

Algorithmic Transparency and Participation through the Handoff Lens: Lessons Learned from the U.S. Census Bureau’s Adoption of Differential Privacy

Amina A. Abdu
aabdu@umich.edu
University of Michigan
Ann Arbor, Michigan, USA

Deirdre K. Mulligan
dmulligan@berkeley.edu
University of California, Berkeley
Berkeley, California, USA

Lauren M. Chambers
lauren@ischool.berkeley.edu
University of California, Berkeley
Berkeley, California, USA

Abigail Z. Jacobs
azjacobs@umich.edu
University of Michigan
Ann Arbor, Michigan, USA

ABSTRACT

Emerging discussions on the responsible government use of algorithmic technologies propose transparency and public participation as key mechanisms for preserving accountability and trust. But in practice, the adoption and use of any technology shifts the social, organizational, and political context in which it is embedded. Therefore translating transparency and participation efforts into meaningful, effective accountability must take into account these shifts. We adopt two theoretical frames, Mulligan and Nissenbaum’s *hand-off model* and Star and Griesemer’s *boundary objects*, to reveal such shifts during the U.S. Census Bureau’s adoption of differential privacy (DP) in its updated disclosure avoidance system (DAS) for the 2020 census. This update preserved (and arguably strengthened) the confidentiality protections that the Bureau is mandated to uphold, and the Bureau engaged in a range of activities to facilitate public understanding of and participation in the system design process. Using publicly available documents concerning the Census’ implementation of DP, this case study seeks to expand our understanding of how technical shifts implicate values, how such shifts can afford (or fail to afford) greater transparency and participation in system design, and the importance of localized expertise throughout. We present three lessons from this case study toward grounding understandings of algorithmic transparency and participation: (1) efforts towards transparency and participation in algorithmic governance must center values and policy decisions, not just technical design decisions; (2) the handoff model is a useful tool for revealing how such values may be cloaked beneath technical decisions; and (3) boundary objects alone cannot bridge distant communities without trusted experts traveling alongside to broker their adoption.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

FACCT ’24, June 03–06, 2024, Rio de Janeiro, Brazil

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0450-5/24/06

<https://doi.org/10.1145/3630106.3658962>

CCS CONCEPTS

• **Social and professional topics** → **Government technology policy**; • **Security and privacy** → **Human and societal aspects of security and privacy**; • **Applied computing** → *Law, social and behavioral sciences*.

KEYWORDS

critical transparency studies, participatory AI, differential privacy, census

ACM Reference Format:

Amina A. Abdu, Lauren M. Chambers, Deirdre K. Mulligan, and Abigail Z. Jacobs. 2024. Algorithmic Transparency and Participation through the Handoff Lens: Lessons Learned from the U.S. Census Bureau’s Adoption of Differential Privacy. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FACCT ’24)*, June 03–06, 2024, Rio de Janeiro, Brazil. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3630106.3658962>

1 INTRODUCTION

Recent work on values in technology attempts to understand how technological changes can produce fairer, more accountable, and more trustworthy systems [38, 78, 110]. Transparency and participatory design are often proposed to advance these goals in both academic work and policy [16, 26, 27, 44, 67]. However, scholars, critics, and advocates have raised complications and limitations of transparency and participation, particularly when adopted uncritically [13, 30, 33, 112]. We argue that prevailing models of algorithmic transparency and participation stand to benefit from sociotechnical analysis of transparency and participation on-the-ground. As such, we focus on a single case study: the adoption of differential privacy in the 2020 Decennial Census. Differential privacy (DP) is a mathematical definition of privacy that leverages statistical uncertainty to provably limit leakage of any individual’s sensitive information— in other words, random noise is added to data in order to reduce the possibility for re-identification [43]. Under the leadership of Chief Scientist John Abowd, the Census Bureau implemented DP in its 2020 disclosure avoidance system (DAS), the

The content herein represents the personal views of the authors and is not intended to reflect the views of the United States Government or any Federal agency.

mechanism used to manipulate census response data prior to publication to ensure confidentiality in accordance with Title 13.¹ Notably, the technical affordances of DP allowed the Bureau to make details of the DAS public for the first time without undermining confidentiality. The Bureau embraced this possibility, introducing many innovations in transparency and attempting to facilitate participation from a wide variety of experts and the public.

Despite these significant efforts, this newfound transparency did not produce the accountability and trust the Bureau hoped to engender [59, 71, 95, 97]. The resulting controversy attracted the attention of critical scholars, who have attempted to adjudicate its history and implications [18, 19, 92]. We build upon this literature; by employing the *handoff lens* [90], we parse out how a seemingly technological transition - from its previous statistical disclosure (SDL) methods to DP - in fact altered the very function of disclosure avoidance as the Bureau's methods, experts, and values were reconfigured.

Drawing on this case study, we demonstrate the utility of the handoff model for addressing calls from the critical algorithmic transparency literature to examine transparency in context. In particular, we show that the handoff model makes visible *where* decisions about values are embedded within sociotechnical systems and identifies the configurations of human actors surrounding these decisions.

Our contribution is three-fold. First, we shed new insight on a case that has been of significant interest to both researchers and policymakers. We provide an account of the specific values decisions at the core of the adoption of DP and how participation was limited in these decisions despite significant efforts. Second, we argue that values should be at the center of transparency and participation efforts and demonstrate the utility of the handoff model for eliciting these values. Finally, we highlight the need for understanding the role of experts in transparency and participation processes. While the literature has focused on developing documentation artifacts for transparency and participation, the census case highlights the insufficiency of artifacts alone to facilitate meaningful participation. Trusted individuals with requisite expertise must exist, within stakeholder communities, in order for such artifacts to be understood and adopted.

2 RELATED WORK

Calls to make technical systems more trustworthy and accountable point to *transparency* and *participation* as key interventions. Such calls are common across academic [25, 28, 38, 68, 133], industry [34], and governmental and NGO-based initiatives [16, 27, 48, 98, 99]. Transparency efforts have called for visibility around choices about data, processes, or mathematical properties of algorithms [67], e.g., through documentation of the development process [39, 61, 102] or structured disclosures of data properties [51, 57, 88]. Other efforts have been made to release code, make 'inscrutable' algorithms interpretable or their decisions explainable [25, 44, 82].

Yet revealing such inner workings can fail to live up to promises of participation, contestability, or trust. Researchers have consistently cautioned that algorithmic transparency efforts may not

bring the benefits they promise [13, 21, 30, 76, 78]. Making inscrutable algorithms explainable or readily available is insufficient without accounting for the social, organizational, and political power structures that shape their outcomes [15, 44, 74, 76, 81]. Intermediate objects for transparency efforts, such as structured disclosure documents, prioritize and shift ownership, effort, expertise and values [136]. These efforts offer visibility without necessarily offering meaningful paths to accountability for public administration [26, 67] or governance more broadly [30, 31].

Accountability therefore relies on enabling substantive participation and contestation [28, 31, 67, 73, 129]. However, approaches to participation are complicated by the technical features of algorithmic systems and the broader social, political, and organizational contexts in which they operate [66, 80, 104]. Participation efforts - whether participatory design activities, communicating with a wider range of stakeholders, or creating structured opportunities for feedback - can then stand in for meaningful stakeholder empowerment. Indeed, critiques have shown how participation efforts can be exploitative and performative [31, 33, 100, 112], where incentive structures can encourage "participation-washing" [17, 29, 32, 100].

Contextual analysis is needed to support meaningful participation in (algorithmic) governance. Tools to articulate responsibility and decision-making can be disconnected from what practitioners need [83, 114, 131], but articulating the decisions being made in the implementation of systems can better reveal where decision-making is happening in the first place [61, 62, 75, 84, 113]. Transparency is needed not just of design choices, but also of the policy questions these choices often seek to answer [89]. Past literature has pointed to how the introduction or substitution of technologies can reconfigure values and social and political arrangements, and that understanding this context shift is key to responsible technology development [11, 79, 90, 110]. This missing lens on what happens when we substitute a new technology [83, 110], and how that shifts organizational roles [86, 111, 136] with technical and political impacts [e.g., 63, 90], must then be considered to support meaningful participation. A fundamental challenge is to reveal how values and decisions change as a result of the introduction or substitution of a new technology: the handoff model that we discuss in the following section intervenes on this specific challenge.

3 THEORETICAL LENSES

3.1 Handoff Model

Mulligan & Nissenbaum's handoff model illuminates the values embedded in seemingly technical changes within a sociotechnical system, in particular when one part of a system is replaced with another [91]. By surfacing often overlooked reconfigurations of political and social relationships between technologies and people, the handoff model challenges narratives of linear technological progress. The handoff model defines five elements of a system to interrogate in order to expose the changing values that it embodies: the *functions* of a system and sub-functions of its components (what does it do?); the *components* that are involved, be they technical or human (what are its pieces?); the ways in which components *engage with* or *act on* one another (where do its pieces connect?);

¹The Census Bureau is obligated to protect the privacy and confidentiality of individual data according to Title 13 of the U.S. Code

the *modes of action* by which one component acts on or engages another (in what ways do its pieces connect?); and finally the *trigger* that spurs the handoff (why did it change?).

The handoff model has previously been adopted to understand shifting values in a variety of sociotechnical systems, including the adoption of design practices within organizations [135] and shifts in technology design. For instance, Mulligan and Nissenbaum use the model to reveal how changes in access control for mobile phones (introducing passcodes and, later, biometric verification) shifted the roles of developers and users as well as the space of privacy threats [91].

In this paper, we apply the handoff model to examine the shift from SDL to DP. The function is ostensibly the same (confidentiality) but the handoff reveals the implicitly value-laden rearrangements of organizations and design decisions.

3.2 Boundary objects

Star and Griesemer introduce the concept of boundary objects as arrangements that allow groups to work together without consensus [116]. Star theorizes boundary objects as having three important characteristics: (1) interpretive flexibility, (2) arising from information needs that are (3) weakly structured in common use but strongly structured in local contexts [115]. Past work on boundary objects in organizations has, for instance, shown that these intermediate arrangements can serve as important tools for learning and communication across diverse stakeholders for difficult or heterogeneous domains.

