Ethical Closure Problems in Formalizing Technologies

AMINA A. ABDU, University of Michigan, USA
JEREMY SEEMAN, Urban Institute, USA
ABIGAIL Z. JACOBS, University of Michigan, USA

When putting policy into practice, government agencies must make a range of decisions about how to implement and operationalize said policies. Ideally, such administrative decisions can be instrumented in a bureaucratically legitimate manner—in line with their intended values, in line with citizen and institutional expectations, and in legally appropriate, transparent, and just ways. We introduce the concept of *formalizing technologies*, which offer mathematical frameworks to conceptualize and operationalize fundamentally contested policy ideals, such as efficient markets (e.g., through mechanism design), fair allocation, or democratic participation (e.g., through voting). Notably, these formalizing technologies promise to make the work of administration fair, non-arbitrary, accountable, and efficient. We adopt a comparative approach to examine two cases of formalizing technologies in the U.S. government context: contemporary discourse around privacy-enhancing technology and the historical adoption of game theory for military decision-making. To analyze these cases, we draw on theories of scientific closure from across the science and technology studies (STS) literature and organizational theory, emphasizing the processes by which scientific consensus is achieved and scientific autonomy is preserved. Expanding on the relationship between scientific closure and legitimacy, we show that formalizing technologies are rhetorically powerful tools for producing both administrative legitimacy and closure over fundamentally value-laden policy questions. Through two historically distinct case studies, we identify shared rhetorical patterns and outline how formalizing technologies combine and extend features of law and quantification to scope the space of legitimate ethical decisions in bureaucratic contexts—i.e., producing *ethical closure*. We describe how this process of closure occurs, its impact on shaping regulatory decision-making and practice, and implications for technology and policy imaginations and futures.

CCS Concepts: • **Social and professional topics → Government technology policy**; • **Applied computing** → *Law*; • **Security and privacy** → *Social aspects of security and privacy*.

Additional Key Words and Phrases: Scientific Closure, Formalizing Technologies, Public Administration, Quantification, Legitimacy

1 INTRODUCTION: “PERFECT TECHNOLOGIES OF JUSTICE”

Administrative officials across U.S. government agencies serve the important but complex function of implementing policy. Despite its important role in operationalizing policy “on the ground" [30], the administrative state is plagued by frequent challenges to its legitimacy. Unlike publicly elected representatives, administrative officials are appointed: their role as policymakers (through regulation and policy implementation) is not legitimated through democratic elections. Instead, the state has historically sought to justify agency authority by other means, emphasizing non-arbitrary interpretations of policy, accountable decision-making, and efficient operations [18]. Administrators' implementation decisions are at times desired to align with a policy's intended goals and values; with citizen expectations and political machinations; and with institutional expectations, all while respecting both democratic and technocratic norms of agency policy-making [9, 10]. When such values, expectations, and norms conflict, agencies engage in legitimacy-making exercises to justify their continued policy-implementing capabilities. In this paper, we investigate how mathematical and algorithmic formalisms—specifically what we call formalizing technologies—intersect with such conflicts. We show how formalizing technologies are used to resolve these conflicts to perform legitimacy, but also reshape and create conflicts anew.

We introduce the term *formalizing technologies* (FTs) to refer to mathematical frameworks intended to operationalize and enact policy. We define formalizing technologies as mathematical frameworks that implement policy ideals, such as

Authors' addresses: Amina A. Abdu, University of Michigan, USA; Jeremy Seeman, Urban Institute, USA; Abigail Z. Jacobs, University of Michigan, USA.

efficient markets (e.g., through mechanism design), fair allocation, certain security goals (e.g., through cryptography), or democratic participation (e.g., through voting). Drawing on existing uses of formalizing technologies in government, we build upon the science and technology studies (STS) and quantification literature to explore the mechanisms through which formalizing technologies create legitimacy—and the consequences of this legitimacy for policy-making. Specifically, we show how these formalizing technologies enact *closure*: closure of social problems, of policy solutions, of values and of ethics in the political context.

Here we build on multiple strands of literature. First, we show how multiple theories of *scientific closure* in organizational theory and science and technology studies speak to the ways that consensus and stability are formed around comparable technologies. Because we explore technologies that are used to shape social, policy-oriented outcomes, we also look to how technologies in the administrative state produce closure around values. We connect closure to the production and performance of legitimacy, whereby consensus provides moral justification for the exercise of power. In a context where legitimacy is at stake and a lack of consensus can undermine agency authority, it is important to understand where this legitimacy comes from, and where it is sensitive to challenges.

