notice After Notice-and-Consent: Why Privacy Disclosures Are Valuable Even If Consent Frameworks Aren't. *Journal of Information Policy*, 9(1):148–173, 2019. Publisher: Pennsylvania State University Press.

[42] Daniel Susser. Decision Time: Normative Dimensions of Algorithmic Speed. 2022.

[43] Jun Tang, Aleksandra Korolova, Xiaolong Bai, Xueqiang Wang, and Xiaofeng Wang. Privacy loss in apple's implementation of differential privacy on macos 10.12. *arXiv preprint arXiv:1709.02753*, 2017.

[44] Michael Carl Tschantz, Dilsun Kaynar, and Anupam Datta. Formal verification of differential privacy for interactive systems. *Electronic Notes in Theoretical Computer Science*, 276:61–79, 2011. Publisher: Elsevier.

[45] Michael Carl Tschantz, Shayak Sen, and Anupam Datta. Differential privacy as a causal property. *arXiv preprint arXiv:1710.05899*, 2017.

[46] Emilie Van Den Hoven. Hermeneutical injustice and the computational turn in law. *Journal of Cross-disciplinary Research in Computational Law*, 1(1), 2022.

[47] Salome Viljoen. A relational theory of data governance. *Yale LJ*, 131:573, 2021. Pub lisher: HeinOnline.

[48] Sandra Wachter and Brent Mittelstadt. A right to reasonable inferences: re-thinking data protection law in the age of big data and AI. *Colum. Bus. L. Rev.*, page 494, 2019. Publisher: HeinOnline.

[49] Ari Ezra Waldman. Privacy Law's False Promise. *Wash. UL Rev.*, 97:773, 2019. Pub lisher: HeinOnline.

[50] Langdon Winner. Do artifacts have politics? *Daedalus*, pages 121–136, 1980. Publisher: JSTOR.

[51] Alexandra Wood, Micah Altman, Aaron Bembenek, Mark Bun, Marco Gaboardi, James Honaker, Kobbi Nissim, David R O'Brien, Thomas Steinke, and Salil Vadhan. Differen tial privacy: A primer for a non-technical audience. *Vand. J. Ent. & Tech. L.*, 21:209, 2018. Publisher: HeinOnline.