Past work demonstrates that the benefit of boundary objects also depends on their design, which can highlight gaps in understanding and values across groups or failures to serve different stakeholder needs [23]. Further, the role of expertise is crucial: the design and use of boundary objects unintentionally (and intentionally) shapes stakeholder engagement [72, 105]. Within the handoff model, boundary objects can help us to understand the changing relationships between different actants in a handoff. In particular, we examine how the Bureau shaped participation through the artifacts that it introduced to structure negotiations. By attending to these sites of negotiation within the Census Bureau's adoption of DP, our goal is to illuminate how differences in values and expertise affect sociotechnical handoffs with consequences for participation, trust and accountability.

4 APPLYING THE THEORETICAL LENSES

In this section, we employ the handoff model to highlight the reconstructions which took place during the Census Bureau's change to a new disclosure avoidance system (DAS) as they adopted DP. Specifically, we analyze changes in the elements which relate directly to the DAS's primary function: *protecting the confidentiality of census responses*. In addition, we consider the artifacts that the Bureau introduced as part of the DP implementation process.

4.1 Handoff Triggers: New tech, new threats, new hype

From Census Bureau documents and communications [e.g., 5–7, 10], we identify three triggers that spurred the handoff, i.e., the adoption of DP. The first trigger enabled the handoff: the *development of DP*

in 2006 presented a promising new method [43]. Second, the Bureau became aware of mounting evidence that increased computational power and data access might lead to new *threats to privacy and confidentiality* [7, 69, 128]. These threats focused on future, potentially dangerous threats via reconstruction and/or re-identification attacks [40] as well as current, realizable threats. For instance, Latanya Sweeney showed in 2000 that almost 90 percent of respondents to the 1990 U.S. Census could be identified using zip code, birth date, and gender alone [119]. In response to potential future threats, the Bureau conducted experimental attacks on its own statistical releases. While the experimental details could not be made public, the Bureau claimed their attack demonstrated that the re-identification of census records was indeed a credible threat [7, 56]. Third, the social environment served as another trigger: specifically the *'tech for good' hype*. A heightened cultural interest in framing and 'solving' policy problems using stylized computational methods preceded the Bureau's decision to adopt DP. This interest was evidenced in part by the rise of programs such as Code for America in 2009, the United States Digital Service in 2014, and the Mechanism Design for Social Good initiative in 2016 [53], as well as by technical communities' heightened attention to questions of fairness, accountability, transparency, and ethics of technology, from the FATML origins in 2014 to the establishment of ACM venues such as FAccT (previous FAT*) and AIES in 2018 [2, 14]. DP saw increased in tandem with this hype, promising to encode values such as transparency, accountability, and privacy into a mathematical formulation [108, 138]. In such an environment, the adoption of a theoretical computer science methodology to bolster privacy protection in the national Census - and a methodology which would increase transparency into the Bureau's processes - aligned perfectly with prevailing trends.

4.2 Handoff Components: Shifting experts, techniques, and data

The sociotechnical ecosystem surrounding the handoff of the DAS has many *actor-components*, many of which remained largely unchanged throughout the transition to DP. For instance, the stakeholders and users who depend upon Census data products and the external agencies and groups (demographers, community groups, etc.) with whom the Bureau collaborates remained relatively stable throughout the handoff. However, the handoff also introduced new components, shifting the experts and technologies involved in delivering the DAS's confidentiality function. Below we compared those DAS components before and after the shift from SDL to DP.

4.2.1 Technical Methods. A key shift in the DAS was the substitution of SDL tools with DP mechanisms, a new set of confidentiality-preserving tools built on a definition of privacy from theoretical computer science. Under the previous SDL methods, Bureau statisticians protected census response data through methods like suppressing and swapping individual records. Under DP, however, randomly generated "noise" is algorithmically added to census data to preserve confidentiality. This transition introduced two new sub-components of particular note. First, the tunable *epsilon* (ϵ) parameter is a direct measure of privacy loss in DP, in which small ϵ reflects low privacy loss (i.e., high privacy, low accuracy) while large

ϵ reflects high privacy loss (i.e., low privacy, high accuracy). Second is *post-processing*, a new step added to the data pipeline under DP to further modify the confidential data after the injection of randomized noise. This step ensures all final data products are non-negative integers, in order to assuage human interpreters who might be confused or put off by, for instance, a table reporting -48.12 people residing within a particular geography.

4.2.2 Data Invariants. The Census's outputs have significant consequences: because certain statistics inform resource allocation, accurate representation of population is important to a number of stakeholders, including voting rights advocates, state and municipal governments, tribal leadership, and even disaster recovery and public health personnel [95, 97, 137]. For the 2010 census and those prior, some particularly significant counts (such as total state populations) were held *invariant* under the DAS;² in other words, they were not manipulated from their value 'as counted' [93, 94]. However, invariants are incompatible with traditional DP – zero noise requires an infinite privacy budget – meaning that any count held invariant complicates DP's confidentiality guarantees [3, 93].³ As a result, the 2020 DAS reduced the number of counts that would be held invariant, most notably no longer publishing the population of census blocks as counted. Thus, the reported population - as well as demographic characteristics such as race and age - of all geographies smaller than a state would be altered with DP before publication.

4.2.3 Experts. Finally, the adoption of DP thus meant the introduction of a new class of experts to the DAS: theoretical computer scientists, specifically those well-studied in DP formulations. As a result, computer scientists were slotted into DAS design processes, for instance: serving alongside social scientists, policy researchers, political advocates, and corporate leaders on DP working groups for two census oversight committees; and completing contractual work to directly assist in implementing DP for the 2020 Demographic and Housing Characteristics tabulations [55].

4.3 Handoff Modes: Abstraction and constrained expertise

The handoff model pays particular attention to differences between *modes* of interaction between components. Attending to how these modes shift with the introduction of DP reveals how underlying values shift as well.

Some of the ways the Bureau manipulated response data remained fundamentally the same for the 2020 census, such as using imputation to compensate for non-responses [22]. However, the introduction of DP changes the mode in which the *DAS acts on the Census response data* in order to protect respondents' confidentiality, shifting the roles of *expert decision-making*. Under the Bureau's

²Specifically, we know that "total population, voting-age population, number of housing units, number of occupied housing units, and number and type of group quarters were all invariants at the block level in 2000 and 2010 Census publications" [4]. Blocks are the smallest unit of geography recorded by the Census [7] In 2010, for instance, the Census divided the country into over 11 million blocks.

³We note that Bureau messaging regarding invariants is inconsistent. Officials have said in some instances that invariants can be reconciled with DP but that they "eat[...] the privacy-loss budget" [93, p. 27], but in others that invariants "fundamentally violate[...] the central promise of differentially private solutions to controlling disclosure risk" [6, p. 32]. As such, it is unclear whether invariant values should or should not be considered within the scope of the DAS.

prior SDL methods, Bureau experts researched, debated, and determined the process for matching and swapping individual 'micro-data' records and for suppressing tables [85]. With the change to DP, however, the numerous swapping, rounding, or suppression decisions that a given DAS researcher might make regarding the particulars of disclosure avoidance are condensed down to fewer, more abstracted decisions - for instance, determining the value of the privacy budget ϵ and the allocation of that privacy budget across the dataset.⁴

The changes to the DAS also reconfigured the relevance of *disciplinary expertise*. Many kinds of pre-DP census expertise (such as that of demographers and political advocates) was no longer sufficient to afford a confident understanding of, or even engagement with, how the DAS operates [95, 97]. This is to say, the responsibility to effectively design privacy protections was displaced from statisticians and the DAS's prior experts onto computer scientists. This reorientation leads to not only a new balance of power across the landscape of Census experts, but also a fundamental shift in the rhetorical and epistemological configuration of the Census and the DAS [19, 92].

Finally, the primary mode of *securing confidentiality* transformed under DP. Under prior SDL methods, revealing the details of DAS methods would entirely undermine their effectiveness by allowing a motivated attacker to reverse strategies and reconstruct the unprocessed data. Consequently, SDL methods provided confidentiality via maintaining secrecy and withholding details about the DAS from the public – an approach commonly referred to as "security by obscurity." Thus the degree of confidentiality which SDL provided could only be determined by ex-post evaluation by internal bureau experts, and it was not future-proofed. With DP, however, confidentiality becomes mathematically provable, ex-ante, under the guarantees provided by statistical uncertainty. Secrecy and public withholding are no longer a concern. What's more, this reconfiguration also shifts how - and even if - experts evaluate confidentiality. While SDL required expert judgement in order to evaluate how much confidentiality protection a particular configuration of the DAS afforded, DP has displaced that judgement onto a statistical guarantee. Experts are no longer required to assess how effective the DAS is at preserving confidentiality; DP does it for them via epsilon (ϵ), the privacy budget.

4.4 Handoff Function: Interrogating the *how*

At first blush, it seems that the disclosure avoidance system's primary *function* remained the same before and after DP: preserving the confidentiality of census responses. However, the handoff model encourages a richer understanding of a system's function, including not only its "goals, purposes, or [...] values" but also "*how* it does what it does, as a designer or engineer might explain it" [91, p. 6, original emphasis]. The handoff lens ultimately reveals that DP, by shifting *how* the function of confidentiality preservation was enacted, expanded the function of the DAS in many value-laden ways.

⁴To be sure, there are still many fundamental choices that remain for designers of the DAS under DP. Researchers must draw block and tract boundaries, determine which data will be kept invariant, as well as which kinds of geographies will and will not include in the hierarchy known as the 'spine,' to name a few [7, 94, 124]. Nevertheless, epsilon-DP greatly reduces the dimensionality of *disclosure avoidance decisions* left to Bureau experts.

In particular, DP (1) created new opportunities for transparency between the Bureau and interested publics; (2) allowed for formal, quantifiable validation of the privacy and confidentiality commitments actualized by the Bureau; and (3) replaced one form of expertise with another, precipitating the rise of theoretical computer science professionals and the decline of statisticians in the design, operation, evaluation of disclosure avoidance. We explore these shifting functions in greater detail throughout Section 5.