Shifting to the context of the administrative state, liberal political theory presupposes legitimacy as a necessary condition for just functioning of the administrative state [67]. Seminal works in this literature identify three success criteria for justifying productive political dialogue at the heart of U.S. administration: first, rationality, the ability to explain a resource allocation decision; second, consistency, the consistent application of rules across parties and circumstances; and third, neutrality, the equal moral standing of different conceptions of the good and their advocates [68]. In combination, these three criteria can be operationalized by what Bruce Ackerman [1980] calls "perfect technologies of justice", imagined utopian means to implement policies without obstacles, distortions or costs. Such technologies dictate the terms and processes of democratic deliberation, not the content of the deliberation itself, as the point of connection between democratic legitimacy and just administration [63]. Doing so inextricably links these success criteria to the efficient functioning of the state: in Ackerman's view, an efficient liberal state is necessarily a just state.

We argue that FTs most closely align with this liberal utopian vision by scaffolding the terms of democratic deliberation in a technically productive manner. The use and rhetoric of FTs aligns with liberalism's governance goals of accountability, efficiency, and non-arbitrariness. We explore the rhetorical machinery through which formalizing technologies produce ethical closure, and its implications for policy. (Even the language of *technologies*, rather than policy instruments, serves to distance the tool from its political and social function.)

The paper proceeds as follows. In Section 2, we introduce and define *formalizing technologies* (FTs), providing examples and discussing the affordances that make FTs useful in the pursuit of legitimacy. In Section 3, we synthesize existing frameworks of scientific closure, focusing on building consensus and preserving autonomy in science. We subsequently examine how an analogous framework of *ethical closure* describes the process by which policy implementations impose consensus and autonomy in decisions about essentially contested values. In Section 4, we examine how the scientific and ethical closure process plays out through FTs, focusing on two case studies: formal privacy enhancing technology and game theory. In Section 5, we discuss how this closure process creates legitimacy through the rhetorical strength of FTs and examine the consequences of this unique form of legitimacy.

2 FORMALIZING TECHNOLOGIES

In this section we discuss prior work on the use of quantification in public administration before discussing how formal mathematical frameworks have been taken up to operationalize contested policy ideals. We begin with the larger context of technologies of quantification (§2.1) and then define a specific subset of these, which we call *formalizing technologies* (FTs) (§2.2). We discuss examples of FTs and examine how they differ from other forms of quantification.

2.1 Technologies of Quantification

Technologies of *quantification*, which enable the "communication and production of numbers" [36, p. 401], have long been a subject of interest in the history and sociology of science. Significant attention has been devoted to the use of quantification in public administration (see [54] for a more detailed review), focusing on numbers' capacity for standardization across diverse populations and for producing legitimacy.

2.1.1 Technologies of Commensuration. Through the process of commensuration– "the transformation of different qualities into a common metric" [35, p. 314]– complex and disparate information is simplified into a single standardized number. In the context of policy, commensuration can reduce different interests and values to a manageable amount of information. Crucially, commensuration redefines relationships between policy alternatives by ordering them on a single scale. This ordering suggests a hierarchy between alternatives, facilitating a seemingly optimal choice [34]. Moreover, commensuration allows administrators to use the language of economic efficiency to justify policy decisions: by assigning a market price to competing interests, commensuration facilitates market-based economic solutions to value-laden policy questions [26]. This logic has become embedded in administrative decision-making through cost-benefit analysis, which forces the commensuration of policy inputs and outputs to compare the value of policy proposals.

2.1.2 Technologies of Legitimacy. In addition to noting the utility of quantification for facilitating measurement and comparison, the quantification literature also examines the use of numbers to establish legitimacy. Across scientific and bureaucratic contexts, quantification has been particularly attractive in low-trust environments for its capacity to establish trust through appeals to objectivity [65]. Trust is central in particular to legitimacy in democratic societies, where government authority must be willingly accepted by its citizens rather than imposed upon them. The authority of quantification aligns with the U.S. administrative state's historical and ongoing pursuit of legitimacy through promises of efficiency, accountability, and non-arbitrariness [18].