4.5 Transparency artifacts at the boundaries: Spaghetti at the wall

To seize the benefits of the transparency that DP allows, and to enable stakeholder participation in the DAS design, the Bureau created many new artifacts to facilitate public understanding and input of the DAS. First, the Bureau released an unprecedented degree of technical detail, sharing the DAS *source code* via GitHub [54]. Realizing that code was not sufficient for providing transparency given stakeholder capacity, the Bureau released *demonstration data* that would allow demographers and social scientists who use census products to interact with the new system in a way that was familiar to them. Ultimately releasing six sets of data between 2018 and 2021, these public datasets were the result of applying the Bureau's 2020 DP algorithm to data from the 2010 Census.

The Bureau also engaged external experts in formal and informal *co-design processes*. Specifically the Bureau solicited written comments from data users via the Federal Register [120]; encouraged user feedback after publishing each round of demonstration data; hosted and participated in workshops devoted to discussing the use of DP [45, 93]; and held multiple consultations with tribal leadership [121]. Reaching beyond those experts, the Bureau provided an impressive array of *educational resources* designed for more diverse stakeholders and the interested public about the new DAS. Blogs narrating the Bureau's plans and progress as they worked to implement DP were authored by the Census' chief scientist himself [5, 9, 10]. To build up stakeholders' understanding of what DP is and why it is a worthwhile tool, the Bureau developed interactive Python Jupyter notebooks [36], webinars [127], handbooks [124], and videos [87] designed for a lay audience.

However, while the bureau created many artifacts and processes to bolster the public's understanding and participation in the DAS design, they withheld one artifact that external DP experts needed to evaluate the DAS. *Noisy measurement files* are an interim data product which contain the census data after the application of DP, but before post-processing removes negative or non-integer values. These files were not originally released by the Bureau but became an object of great interest. In 2021, a group of over 50 researchers, technologists, and city officials wrote to the Bureau requesting publication of the noisy measurement file, arguing that the release of this noisy data would expedite evaluations of the downstream effects of DP while still adhering to Title 13 privacy requirements [41]. Initially, the Bureau denied a FOIA request to release this data, citing concerns about confusing the public by revealing the existence of more than "one 'true' data set" [19, p. 15]; researchers were further frustrated [60]. Yet a year and a half later, in April 2023, the Bureau did release noisy measurements of 2010 demonstration data [126].

While the transparency artifacts listed in this section originated in the Bureau's interest in helping engage, educate, and involve diverse stakeholder groups; they ultimately ended up being sites of active negotiation within the handoff to DP. This reflects an implicit understanding within the Bureau of the need for boundary objects [116] to bridge the many stakeholder groups they hoped to engage in the design of the DAS. We do not intend to interrogate the degree to which individual artifacts were or were not successful in acting as boundary objects. Rather, we present and emphasize the *volume and variety* of artifacts which the Bureau developed to facilitate stakeholder participation. Building upon this foundation, in Section 5 we will evaluate the artifacts' overall effectiveness in *doing boundary work* [77] – that is, allowing a variety of stakeholders to participate in and negotiate the handoff from SDL to DP.

5 UNCOVERING THE STAKES OF THE HANDOFF

Applying the handoff model to our case as outlined in section 4, we surface several value-laden shifts in the Census Bureau's adoption and implementation of DP, as well as the participatory processes that the Bureau introduced to negotiate this transition. We summarize these findings in Table 1.

5.1 Confidentiality is the tip of the iceberg

Switching from earlier SDL methods to DP, the DAS maintained the same function of protecting Census respondents' confidentiality. The Bureau emphasized this change as a narrative of progress and increased effectiveness: the Bureau framed DP as a "modern" alternative to prior methods and "a new, advanced, and far more powerful confidentiality protection system" [5]. The handoff model allows us to look beyond this narrative of linear progress to understand the larger social and political implications of the new DAS. While the DAS's *function* remained focused on confidentiality, examining the changing *components and modes* used to achieve this function reveals a more complicated story. First, **the notion of confidentiality is itself contested**. The Bureau's decision to pursue confidentiality through *disclosure avoidance* is itself a value-laden choice, shaped by its interpretation of the confidentiality requirement outlined in Title 13. The handoff model demonstrates that changing one part of the DAS is not merely a modular replacement of one technical component with another. Instead, the adoption of DP changed the meaning of the system's core confidentiality function by shifting what harms the DAS was designed to protect against. (Indeed, different conceptualizations of confidentiality led to significant conflict around the Bureau's use of DP [92].) In particular, the turn to DP enables two distinct confidentiality functions: 1) empirical protections against external reconstruction of individual records and 2) because of DP's emphasis on future-proof theoretical guarantees, plausible deniability for the Bureau against any *future harms*. The latter function is a shift from earlier versions of the DAS where these guarantees could not be rigorously formalized. This expansion of the confidentiality function aligns with the Bureau's interpretation of the Title 13 confidentiality mandate, targeting worst-case risk and insulating the Bureau from both present and future legal liability. The choice to operationalize the DAS's confidentiality function using DP is an important upstream policy

Table 1: Applying the handoff model to uncover shifting values and functions in the case of the 2020 Census

The Census Bureau...	Through the hand-off lens	Conclusions about shifting values & functionality
Switched from statistical disclosure methods (SDL) to differential privacy (DP)	The function of confidentiality is preserved, but how confidentiality is operationalized has changed in response to triggers	Reveals the contested nature of <i>confidentiality</i> (§5.1)
Solicited feedback about what use cases data users value <ul style="list-style-type: none"> • Reduced the total number of counts that are held invariant 	Changing the boundaries of the system through decisions about what is inside and outside the scope of confidentiality protections	Demonstrates the significance of <i>data utility</i> (§5.2) as a function of the DAS <ul style="list-style-type: none"> • Concerns about access to political and economic resources are in tension with concerns about confidentiality
Introduced decision about the parameter epsilon as a locus of stakeholder participation	Functions (confidentiality and data utility) are now explicitly measured, formalized as a quantifiable trade-off <ul style="list-style-type: none"> • Experts evaluate and enact confidentiality through different modes, now with data-independent, ex-ante characterizations of privacy risk 	Prioritizing <i>formalized</i> (§5.3) notions of privacy and accuracy re-scope those concepts and imply the existence of an optimal trade-off <ul style="list-style-type: none"> • Expert decisions about data manipulation (through e.g. swapping) are disintermediated by DP implementation
Released significantly more information about the DAS (e.g., source code, demo data, blog posts)	Due to new DP methods, some forms of transparency no longer a threat to the DAS' confidentiality function	<i>Transparency</i> (§5.4) emerges as a value of the DAS political process, but not clear that transparency can meet all of its second-order goals <ul style="list-style-type: none"> • Expert autonomy curtailed by public scrutiny
Attempted to solicit and scaffold both expert and the public participation	Introduction of new boundary objects as components in the DAS policy process	<i>Participation</i> (§5.5) was ostensibly broadened, but with insufficient support by trusted experts

decision but, because of the Bureau’s interpretation of their legal mandate, not one over which stakeholders outside of the Bureau had input. Moreover, beyond the contested meaning of confidentiality, the handoff lens reveals that **confidentiality was far from the only value implicated** by the shift to DP.

5.2 Data Utility

The DAS attempts to balance confidentiality with *data utility*. The goal is to implement confidentiality protections that, in theory, allow a wide range of stakeholders to access and use census data, while ensuring that census takers trust these confidentiality protections enough to disclose their information. But in response to the Bureau’s decision to update the DAS, many stakeholders expressed concerns that the noise added under DP would render Census data unusable for many use cases [130, 137]. Through the handoff lens, we note that the adoption of DP shifted other components of the DAS: in particular, in order to minimize privacy loss under DP, the Bureau reduced the number of published statistics and the number of counts that were held invariant (i.e., not affected by confidentiality protections; recall §4.2.2). The decision to report specific invariant statistics reflect policy decisions about what use cases are most important and where data utility should be preserved above confidentiality – what statistics are understood to be essential for democratic representation, versus those which are malleable. Notably, the Bureau rejected a request from the National Congress of American Indians to keep state-level data for tribal areas invariant [122].

Because Census data are closely tied to the allocation of political and economic resources, decisions about data utility impact the pursuit of values like equity and justice. When the Bureau solicited feedback from data users in the Federal Register, the solicitation and the responses revealed unspoken agreements about data access and utility for a wide range of applications, including state and local government, public health, anti-discrimination efforts, research, and education efforts. Stakeholders had differing epistemic perspectives about what makes data “good enough” to be useful [19]. These ethical and epistemic questions underlie a number of important policy decisions about how the DAS should operationalize and prioritize data utility across different settings. Notably, our analysis highlights a key reframing around utility: the switch to DP and its focus on formalism (§5.3) meant that *utility* was largely operationalized as *accuracy*, thus collapsing this epistemic debate.

5.3 Formalism

The shift to DP introduced a formal definition of both privacy and data utility. While SDL methods could be formalized as a series of rules, they did not allow the Bureau to quantify the resulting confidentiality protections. The Bureau highlighted the advantages of formalism, citing provable and externally verifiable guarantees [5] as well as precision in balancing competing values [9]. An extensive literature on quantification in the history and sociology of science examines why the call to numbers has been so powerful, particularly in bureaucratic contexts [35, 37, 46, 101]. Quantitative approaches to ethical questions promise to make political decisions visible [46, 118] and facilitate debate in a common language [35, 37], creating an avenue toward accountability by facilitating participation in democratic deliberation.