2.2 Formalizing Technologies

What are formalizing technologies? Despite the rich literature on quantification, comparatively little attention has been given to variation between different forms of quantification, and different types of quantifying technologies. We conceptualize formalizing technologies (FTs) as a particular form of quantification with unique implications for public administration.

Specifically, we define formalizing technologies as mathematical frameworks that quantitatively operationalize some policy-making ideal (e.g., deservedness, fairness, efficiency, privacy preservation) and enable the design of systems that guarantee, under some explicitly stated conditions, that the outputs of the system achieve the ideal at a level pre-specified by the system's parameterization. Fundamentally, such ideals—deservedness of benefits, fairness of allocation, market efficiency, privacy—represent concepts that are value-laden and essentially contested [56, 57, 76].

Examples of formalizing technologies. We can identify and group FTs by the values they quantify and the algorithmic means through which they are expressed. Consider the following examples. FTs that quantify privacy and security principles create measurements of how much data sharing algorithms make information accessible or inaccessible to different parties as tunable functions of their parameters; example FTs include formal privacy tools like differential privacy [29] and secure multiparty computation [23]. FTs that quantify equitable outcomes create measurements of inequality, unfairness, or disparity between algorithmic outputs that can be tuned with different parameters; example FTs include optimization-based modeling under fairness constraints [55] and algorithmic solutions to fair allocation problems [45]. FTs that quantify equitable model performance create measurements of disparities in errors, variability, or stochasticity associated with algorithmic outputs; example FTs include multicalibration algorithms [40] and equity-aware privacy analyses [24].

Examples of quantification that are not formalizing technologies. Quantification captures a wide range of activities with numbers. FTs represent only a small subset of technologies of quantification. While significant attention has been devoted to quantification in government, the majority of this work focuses on empirical, evaluative technologies, e.g., state censuses, population statistics, taxation, and cost-benefit analysis. Like these technologies, FTs facilitate comparisons, produce trust and legitimacy, and facilitate the language of efficiency. However, despite this significant overlap, FTs are distinguished from other forms of government quantification by their emphasis on *ex ante* guarantees rather than *ex post* assessment.

Specifically, FTs enact policy through formally parameterized algorithmic mechanisms for quantifying essentially contested ethical concepts. Many policy interventions enact technologies of quantification but not all of these properties, placing them beyond the scope of our paper. Quantification technologies that are not FTs may operationalize contested concepts, for instance to empirically measure productivity costs or to evaluate the success of privacy attacks. However, these tools lack a *formal*—i.e., provable, causal, *ex ante*—relationship between the system design and the values being quantified. For example, many technologies quantify concepts whose contestation is inessential. Some water monitoring systems may *guarantee* certain upper bounds on the concentration of lead in drinking water [74]. However, such a system treats the concentration of lead, not the *contestable, policy-driven* assessment of what constitutes a safe concentration of lead, as the *ex ante* parameterized variable of interest. In contrast, whether a concept like "fairness" should be understood as equity, equality, justice, or arbitrarily many other understandings might be more fundamentally contested, and the technology's guarantees correspond to the policy concept of interest. Finally, FTs are necessarily parameterized, but many policy decisions are formulaic or mechanical without being parameterized; for example, allocation formulas for dispersing funding have implicit, but not explicit, parameters that dictate the policy problem's ethical dimensions.

Affordances of formalizing technologies. Describing what FTs do can be surprisingly challenging. We can reasonably intuit that FTs are technological artifacts with their own politics [77]; yet the formal logic underlying FTs produces theoretical mathematical results independent from their political context. Even asking what FTs "afford" could refer to either "how people use FTs in practice," or "how the properties of FTs admit a range of possible uses" [39, 61]. Here we refer to FT affordances in the latter sense, specifically highlighting the rhetoric of how FT developers and users engage with FTs.

FTs make explicit the algorithmic mappings between inputs and outputs. Thus these mappings can easily be made *transparent*, such that the space of decisions induced by specific FTs can be made visible to a wide range stakeholders, including participants, data subjects, and regulators. This enhanced transparency allows stakeholders to demarcate possible interventions and their inherent consequences. This transparency is apparently empowering, revealing a defined space of potential interventions.