Specifically, the Bureau’s embrace of formalism allowed stakeholders to reason about previously hidden policy decisions and made salient the balance and trade-off between privacy and accuracy (the latter often standing in for data utility §5.2). This is particularly salient in discussions about the privacy parameter ϵ , which promised to let the Bureau trade off between privacy protection and accuracy by operationalizing these complex values in an explicitly quantifiable way. The definition of DP does not dictate any particular value of ϵ , which must instead be tuned in accordance with the specific normative context of a particular data use case and a (deemed-acceptable) degree of privacy loss [42], i.e., “ ϵ is a parameter chosen by policy” [43]. As a result, determining the value of ϵ became a central focus of policy discussions surrounding the Bureau’s adoption of DP. The choice of what to quantify reflects decisions about *where to draw attention* [47]. We argue that the Bureau’s focus on formalism in the new DAS narrowed the scope of stakeholder participation by drawing attention to a single parameter of interest rather than a suite of decisions involved in the DAS handoff. Although there are many implementation decisions involved in DP, the formalization of the privacy-accuracy tradeoff in a single parameter focused attention on a single highly visible policy decision.

Stakeholders had previously largely ignored this tradeoff, effectively taking the Bureau’s prior statistical releases as ground truth [19]. Yet in describing the choice facing stakeholders as a quantifiable tradeoff between accuracy and privacy, the Bureau *implied that an optimal choice was possible* [8, 15]. This optimization discourse fit poorly onto the realities of the DAS. The Bureau acknowledged this challenge, noting that in order to get what the Data Stewardship Executive Policy considered a reasonable level of accuracy, they had to select a value of ϵ that was “far higher than those envisioned by the creators of differential privacy” [50, p. 3]. Indeed, some critics posited that the choice of ϵ created an impossible problem: “There may not be an overlap between the values of ϵ that are considered stringent enough for privacy purposes and high enough for redistricting purposes” [71, p. 2]. In this way, we see how the constraints of a more-formalized DAS backed the Bureau into a rhetorical corner.

The quantification literature cautions that numbers can hide policy decisions beneath a veneer of scientific objectivity, producing legitimacy in highly contested decision-making settings and, at times, foreclosing external intervention [101]. Moreover, the choice to quantify privileges that which is easily measurable [46]. In the context of the DAS, the Bureau called for “pre-specified, objective criteria” [65, p. 2] to compare privacy methodologies. We argue that privileging formally quantifiable confidentiality guarantees led to a *sociologically unintuitive conceptualization of privacy* [109], and one which does not capture, e.g., notions of data privacy that are dependent on context or social relations [96, 132]. Thus, the Bureau’s decision to formalize privacy harms as reidentification risk—and data utility as (lack of) statistical uncertainty—in accordance with DP were not neutral, but reflect particular assumptions about the values at the center of the DAS.

5.4 Transparency

The Bureau initially emphasized confidentiality, not transparency, as the benefit of the new DAS [5]. However, the handoff lens reveals that the shift to DP also changed the DAS’s capacity to support transparency. Because DP enabled technical details of the new DAS to be made public without compromising confidentiality, transparency emerged as a principal value of the political process surrounding the DAS. This shift opens up new possibilities for transparent relationships between Bureau researchers and the public.

A closer examination of this handoff reveals that the notion of “transparency” is in fact standing in for, and masking, many different values. In particular, there were many different ideas about the goal of transparency efforts, making it difficult for the Bureau to succeed in achieving each simultaneously. The Bureau engaged in many types of transparency, going beyond simplistic information disclosures and attempting to engage multiple audiences. Despite this, some stakeholders maintained that the Bureau was not being sufficiently transparent [12, 19, 97]. Understanding the Bureau’s various information releases as efforts to create boundary objects where competing notions of transparency were negotiated, we can unpack these transparency efforts and understand the many values and conflicts subsumed under the umbrella of transparency.

5.4.1 Transparency for Data Utility. Transparent privacy mechanisms can enable well-informed data users to make valid statistical inferences using privatized data by properly accounting for the uncertainty introduced by the privacy mechanism [52, 134]. Because of this, transparency in the DAS can make Census data more useful for statistical applications [52]. If the purpose of transparency is to enhance data utility through appropriate uncertainty quantification, transparency can be narrowly defined. In this case, only technical details are relevant objects of transparency, while disclosures about why a particular decision was made or who made a decision are outside of the scope of transparency for data utility.

Yet, even under this narrowed scope, transparency can be complicated. While transparency under DP allowed the specifics of the algorithmic design, such as the value of ϵ , to become direct objects of public scrutiny and discussion, in the face of substantial uncertainty around what might constitute an appropriate value of ϵ , evaluations of the new DAS’s privacy protections and utility constraints varied substantially across experts [60, 107]. Additionally, any data-dependent post-processing (for example, the Bureau’s decision to enforce non-negative counts) undermines analysts’ ability to estimate uncertainty [59]. To address this challenge, a group of experts requested access to the noisy measurement files, as described in §4.5, which did not include the Bureau’s post-processing steps. However, the Bureau did not initially release these files, preventing external experts from realizing the promised data utility benefits of the DAS’s transparency.⁵

⁵Further complicating the issue of transparency for data utility, many Census advocates argued that uncertainty caused by DP noise injection was minor compared to other sources in the Census’s data collection and processing unrelated to confidentiality [117]. However, with some partial exceptions, these sources of uncertainty were not made transparent, undermining the transparency efforts and foreclosing comparison to DP uncertainty. This highlights the importance of considering transparency efforts and sociotechnical systems within their larger context.

5.4.2 Transparency for Trust. Increased trust is often cited as a primary benefit of transparency efforts [e.g., 64, 103]. We can see that the Bureau’s decisions about what to make transparent— and what to keep hidden— were shaped by the importance of trust in census products. For example, the decision not to release the noisy measurement files (at odds with the pursuit of data utility through transparency as outlined in the previous subsection) was intended to preserve trust in census counts by hiding implausible counts produced by the original DP processing. However, by keeping the noisy measurement files hidden, external DP experts were not able to fully evaluate the Bureau’s implementation [19], ultimately undermining the trust the Bureau had hoped to preserve.

Outside of the noisy measurement files, the Bureau made many elements of the DAS visible during the transition to DP. Despite the Bureau’s increased transparency, however, a number of key stakeholders expressed distrust in census data products during the DAS handoff. The National Congress of American Indians expressed concern that the 2020 census data would be “inaccurate and unusable” [137, p. 3]; similarly, organizers working to increase participation in the census questioned “why they should bother putting in all this effort if the end data are going to be so noisy” [93].

Schnackenberg and Tomlinson suggest that trustworthiness perceptions are enhanced through disclosure, clarity, and accuracy [106]. Because of limited ability to disclose all relevant information—including previous details of disclosure avoidance systems, details of the reconstruction attack, and ground-truth data— stakeholders could not evaluate the Census Bureau’s choices through the information disclosures and demonstration data. In the absence of this additional information, the complex technical details of the system, along with bugs in the demonstration data products caused by the post-processing system, damaged, rather than enhanced, trust [18].

Importantly, the shift from secrecy to transparency about the perturbations of census data drew attention to data alterations and their implications that had gone unnoticed, or at least unexamined, by many stakeholders. Thus, transparency undermined trust not only in the Bureau’s implementation of DP, but also in the value of insights gained from previous census products [19]. Freeman argues that when trust between stakeholders and agencies is low, negotiations over policy implementation take on an adversarial character under which transparency can become dangerous [49]. By frontstaging the hidden work involved in the disclosure avoidance system, the Bureau revealed that decisions involved in its design were not merely sparing stakeholders mundane technical details but were in fact obscuring important policy choices. While the introduction of DP allowed the Bureau to make behind-the-scenes decision-making processes visible, this visibility exposed the slip-page between the backstage and the frontstage of agency discretion — to the detriment of trust.

5.4.3 Transparency for Accountability. Stakeholders’ ability to interrogate the data and report on its limitations helped the Bureau identify what aspects of the DAS were limiting the utility of the data for different purposes. Allowing stakeholders to engage with the data during a Census workshop revealed the post-processing stage of the DAS was introducing “unacceptable and problematic data biases and distortions” [10] and required structural changes.

This insight demonstrates the value of the Census Bureau’s transparency efforts in producing a more accountable DAS.

Yet, accountability was often limited because of a lack of transparency in what DP implementation decisions were feasible for the Census, or what limiting factors were effectively immutable. For instance, only the Bureau had access to the details of the previous SDL methods and the DP framework does not readily allow for comparisons to non-DP methods, making it challenging to assess critiques that did not agree with the DP formalization of privacy as a starting point [58, 70]. Without knowledge of what policy levers were available to them, stakeholders were constrained in their ability to change the DAS.

Additionally, accountability was further hampered by difficulties bridging different expert groups. The Bureau needed to communicate in expert language to the relevant theoretical computer science community to convey expertise and facilitate feedback. Yet, the technical jargon necessary to elicit solid feedback from that expert community yielded communications that alienated other expert stakeholders. A letter in July 2022 from the National Congress of American Indians specifically requested that the Bureau avoid the use of jargon and technical terms in their communications with tribal leadership, citing that prior tribal consultations were “at far too high literacy levels for a lay audience and were therefore not meaningful consultation sessions” [137]. While the Bureau recognized the importance of translating across varied stakeholder groups [10, 121], the challenge of doing so proved difficult to overcome and presented a persistent obstacle to accountability.

5.5 Participation

The Bureau’s process for engaging stakeholders around the 2020 Census included a number of innovations to support both democratic and technocratic elements of agency policy-making [90]. As outlined in §5.4, DP newly allowed transparency in the DAS, which in turn enabled a wider range of actors to be made aware of and participate in policy decisions embedded within the DAS.⁶

Increased technocratic participation became clear: during this shift, the Bureau brought in a range of experts and opened itself up to external expert review. These experts considered not only the technical details of DP and the DAS, but also provided input and review of the Bureau’s communications around the system.

More democratic participation was less clear. Such participation was mediated by Bureau’s choices about who constituted a relevant public and how to communicate with them. While the Census Scientific Advisory Committee’s DP working group applauded the Bureau for their efforts to include multiple perspectives, the committee also noted that it was difficult to assess what perspectives were not included and that many relevant stakeholders might not have the awareness, time, or energy to engage in policy decisions around the Census’s implementation of DP [20].

Nevertheless, the Bureau introduced multiple innovations to facilitate democratic participation around complex technical artifacts. First, the Bureau moved beyond static notions of engagement like

⁶Stakeholders have participated in and influenced prior iterations of the DAS. For example, “data user dissatisfaction” led the Bureau to pursue alternative methods to table suppression in the wake of the 1970 and 1980 Censuses [85, 125, p. 4-6]. However the DAS’s reliance on “security by obscurity” meant that external stakeholders could not participate directly in many design decisions.

the traditional notice-and-comment process (i.e., the Bureau releases information at one moment in time, after which the public provides feedback). Instead, they introduced dynamic and ongoing engagement using a mix of videos, webinars and other educational materials coupled with listening sessions and discussions throughout development. Second, the Bureau published data artifacts produced by different policy choices, to scaffold better understanding of those policy choices. These data sets allowed stakeholders to interactively and intuitively explore the impact of potential implementation choices on their equities. Crucially, these demonstration products revealed a desire for boundary objects that would bridge between SDL methods and DP and allow users to interact with the new system in their own varied contexts. Through these artifacts, the Bureau attempted to surface implementation decisions that would be understood and shaped by multiple communities of practice. The Bureau ultimately did incorporate user feedback into the DAS design process in several cases. When the Bureau announced their final choice of ϵ , they emphasized that it was selected in response to user feedback from demonstration data analysis and that it was “exponentially higher” than the value of ϵ proposed alongside earlier (expert-designed, expert-led) artifacts. Yet, the choice to focus on a single parameter, rather than a range of policy decisions, limited where democratic participation was possible. [123].

6 BEYOND THE CENSUS: LESSONS FOR TRANSPARENCY AND PARTICIPATION

Technological changes in a system are not only technical: they reconfigure the social, political, and organizational contexts in which they occur. Understanding these reconfigurations is crucial for responsible, trustworthy, and accountable systems. [79, 89]. We offer three lessons about how understanding these reconfigurations can enable meaningful governance, where transparency and participation interventions would otherwise fall short.

6.1 Lesson 1: *The handoff lens is a critical tool for surfacing values*

Applying of the handoff model in the census case enabled us to systematically untangle the daunting knot of actors, components, modes, functions, and values that were involved. In particular, the handoff model allows us to understand that the Census Bureau’s shift to DP did not merely produce a functionally equivalent disclosure avoidance system (DAS), preserving the core function of confidentiality protection. Instead, the shift toward DP was a deeper sociotechnical shift, reconfiguring the human and technical actors involved in the DAS and, ultimately, the values and forms of expertise embedded within this system. While working toward more accountable and trustworthy technology, it is critical to understand how the mere introduction of interventions— such as new efforts towards transparency and participation— may change the system in unintended ways. The handoff model can help us to identify these changes. For instance, the Census case reveals how the abandonment of ‘security by obscurity,’ intended as a win for transparency, precluded transparency along certain dimensions. The handoff model makes space for critical and complicated invocations of transparency and participation (§2), in realistic, on-the-ground

contexts. We suggest the handoff model as a tool that can help researchers and technologists to systematically move lessons from critical algorithm studies into practice.

6.2 Lesson 2: *Beware objects without experts*

While the FAccT community has advocated for a range of artifacts as interventions toward transparency and participation [e.g., 51, 57, 88], these artifacts have largely been divorced from the contextual changes they introduce [89, 113]. In our case the Census Bureau invested significantly in such interventions towards transparency and participation. Going beyond simplistic information disclosures, they created an impressive variety of boundary objects through which stakeholders could negotiate decisions about the DAS (§4.5). Further, the Bureau implemented many considerations that members of the FAccT community (and beyond) have long advocated for: toward *explainability* through stakeholder education efforts, toward *contextual* transparency through products like the demonstration data, and toward *contestability* through ongoing dialogue and levers for change (namely, the value and allocation of the privacy budget). Despite the Bureau’s enormous efforts, however, these boundary objects were only partially successful in facilitating meaningful participation and accountability, and in some cases they ultimately undermined trust.

The mobilization of any given boundary object is dependent not only upon the object itself, but also upon the motivation and orientation of those *brokers* that span and connect communities [72]. The Bureau’s boundary objects were in need of **trusted local experts to carry them** across community divides. Nurturing such experts is not a trivial task. Yet without them, the collaborative outcomes for which boundary objects are created in the first place might never come to fruition. Future work should explore in more detail what effective boundary object brokerage might look like in practice. A too narrow focus on artifacts can overlook the processes needed to engage them.

As the Census Bureau case demonstrates, **boundary objects cannot travel alone**. The Bureau’s focus on creating boundary objects, however innovative, was insufficient to build trust and comprehension among a diverse ecosystem of stakeholders. The epistemological and disciplinary chasms separating the communities which the Census was attempting to bridge were just too wide [19]. We encourage the FAccT community to think about the expertise needed to shepherd and use such boundary objects effectively in order to broker meaningful trust and participation.

6.3 Lesson 3: *Transparency and participation should center values and policy*

Through our case study, we can expand theoretical critiques of transparency and participation to better understand tensions on the ground. We highlight complexity of actualizing transparency and participation in practice: despite efforts to solicit feedback over technical and design decisions, the Bureau faced criticism for not being sufficiently transparent or enabling sufficient participation. While the lessons from any one case is necessarily limited, we argue that a significant revelation from this case is that transparency efforts should not be purely about technical decisions, and

that participation efforts should not be purely about design decisions. Rather, **both transparency and participation efforts should foreground decisions about values**. Importantly, providing transparency into technical decisions alone is not enough to reveal these values decisions. In fact, focusing on technical decisions can bound participation by making certain policy choices visible while neglecting others. In the Census Bureau’s adoption of DP, for instance, a narrow focus on the privacy-loss parameter, ϵ , privileged the privacy-accuracy trade-off (and with narrow conceptualizations of both privacy and accuracy). Meanwhile other value-laden policy levers—including how confidentiality should be conceptualized and operationalized, what data should be within the scope of the DAS protections and what should be held invariant, and how the Bureau might advance values like equity and collective benefit [24]—were often less visible and therefore less accessible to participation. We argue that by prioritizing the visibility of values and policy on the same level, or even above, the visibility of technical details, the FAcCT community can better leverage transparency and participation toward accountability and trust.

7 CONCLUSION

The adoption of differential privacy by the U.S. Census Bureau marked a pivot in their practices around transparent and participatory algorithmic governance. The complex nature of this adoption, and its subsequent impacts revealed the ways in which *handoffs* in algorithmic adoption in government must be mediated by different stakeholders with different levels of expertise, including via the use of carefully-designed boundary objects, to allow for meaningful participation. The lessons learned here apply more broadly to processes of algorithmic adoption, well-intentioned (and carefully planned) shifts towards transparency, and practices for successful handoffs in modern algorithmic governance.

8 RESEARCH ETHICS AND SOCIAL IMPACT

8.1 Ethical concerns

8.1.1 Methods. To the best of our understanding, there are no significant ethical concerns inherent to this work, as it is based in analysis of only publicly available documents. No interviews or sensitive data were collected for this paper. As such, no IRB approval was sought, as this work does not interface with human subjects.

8.1.2 Fairness. As we undertook our analysis, we took care to consider and portray the opinions of the various communities involved as fairly and equitably as possible, while understanding that some of these communities have been in active disagreement around the specifics of the 2020 DAS for years. Due to representation in the public archive and space limitations, we acknowledge that we were not able to represent every stakeholder viewpoint nor every notable event in the history of the Bureau’s DP implementation.

8.2 Positionality

All four authors are U.S. citizens. While we are all thus implicated in matters of American legislative representation and voting rights, we all also reside in well-resourced regions which are not threatened by census undercounts or exclusion, nor by active infringements upon voting rights. One author participated in processes around the 2020

DAS in real time; the other three became involved the project post-2020, and do not belong to any of the primary stakeholder groups that were most active in the census debate. One author is trained in computer science, two in the mathematical and physical sciences, and one as a lawyer; these backgrounds informed our comprehension of and perspectives on the legal and technical processes at play.

8.3 Adverse impact statement

The primary adverse impact that this work could have would be playing into the hands of those who would weaponize the census for political gain. Given the heavily politicized nature of the census in general, and of the DP debate in particular, we cannot anticipate how or whether this work could be used to undermine the legitimacy of the census. Further, given the importance of the census for essential societal processes such as redistricting and resource allocation (which we address in our paper), we cannot dismiss the potential for such weaponization as inconsequential.

Unfortunately, there is indeed precedent for such adverse impact. During the DAS development process, the Bureau faced direct political threats to its data products, the most serious of which arose in March 2021 when the state of Alabama sued the Department of Commerce and the Census Bureau in federal district court, alleging that by adopting DP the Bureau had “manipulated” and “intentionally skewed” the redistricting data that they provided to states [1]. Furthermore, the coincidence of the decennial count with the 2020 presidential election, as well as the uncertainty around the Trump administration’s proposal to include a citizenship question on the census, drew political attention to the count. In a time when political actors were searching for any chinks in governmental armor, a Federal agency which was public about internal sources of error became an easy target. Indeed, the Bureau has faced bipartisan scrutiny for the troubles made evident by the implementation of DP – including allegations that DP was a Trump administration tactic attempting to ‘game’ federal funding allocations, and directly contradictory allegations that DP was a Democratic tactic to destabilize the Trump administration [18, p. 32]. Of course, such critiques undermine the ultimate role of the Bureau – to produce representative population counts – and further muddy the already-cloudy waters when it comes to identifying an appropriate implementation of DP.

Ultimately, we believe that the benefits that publishing our analysis might provide – hopefully, insights for both more effective algorithmic governance and more critical algorithmic scholarship – outweigh any potential risks for further weaponization.

ACKNOWLEDGMENTS

We would like to thank Jeremy Seeman, the participants at Yale ISP’s 2023 Data (Re)Makes the World Conference and the 2023 Privacy Law Scholars Conference. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 2146752. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Deirdre K. Mulligan is a Professor at UC Berkeley School of Information, a faculty director of the Berkeley Center for Law and Technology, and currently serving as Principal Deputy US Chief Technology Officer in the White

House Office of Science and Technology Policy. The content herein represents the personal views of the authors and is not intended to reflect the views of the United States Government or any Federal agency.