An appealing, formalizing aspect of FTs is that they entail user-specified *parameters*—i.e. design choices—dictating how algorithms function. Because the formal reasoning underlying FTs is independent of any realized quantitative inputs or outputs to the system, FTs map the space of hypothetical inputs to hypothetical outputs. These parameters can produce constraints and locate specific policy choices—quantitative thresholds that might relate to values, such as tolerable error, trade-offs of risk or accuracy. This allows these quantitative parameters to be the locus of policy negotiation: if a value (e.g., "group fairness") were a dial, the parameters define the "notches" to which a policymaker could set the dial. Notably, FTs reflect *ex ante* choices that prescribe how algorithmic systems achieve specific quantitative properties via these notches, as opposed to *ex post* descriptions of existing systems. For example, formal privacy methods like differential privacy (DP) include pre-specified privacy loss budgets. Similarly, formal fairness methods like multi-calibration rely on multiple accuracy criteria and group definitions.

3 FROM SCIENTIFIC TO ETHICAL CLOSURE

The process by which FTs produce legitimacy is closely linked to the process of *scientific closure*, a concept developed by sociologist of science Harry Collins [1981] to describe the social mechanisms through which scientific knowledge comes to be accepted. We argue that two forms of scientific closure, consensus and exclusivity, work together around FTs. We further argue that this scientific closure creates closure over the ethical questions that FTs purport to solve by narrowing the space of legitimate policy interventions.

3.1 Scientific Closure

Here we draw on two related but distinct notions of scientific closure. The first, developed in science and technology studies (STS), describes how social groups reach consensus about the interpretation of scientific and technological artifacts. In their foundational work on the social construction of technology (SCOT), Pinch & Bijker [1984] describe the three stages of meaning-making that shape the acceptance of scientific knowledge and the adoption of technologies. During the first stage, interpretive flexibility, both the interpretation and design of a technology are malleable. In the second stage, closure, consensus develops around a single meaning and design of the technology, either through rhetorical closure or redefinition of the problem. In the third and final stage, the now-closed technology becomes embedded in the wider sociocultural context [64]. This framework has been taken up and extended in the STS literature to examine the resolution of scientific controversies.

The second form of scientific closure we draw upon refers to the hidden backstage of knowledge production, in contrast to open science. From an institutional perspective, closure refers not to consensus-making but instead to limiting external intervention in science [20], for example by maintaining confidentiality during peer review, delaying the public dissemination of research findings, or keeping details of the scientific process hidden. Ostensibly intended to preserve expert autonomy, this form of closure is susceptible to co-optation by other interests, for example through industry partnerships or government funding of science [20]. While the STS literature has dealt with issues of control over science it is worthwhile to consider this exclusivity in the context of consensus-building in order to understand how both consensus and exclusivity work together under the umbrella of scientific closure. Scientific closure, through both consensus and exclusivity, fosters legitimacy by minimizing contestation and obscuring prior disagreements to elevate certain claims as scientific. However, this legitimacy is not stable. Increasingly, waning trust in science's ability to answer policy questions had led to demands for active public participation in matters where science and technology intersect with the public interest [59].

3.2 Ethical Closure

In contrast to scientific closure, we use the term *ethical closure* to refer to processes that operationalize value-laden issues. Different approaches to conceptualizing and operationalizing value-laden issues may rest upon fundamentally conflicting values with social, political, and context-varying dimensions, rendering many such issues essentially contested [38]. To handle this problem, legal and policy interventions necessarily "constrain" ethics—specifying how values-based decisions can be engaged or intervened upon—so that institutions can act "with the force of law" [41]. For example, legal interventions on discrimination (e.g., Title IX, re: discrimination "on the basis of sex") offer institutions such force of law (e.g., through coordinators tasked with overseeing and reporting such cases of discrimination).

We argue that, analogously, **FTs invoke ethical closure by constraining ethics with the force of technology.** While legal and policy interventions constrain how political institutions might set and enforce rules or standards, FTs leverage their technological affordances and disaffordances to constrain the space of acceptable interventions. Tools that algorithmically assess the discriminatory nature of predictions, for example, outline the space of "feasible" algorithmic configurations that trade off predictive accuracy and a quantification of their discriminatory impact [22]. The adoption of a particular FT closes out policy alternatives, not on the grounds of their ethical merit, but on the grounds that they are technically incompatible with a given tool. Consequently, technical and scientific closure imposes ethical closure over otherwise open ethical questions.