REFERENCES

- [1] 2021. *Alabama v. U.S. Dep't of Commerce*. 546 F. Supp. 3d 1057 (M.D. Ala.).
- [2] Rediet Abebe, Solon Barocas, Jon Kleinberg, Karen Levy, Manish Raghavan, and David G. Robinson. 2020. Roles for computing in social change. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAccT '20)*. Association for Computing Machinery, New York, NY, USA, 252–260. <https://doi.org/10.1145/3351095.3372871>
- [3] John Abowd, Robert Ashmead, Ryan Cumings-Menon, Simson Garfinkel, Micah Heineck, Christine Heiss, Robert Johns, Daniel Kifer, Philip Leclerc, Ashwin Machanavajhala, Brett Moran, William Sexton, Matthew Spence, and Pavel Zhuravlev. 2022. The 2020 Census Disclosure Avoidance System Top-Down Algorithm. *Harvard Data Science Review* Special Issue 2 (jun 24 2022). <https://hdr.mitpress.mit.edu/pub/7evz361i>.
- [4] John M. Abowd. 2018. Disclosure Avoidance for Block Level Data and Protection of Confidentiality in Public Tabulations. <https://www2.census.gov/cac/sac/meetings/2018-12/abowd-disclosure-avoidance.pdf>
- [5] John M. Abowd. 2018. Protecting the Confidentiality of America's Statistics: Adopting Modern Disclosure Avoidance Methods at the Census Bureau. https://www.census.gov/newsroom/blogs/research-matters/2018/08/protecting_the_conf.html Section: Government.
- [6] John M. Abowd. 2021. Declaration of John M. Abowd. In *State of Alabama v. U.S. Department of Commerce*. <https://censusproject.files.wordpress.com/2021/04/2021.04.13-abowd-declaration-alabama-v-commerce-ii-final-signed.pdf>
- [7] John M. Abowd and Michael B. Hawes. 2023. Confidentiality Protection in the 2020 US Census of Population and Housing. *Annual Review of Statistics and Its Application* 10, 1 (March 2023), 119–144. <https://doi.org/10.1146/annurev-statistics-010422-034226>
- [8] John M Abowd and Ian M Schmutte. 2019. An economic analysis of privacy protection and statistical accuracy as social choices. *American Economic Review* 109, 1 (2019), 171–202.
- [9] John M. Abowd and Victoria A. Velkoff. 2019. Balancing privacy and accuracy: New opportunity for disclosure avoidance analysis. *Census Blogs* (2019).
- [10] John M. Abowd and Victoria A. Velkoff. 2020. Modernizing disclosure avoidance: What we've learned, where we are now. *Census Blogs* (2020).
- [11] Madeleine Akrich. 1992. The De-Scriptio of Technical Objects. In *Shaping Technology / Building Society: Studies in Sociotechnical Change*, Wiebe E. Bijker, John Law, Trevor Pinch, and Rebecca Slayton (Eds.). MIT Press, Cambridge, MA, USA, 208.
- [12] Kevin Allis. 2020. [Letter from Kevin Allis to Steven D. Dillingham]. https://archive.ncai.org/policy-research-center/research-data/recommendations/NCAl_Letter_to_US_Census_Bureau_on_DAS_6_25_2020_FINAL_signed.pdf
- [13] Mike Ananny and Kate Crawford. 2016. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society* 20, 3 (2016), 973–989.
- [14] Solon Barocas and Moritz Hardt. 2014. Scope. <https://www.fatml.org/schedule/2014/page/scope-2014>
- [15] Andrew Bell, Ian Solano-Kamaiko, Oded Nov, and Julia Stoyanovich. 2022. It's Just Not That Simple: An Empirical Study of the Accuracy-Explainability Trade-off in Machine Learning for Public Policy. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 248–266. <https://doi.org/10.1145/3531146.3533090>
- [16] Joseph R. Biden. 2023. Executive order on the safe, secure, and trustworthy development and use of artificial intelligence. (2023).
- [17] Abeba Birhane, William Isaac, Vinodkumar Prabhakaran, Mark Diaz, Madeleine Clare Elish, Jason Gabriel, and Shakir Mohamed. 2022. Power to the people? opportunities and challenges for participatory AI. *Equity and Access in Algorithms, Mechanisms, and Optimization* (2022), 1–8.
- [18] Dan Bouk and danah boyd. 2021. Democracy's Data Infrastructure. <http://knightcolumbia.org/content/democracys-data-infrastructure>
- [19] danah boyd and Jayshree Sarathy. 2022. Differential Perspectives: Epistemic Disconnects Surrounding the U.S. Census Bureau's Use of Differential Privacy. *Harvard Data Science Review* Special Issue 2 (June 2022). <https://doi.org/10.1162/99608f92.66882f0e>
- [20] Jay Breidit, Deborah Balk, John Czajka, Kathy Pettit, Allison Plyer, Kunal Talwar, Richelle Winkler, and Joe Whitley. 2020. Differential Privacy Working Group Deliverables: Report of the CSAC Differential Privacy Working Group. <https://www2.census.gov/cac/sac/differential-privacy-wg-deliverables.pdf>
- [21] Jenna Burrell. 2016. How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society* 3, 1 (June 2016), 205395171562251. <https://doi.org/10.1177/205395171562251>
- [22] Pat Cantwell. 2021. How We Complete the Census When Households or Group Quarters Don't Respond. <https://www.census.gov/newsroom/blogs/random-samplings/2021/04/imputation-when-households-or-group-quarters-dont-respond.html> Section: Government.
- [23] Paul R. Carlile. 2002. A pragmatic view of knowledge and boundaries: Boundary objects in new product development. *Organization Science* 13, 4 (2002), 442–455.
- [24] Stephanie Russo Carroll, Ibrahim Garba, Oscar L Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materchera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, et al. 2020. The CARE principles for indigenous data governance. *Data Science Journal* 19 (2020), 43–43.
- [25] Danielle Keats Citron. 2007. Technological due process. *Wash. UL Rev.* 85 (2007), 1249.
- [26] Cary Coglianese and David Lehr. 2019. Transparency and algorithmic governance. *Administrative law review* 71, 1 (2019), 1–56.
- [27] European Commission. 2021. Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. *Eur Comm* 106 (2021), 1–108.
- [28] A Feder Cooper, Emanuel Moss, Benjamin Laufer, and Helen Nissenbaum. 2022. Accountability in an algorithmic society: relationality, responsibility, and robustness in machine learning. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 864–876.
- [29] Ned Cooper, Tiffanie Horne, Gillian R Hayes, Courtney Heldreth, Michal Lahav, Jess Holbrook, and Lauren Wilcox. 2022. A systematic review and thematic analysis of community-collaborative approaches to computing research. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [30] Eric Corbett and Emily Denton. 2023. Interrogating the T in FAccT. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 1624–1634.
- [31] Eric Corbett, Emily Denton, and Sheena Erete. 2023. Power and Public Participation in AI. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*. 1–13.
- [32] Sasha Costanza-Chock. 2020. *Design justice: Community-led practices to build the worlds we need*. The MIT Press.
- [33] Fernando Delgado, Stephen Yang, Michael Madaio, and Qian Yang. 2021. Stakeholder Participation in AI: Beyond “Add Diverse Stakeholders and Stir”. *arXiv preprint arXiv:2111.01122* (2021).
- [34] Deloitte. 2020. Trustworthy AI: Bridging the ethics gap surrounding AI. <https://www2.deloitte.com/us/en/pages/deloitte-analytics/solutions/ethics-of-ai-framework.html>
- [35] William Deringer. 2018. *Calculated values: Finance, politics, and the quantitative age*. Harvard University Press.
- [36] Uma Desai. 2019. [usensusbureau/census-dp](https://github.com/usensusbureau/census-dp). <https://github.com/usensusbureau/census-dp>
- [37] Alain Desrosières. 1998. *The politics of large numbers: A history of statistical reasoning*. Harvard University Press.
- [38] Nicholas Diakopoulos. 2016. Accountability in algorithmic decision making. *Commun. ACM* 59, 2 (2016), 56–62.
- [39] Nicholas Diakopoulos and Michael Koliska. 2017. Algorithmic transparency in the news media. *Digital journalism* 5, 7 (2017), 809–828.
- [40] Irit Dinur and Kobbi Nissim. 2003. Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, San Diego California, 202–210. <https://doi.org/10.1145/773153.773173>
- [41] Cynthia Dwork, Gary King, Ruth Greenwood, William T. Adler, and Joel Alvarez. 2021. Re: Request for release of “noisy measurements file” by September 30 along with restricting data products. https://gking.harvard.edu/files/gking/files/2021.08.12_group_letter_to_abowd_re_noisy_measurements.pdf
- [42] Cynthia Dwork, Nitin Kohli, and Deirdre Mulligan. 2019. Differential Privacy in Practice: Expose your Epsilons! *Journal of Privacy and Confidentiality* 9, 2 (Oct. 2019). <https://doi.org/10.29012/jpc.689>
- [43] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory of Cryptography (Lecture Notes in Computer Science)*, Shai Halevi and Tal Rabin (Eds.). Springer, Berlin, Heidelberg, 265–284. https://doi.org/10.1007/11681878_14
- [44] Upol Ehsan, Q Vera Liao, Michael Muller, Mark O Riedl, and Justin D Weisz. 2021. Expanding explainability: Towards social transparency in ai systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [45] John Eltinge, Robert Sienkiewicz, Michael B. Hawes, Quentin Brummet, Edward Mulrow, Kurt Wolter, David Van Riper, Tracy Kugler, Johnathan Schroeder, JosA© Pacas, Steven Ruggles, Brian Asquith, Brad Hershbie, Shane Reed, and Steve Yesiltepe. 2019. Differential Privacy for 2020 US Census. <https://assets.pubpub.org/j2yr11kl/11587735061843.pdf>
- [46] Wendy Nelson Espeland and Mitchell L. Stevens. 1998. Commensuration as a social process. *Annual Review of Sociology* 24, 1 (1998), 313–343.

- [47] Wendy Nelson Espeland and Berit Irene Vannebo. 2007. Accountability, quantification, and law. *Annu. Rev. Law Soc. Sci.* 3 (2007), 21–43.
- [48] International Organization for Standardization. 2020. ISO/IEC TR 24028:2020 Overview of trustworthiness in artificial intelligence. <https://www.iso.org/standard/77608.html>
- [49] Jody Freeman. 1997. Collaborative governance in the administrative state. *UCLA L. Rev.* 45 (1997), 1.
- [50] Simson L. Garfinkel, John M. Abowd, and Sarah Powazek. 2018. Issues Encountered Deploying Differential Privacy. In *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*. ACM, Toronto Canada, 133–137. <https://doi.org/10.1145/3267323.3268949>
- [51] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for datasets. *Commun. ACM* 64, 12 (Nov. 2021), 86–92. <https://doi.org/10.1145/3458723>
- [52] Ruobin Gong. 2022. Transparent privacy is principled privacy. *Harvard Data Science Review Special Issue 2* (June 2022). <https://doi.org/10.1162/99608f92.b5d3faaa>
- [53] Ben Green. 2019. “Good” isn’t good enough. In *Conference and Workshop on Neural Information Processing Systems (AI for Social Good Workshop)*. Vancouver. https://aiforsocialgood.github.io/neurips2019/accepted/track3/pdfs/67_aisg_neurips2019.pdf
- [54] Kenneth Haase. 2021. `uscensusbureau/DAS_2020_Redistricting_Production_Code`. https://github.com/uscensusbureau/DAS_2020_Redistricting_Production_Code
- [55] Sam Haney, William Sexton, Ashwin Machanavajjhala, Michael Hay, and Gerome Miklau. 2021. Differentially Private Algorithms for 2020 Census Detailed DHC Race & Ethnicity. <https://doi.org/10.48550/arXiv.2107.10659> [cs, stat].
- [56] Michael Hawes. 2021. The Census Bureau’s Simulated Reconstruction-Abetted Re-identification Attack on the 2010 Census. <https://www.census.gov/data/academy/webinars/2021/disclosure-avoidance-series/simulated-reconstruction-abetted-re-identification-attack-on-the-2010-census.html> Section: Government.
- [57] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, and Kasia Chmielinski. 2020. The Dataset Nutrition Label: A Framework to Drive Higher Data Quality Standards. In *Data Protection and Democracy*, Dara Halilinan, Ronald Leenes, Serge Gutwirth, and Paul De Hert (Eds.). Data Protection and Privacy, Vol. 12. Bloomsbury Publishing, 1–26. Google-Books-ID: F2HRDwAAQBAJ.
- [58] V Joseph Hotz, Christopher R Bollinger, Tatiana Komarova, Charles F Manski, Robert A Moffitt, Denis Nekipelov, Aaron Sojourner, and Bruce D Spencer. 2022. Balancing data privacy and usability in the federal statistical system. *Proceedings of the National Academy of Sciences* 119, 31 (2022), e2104906119.
- [59] V. Joseph Hotz and Joseph Salvo. 2022. A Chronicle of the Application of Differential Privacy to the 2020 Census. *Harvard Data Science Review Special Issue 2* (June 2022). <https://hdsr.mitpress.mit.edu/pub/ql9z7ehf>.
- [60] Jessica Hullman. 2022. Show me the noisy numbers! (or not). <https://statmodeling.stat.columbia.edu/2022/12/28/show-me-the-noisy-numbers-or-not/>
- [61] Ben Hutchinson, Andrew Smart, Alex Hanna, Emily Denton, Christina Greer, Oddur Kjartansson, Parker Barnes, and Margaret Mitchell. 2021. Towards accountability for machine learning datasets: Practices from software engineering and infrastructure. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 560–575.
- [62] Abigail Z Jacobs. 2021. Measurement as governance in and for responsible AI. *arXiv preprint arXiv:2109.05658* (2021).
- [63] Abigail Z Jacobs and Deirdre K Mulligan. 2022. The Hidden Governance in AI. *The Regulatory Review* (July 2022).
- [64] Julia Jahansoozi. 2006. Organization-stakeholder relationships: exploring trust and transparency. *Journal of management development* 25, 10 (2006), 942–955.
- [65] Ron S Jarmin, John M Abowd, Robert Ashmead, Ryan Cummings-Menon, Nathan Goldschlag, Michael B Hawes, Sallie Ann Keller, Daniel Kifer, Philip Leclerc, Jerome P Reiter, et al. 2023. An in-depth examination of requirements for disclosure risk assessment. *Proceedings of the National Academy of Sciences* 120, 43 (2023), e2220558120.
- [66] Sheila Jasanoff. 2016. Reclaiming the Future. In *The Ethics of Invention: Technology and the Human Future*. W. W. Norton & Company, New York, 211–245.
- [67] Margot E. Kaminski. 2020. Understanding transparency in algorithmic accountability. In *Cambridge Handbook of the Law of Algorithms*, Woodrow Barfield (Ed.). Cambridge University Press, 20–34.
- [68] Davinder Kaur, Suleyman Uslu, Kaley J Rittichier, and Arjan Durrresi. 2022. Trustworthy artificial intelligence: a review. *ACM Computing Surveys (CSUR)* 55, 2 (2022), 1–38.
- [69] Sallie Ann Keller and John M. Abowd. 2023. Database reconstruction does compromise confidentiality. *Proceedings of the National Academy of Sciences* 120, 12 (March 2023), e2300976120. <https://doi.org/10.1073/pnas.2300976120> Publisher: Proceedings of the National Academy of Sciences.
- [70] Christopher T Kenny, Shiro Kuriwaki, Cory McCartan, Evan TR Rosenman, Tyler Simko, and Kosuke Imai. 2022. Comment: The Essential Role of Policy Evaluation for the 2020 Census Disclosure Avoidance System. *arXiv preprint arXiv:2210.08383* (2022).
- [71] Christopher T. Kenny, Shiro Kuriwaki, Cory McCartan, Evan T. R. Rosenman, Tyler Simko, and Kosuke Imai. 2021. The use of differential privacy for census data and its impact on redistricting: The case of the 2020 U.S. Census. *Science Advances* 7, 41 (Oct. 2021). <https://doi.org/10.1126/sciadv.abk3283>
- [72] Chris Kimble, Corinne Grenier, and Karine Goglio-Primard. 2010. Innovation and knowledge sharing across professional boundaries: Political interplay between boundary objects and brokers. *International Journal of Information Management* 30, 5 (2010), 437–444.
- [73] Daniel N Kluttz, Nitin Kohli, and Deirdre K Mulligan. 2022. Shaping our tools: Contestability as a means to promote responsible algorithmic decision making in the professions. *Ethics of Data and Analytics. Auerbach Publications* (2022), 420–428.
- [74] Joshua A Kroll. 2018. The fallacy of inscrutability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, 2133 (2018), 20180084.
- [75] Joshua A Kroll. 2021. Outlining traceability: A principle for operationalizing accountability in computing systems. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 758–771.
- [76] Joshua A Kroll, Joanna Huey, Solon Barocas, Edward W Felten, Joel R Reidenberg, David G Robinson, and Harlan Yu. 2017. Accountable Algorithms. *University of Pennsylvania Law Review* 165, 3 (2017), 633.
- [77] Ann Langley, Kajsa Lindberg, Björn Erik MÅžrk, Davide Nicolini, Elena Raviola, and Lars Walter. 2019. Boundary Work among Groups, Occupations, and Organizations: From Cartography to Process. *Academy of Management Annals* 13, 2 (July 2019), 704–736. <https://doi.org/10.5465/annals.2017.0089> Publisher: Academy of Management.
- [78] Bruno Lepri, Nuria Oliver, Emmanuel Letouzé, Alex Pentland, and Patrick Vinck. 2018. Fair, transparent, and accountable algorithmic decision-making processes: The premise, the proposed solutions, and the open challenges. *Philosophy & Technology* 31 (2018), 611–627.
- [79] Nancy G Leveson. 2016. *Engineering a safer world: Systems thinking applied to safety*. The MIT Press.
- [80] Karen Levy, Kyla E Chasalow, and Sarah Riley. 2021. Algorithms and decision-making in the public sector. *Annual Review of Law and Social Science* 17 (2021), 309–334.
- [81] Gabriel Lima, Nina Grgić-Hlača, Jin Keun Jeong, and Meeyoung Cha. 2022. The conflict between explainable and accountable decision-making algorithms. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 2103–2113.
- [82] Michele Loi, Andrea Ferrario, and Eleonora Viganò. 2021. Transparency as design publicity: explaining and justifying inscrutable algorithms. *Ethics and Information Technology* 23, 3 (2021), 253–263.
- [83] Michael Madaio, Lisa Egede, Hariharan Subramonyam, Jennifer Wortman Vaughan, and Hanna Wallach. 2022. Assessing the Fairness of AI Systems: AI Practitioners’ Processes, Challenges, and Needs for Support. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–26.
- [84] Kirsten Martin. 2019. Ethical implications and accountability of algorithms. *Journal of Business Ethics* 160 (2019), 835–850.
- [85] Laura McKenna. 2018. *Disclosure Avoidance Techniques Used for the 1970 through 2010 Decennial Censuses of Population and Housing*. Technical Report. U.S. Census Bureau Research & Methodology Directorate. <https://www2.census.gov/ces/wp/2018/CES-WP-18-47.pdf>
- [86] Jacob Metcalf, Emanuel Moss, and danah boyd. 2019. Owning ethics: Corporate logics, silicon valley, and the institutionalization of ethics. *Social Research: An International Quarterly* 86, 2 (2019), 449–476.
- [87] minutephysics. 2019. Protecting Privacy with MATH (Collab with the Census). <https://www.youtube.com/watch?v=pT19VwBAqKA>
- [88] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* ’19)*. Association for Computing Machinery, New York, NY, USA, 220–229. <https://doi.org/10.1145/3287560.3287596>
- [89] Deirdre K Mulligan and Kenneth A Bamberger. 2018. Saving governance-by-design. *California Law Review* 106, 3 (2018), 697–784.
- [90] Deirdre K. Mulligan and Kenneth A. Bamberger. 2019. Procurement as policy: Administrative process for machine learning. *Berkeley Tech. LJ* 34 (2019), 773.
- [91] Deirdre K. Mulligan and Helen Nissenbaum. 2020. The concept of handoff as a model for ethical analysis and design. *The Oxford handbook of ethics of AI* 1, 1 (2020), 233.
- [92] Priyanka Nanayakkara and Jessica Hullman. 2022. What’s driving conflicts around differential privacy for the U.S. Census. *IEEE Security & Privacy* 01 (2022), 2–11.