4 CLOSURE IN PRACTICE

The way textual policy problems are framed and technological problems are framed [42] necessarily has social and political effects on how they are intervened upon. However, the manner in which textual policies and FTs induce ethical closure are substantively different. We characterize how FTs create ethical closure, and the work that does to produce, shape, and even undermine legitimacy. Using examples, we show how this process that merits distinct attention from our community. We draw on two case studies in particular: formal privacy-enhancing technologies and game theoretic decision-making for military applications. We select these two cases to leverage variation in time scale, technical maturity, and the range of possible applications. Formal privacy is relatively new, and has a relatively narrow range of applications to privacy questions. Meanwhile, game theory has been deployed for decades in varying contexts, and with emphasis on a wider range of ethical problems. Despite these differences, we observe similarities across both cases.

4.1 Case 1: Formal Privacy-Enhancing Technologies

Privacy has been enshrined in U.S. public policy through laws like the Health Insurance Portability and Accountability Act (HIPAA), the Family Educational Rights and Privacy Act (FERPA), and the Confidential Information Protection and Statistical Efficiency Act (CIPSEA). Implementing these laws is often under the purview of federal agencies (see e.g., the eGov Act, Title 13 of the U.S. Code, and the Federal Trade Commission's role in privacy enforcement), with little guidance on how such implementation occurs, sometimes leading to ineffective implementations [11] and bouts of contestation around privacy norms [56]. In particular, U.S. data privacy law has historically struggled with defining the scope of "personal information" for data sharing purposes, since modern computing and data infrastructures make it easier than ever to associate data with individuals, known as the "failure of anonymization" in law [62, 78]. Agencies with dual mandates to make data accessible while safeguarding privacy must resolve these issues to function successfully. Recently, agencies have looked to formal methods in privacy enhancing technologies as a promising response to this failure. We focus on policy discussions surrounding the adoption of formal privacy for the 2020 Decennial Census in this case study.

Formal privacy (FP) technologies, often referring to DP [27] and related techniques [25], are FTs aimed at providing ex ante measures of information leakage from data processing systems. Most FP technologies use noise injection to characterize formal claims about the ability for adversaries to distinguish between results in the presence or absence of one person's data. The degree of noise injection, and thus the formal limitation on information leakage, is parameterized by what's known as a privacy loss budget (PLB); smaller PLBs refer to greater formal information leakage protections, and vice versa. In other words, PLBs associate with each data release an upper bound on the amount of information leakage that could occur. Previous work on formal privacy for the U.S. Census Bureau has examined epistemic attitudes towards Census data [17], communication practices around data privacy for data users [58], and the operationalization of transparency and accountability [1]. However, little attention has been paid to the rhetoric used by the U.S. Census Bureau to justify the adoption of specifically *formal* standards.

The Census Bureau used the formal nature of FP's guarantees to justify their macroscopic decision-making process, but notably *not* their specific FP implementation decisions (i.e., what data to release and how to set the associated PLBs for each release). To explain the change from non-formal methods to formal methods, the Bureau had to empirically demonstrate that the previous methodology failed to protect data subject confidentiality. While empirical experiments correctly demonstrated that previous decennial censuses indeed posed privacy risks [2, 3], no such empirical demonstrations were publicly produced by the FP system. Instead, the Bureau relied exclusively on FP's guarantees in both technical [49] and non-technical [46] communications with stakeholders. This is surprising because the Bureau implemented FP with what was considered a large PLB, implying the formal upper bound on privacy loss was too large for data subjects to infer anything useful about their particular risk; in other FP deployment scenarios, like Apple's user analytics FP deployment, such large PLBs rightfully raise concerns about "privacy-washing" through formal, but ultimately vacuous, privacy guarantees [73]. Yet regardless of these particular implementation choices, the Bureau explicitly argued that FP was the right tool for solving disclosure risk problems because of its formal properties [43].

This logic proved ultimately beneficial for the Census Bureau. First, it helped transparently demonstrate to the Department of Justice that the Census Bureau could ex ante satisfy its legal confidentiality mandates specified in Titles 13 and 26 of the federal code. Second, the formal guarantees ensured that no matter the background conditions by which adversarial data users might exploit Census data, the Census Bureau would not be held responsible for aiding and abetting privacy harms beyond their control. And finally, by focusing on selecting a sufficiently large PLB, the Census could demonstrate a commitment to producing useful (read: efficient) statistical estimates needed to fulfill its role managing the core statistics supporting democratic representation, appropriations for federal funding, and other official statistics uses.