- [93] National Academies of Sciences, Engineering, and Medicine. 2020. *2020 Census Data Products: Data Needs and Privacy Considerations: Proceedings of a Workshop*. National Academies Press.
- [94] National Conference of State Legislatures. 2021. *Differential Privacy for Census Data Explained*. <https://www.ncsl.org/technology-and-communication/differential-privacy-for-census-data-explained>
- [95] NCAI Policy Research Center. 2021. *Differential Privacy and the 2020 Census: A Guide to the Data Analyses and Impacts on AI/AN Data*. Research Policy Update. National Congress of American Indians, Washington, D.C. https://archive.ncai.org/policy-research-center/research-data/prc-publications/NCAI_PRC_2020_Census_Guide_to_Data_and_Impacts_5_17_2021_FINAL.pdf
- [96] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.
- [97] Steven A. Ochoa and Terry Ao Minnis. 2021. Impact of Differential Privacy & the 2020 Census on Latinos, Asian Americans and Redistricting. <https://www.maldef.org/wp-content/uploads/2021/04/FINAL-MALDEF-AAJC-Differential-Privacy-Preliminary-Report-4.5.2021-1.pdf>
- [98] OECD. 2021. Tools for trustworthy AI: A framework to compare implementation tools for trustworthy AI systems. *OECD Digital Economy Papers* 312 (Jun 2021). <https://doi.org/10.1787/008232ec-en>
- [99] National Institute of Standards and Technology. 2023. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. (Jan 2023). <https://doi.org/10.6028/nist.ai.100-1>
- [100] Jennifer Pierre, Roderic Crooks, Morgan Currie, Britt Paris, and Irene Pasquetto. 2021. Getting Ourselves Together: Data-centered participatory design research & epistemic burden. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [101] Theodore M. Porter. 1995. *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.
- [102] Inioluwa Deborah Raji, Andrew Smart, Rebecca N White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. 2020. Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 33–44.
- [103] Brad Rawlins. 2008. Give the emperor a mirror: Toward developing a stakeholder measurement of organizational transparency. *Journal of public relations research* 21, 1 (2008), 71–99.
- [104] David Gerald Robinson. 2022. *Voices in the code: a story about people, their values, and the algorithm they made*. Russell Sage Foundation, New York.
- [105] Simo Sarkki, Hannu I. Heikkinen, Teresa Komu, Mari Partanen, Karoliina Vanhanen, and Elise Lepy. 2020. How boundary objects help to perform roles of science arbiter, honest broker, and issue advocate. *Science and Public Policy* 47, 2 (2020), 161–171.
- [106] Andrew K. Schnackenberg and Edward C. Tomlinson. 2016. Organizational transparency: A new perspective on managing trust in organization-stakeholder relationships. *Journal of management* 42, 7 (2016), 1784–1810.
- [107] Mike Schneider. 2021. Census releases guidelines for controversial privacy tool. <https://apnews.com/article/business-census-2020-55519b7534bd8d61028020d79854e909> Section: Voting rights.
- [108] Jeremy Seeman. 2023. Framing Effects in the Operationalization of Differential Privacy Systems as Code-Driven Law. In *International Conference on Computer Ethics*, Vol. 1.
- [109] Jeremy Seeman and Daniel Susser. 2023. Between Privacy and Utility: On Differential Privacy in Theory and Practice. *ACM J. Responsib. Comput.* (oct 2023). <https://doi.org/10.1145/3626494> Just Accepted.
- [110] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*. 59–68.
- [111] Mona Sloane and Emanuel Moss. 2022. Introducing a Practice-Based Compliance Framework for Addressing New Regulatory Challenges in the AI Field. *TechReg Chronicle* (2022).
- [112] Mona Sloane, Emanuel Moss, Olaitan Awomolo, and Laura Forlano. 2022. Participation is not a design fix for machine learning. In *Equity and Access in Algorithms, Mechanisms, and Optimization*. 1–6.
- [113] Mona Sloane, Ian René Solano-Kamaiko, Jun Yuan, Aritra Dasgupta, and Julia Stoyanovich. 2023. Introducing contextual transparency for automated decision systems. *Nature Machine Intelligence* 5, 3 (2023), 187–195.
- [114] Mona Sloane and Janina Zakrzewski. 2022. German AI Start-Ups and “AI Ethics”: Using A Social Practice Lens for Assessing and Implementing Socio-Technical Innovation. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 935–947.
- [115] Susan Leigh Star. 2010. This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values* 35, 5 (2010), 601–617.
- [116] Susan Leigh Star and James R. Griesemer. 1989. Institutional ecology, translations and boundary objects: Amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science* 19, 3 (1989), 387–420.
- [117] Ryan Steed, Terrance Liu, Zhiwei Steven Wu, and Alessandro Acquisti. 2022. Policy impacts of statistical uncertainty and privacy. *Science* 377, 6609 (2022), 928–931. <https://doi.org/10.1126/science.abq4481> arXiv:<https://www.science.org/doi/pdf/10.1126/science.abq4481>
- [118] Cass R Sunstein. 2002. The cost-benefit state: the future of regulatory protection. American Bar Association.
- [119] Latanya Sweeney. 2000. *Simple Demographics Often Identify People Uniquely*. Working Paper. Carnegie Mellon University Data Privacy Lab, Pittsburgh. <https://dataprivacylab.org/projects/identifiability/paper1.pdf>
- [120] U.S. Census Bureau. 2018. Soliciting Feedback From Users on 2020 Census Data Products. <https://www.federalregister.gov/documents/2018/07/19/2018-15458/soliciting-feedback-from-users-on-2020-census-data-products>
- [121] U.S. Census Bureau. 2020. *2020 Census Tribal Consultations with Federally Recognized Tribes*. Report. U.S. Census Bureau. <https://www.census.gov/content/dam/Census/library/publications/2020/dec/census-federal-tc-final-report-2020-508.pdf>
- [122] U.S. Census Bureau. 2020. Invariants Set for 2020 Census Data Products. <https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/2020-das-updates/2020-11-25.html>
- [123] U.S. Census Bureau. 2021. Census Bureau Sets Key Parameters to Protect Privacy in 2020 Census Results. <https://www.census.gov/newsroom/press-releases/2021/2020-census-key-parameters.html> Section: Government.
- [124] U.S. Census Bureau. 2021. *Disclosure Avoidance for the 2020 Census: An Introduction*. Handbook. US Government Publishing Office, Washington, D.C. <https://www2.census.gov/library/publications/decennial/2020/2020-census-disclosure-avoidance-handbook.pdf>
- [125] U.S. Census Bureau. 2023. 2020 Decennial Census: Processing the Count: Disclosure Avoidance Modernization. <https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance.html>
- [126] U.S. Census Bureau. 2023. Coming This Spring: New 2010 Redistricting and DHC “Production Settings” Demonstration Microdata with Noisy Measurement Files. <https://www.census.gov/programs-surveys/decennial-census/decade/2020/planning-management/process/disclosure-avoidance/newsletters/new-2010-redistricting-dhc-demo-microdata.html> Section: Government.
- [127] U.S. Census Bureau. 2023. Disclosure Avoidance Webinar Series. <https://www.census.gov/data/academy/webinars/series/disclosure-avoidance.html> Section: Government.
- [128] U.S. Census Bureau. 2023. *Why the Census Bureau Chose Differential Privacy*. Brief C2020BR-03. U.S. Census Bureau. <https://www2.census.gov/library/publications/decennial/2020/census-briefs/c2020br-03.pdf>
- [129] Kristen Vaccaro, Karrie Karahalios, Deirdre K Mulligan, Daniel Kluttz, and Tad Hirsch. 2019. Contestability in algorithmic systems. In *Conference companion publication of the 2019 on computer supported cooperative work and social computing*. 523–527.
- [130] David Van Riper, Jonathan Schroeder, and Steven Ruggles. 2021. Feedback on the April 2021 Census Demonstration Files. https://users.pop.umn.edu/~ruggl001/Articles/IPUMS_response_to_Census.pdf
- [131] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–14.
- [132] Salome Viljoen. 2021. A relational theory of data governance. *Yale Law Journal* 131 (2021), 573.
- [133] Maranke Wieringa. 2020. What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 1–18.
- [134] Aaron R. Williams and Claire McKay Bowen. 2023. The promise and limitations of formal privacy. *Wiley Interdisciplinary Reviews: Computational Statistics* (2023), e1615.
- [135] Richmond Y. Wong. 2020. *Values by Design Imaginaries: Exploring Values Work in UX Practice*. PhD Dissertation. University of California, Berkeley, Berkeley, California.
- [136] Richmond Y Wong, Michael A Madaio, and Nick Merrill. 2023. Seeing like a toolkit: How toolkits envision the work of AI ethics. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–27.
- [137] Larry Wright, Jr. 2022. Letter from National Congress of American Indians CEO to US Census Director. https://www.ncai.org/policy-research-center/research-data/prc-publications/20220728_NCAI_Letter_to_US_Census_Bureau_FINAL.pdf
- [138] Felix T Wu. 2013. Defining Privacy and Utility in Data Sets. *University of Colorado Law Review* 84 (2013), 1117–1177.