At the same time, the Census Bureau's implementation left many policy questions unanswered. Many empirical conceptions of privacy harm, such as those measuring the ability to infer information about individuals (for example, race– see [47]), were labeled as "unscientific" from both the academic community [28] and the Census Bureau [43] because they could not be simultaneously formalized (via ex ante privacy guarantees) and controlled (via choosing a PLB). To them, one could debate how to implement FP, but to suggest measuring privacy empirically outside of FP was "unscientific," "ad hoc", or not "rigorous," even when concretely tied to policy outcomes [48]. FP, in summary, helped the U.S. Census Bureau demonstrate particular commitments to both privacy protections and the production of high-quality data, but the interplay between FP and holistic policy evaluation was largely left unexplored.

4.2 Case 2: Game Theoretic Decision-Making and Mechanism Design

The draw of formalization is unique neither to recent technologies, nor to contemporary bureaucratic cultures. We can compare more ongoing developments in formal privacy-enhancing technologies with an earlier use of FTs to resolve value-laden policy dilemmas: the U.S.'s uptake of game theory during World War II and the Cold War. The rise (and subsequent fall) of game theory as a policy tool demonstrates both the power of FTs to enact closure, and the fragility of this closure.

Institutionally, WWII led to a major renegotiation of the relationship between science and government in the United States. During the war, the Office of Scientific Research and Development organized significant and unprecedented cooperation between scientists and the war efforts. After the war, those same scientists were involved with developing the institutions to support science (both through federal agencies like the National Science Foundation and non-governmental organizations like RAND). The emphasis on legitimacy of these institutions was central in their founding documents, where scientific progress is framed as an essential component of government [19], and only conducted in legitimate ways (from RAND, this must be "strictly nonpartisan, rigorously analytic" [66]).

Importantly for our context, they elevated, and were elevated by, the adoption and legitimation of game theory in geopolitical context, which itself promised an objective, analytic way to settle political problems. In the 1920s and early 1930s, the mathematician John von Neumann and economist Oskar Morgenstern, among others, developed the mathematical machinery to capture social and psychological phenomena around games (literal games: chess, poker, baccarat). Increasing unrest in Europe after 1933 led von Neumann to imagine this formalism as a solution to the rising social instability and a tool to understand the impending international crisis [52, p.189–213].

The rigor of game theory's formalizations, ironically, were a point of academic and political contestation. Concerning scientific and political legitimacy, "the insistence on 'purity' and logical rigour [was] thought necessary by Morgenstern to preserve economics from political influence" [52, p.141]. Yet for von Neumann, the logical rigor of this mathematics of psychology and social coalition-building was explicitly part of a larger movement of "formalist" mathematics, but also a rejection of a contemporaneous anti-formalist and pro-Nazi integration of psychology into mathematics [52, p.221–223]. Other scholars aligned such formal approaches with pro-democracy approaches, rather than intuition-led fascist framings, "equat[ing] the campaign for logical empiricism with the struggle for social reform" [52, p.122].

As described in the STS literature, a period of interpretive flexibility precedes scientific closure (§3). We can view this contentious period in the American and European mathematical community as a period of *interpretive flexibility* surrounding game theory. By the 1940s, this flexibility was no more, having given way to closure. Game theoretic methods were institutionalized in U.S. policy during the Cold War, with the creation of institutions within and adjacent to the U.S. government that would bring social scientific tools to bear on policy issues. These institutions, like RAND and the National Science Foundation (NSF)'s social sciences division, brought game theory into the center of military decision-making. (Note also the central role of game theory in the popular imagination of the Cold War in films such as *Dr. Strangelove* and *Fail-Safe*, both released in 1964.)

Erickson et al. detail what they call the "ideal type of Cold War rationality," an orientation toward formal reasoning abstracted away from context [33]. This culture of rationality reflected the U.S.'s culture of containment, creating highly controlled "symbolic enclosure[s] within which the (il)logic of nuclear politics played itself out" [31, p. 118]. Early rhetoric around military applications of game theory emphasized many of the same promises highlighted in contemporary discussions around formal privacy. Rather than measuring and tuning a specific value, game theory promised to integrate a host of values captured under the umbrella of "utility," or, in the Cold War context, of "military worth" [32].

Boundaries between science, values, and policy were consistently negotiated, with NSF authorities alternating between rhetoric that emphasized the independence of basic science from political pressure on the one hand and rhetoric that emphasized the political relevance of science on the other [50].

Despite the enthusiasm around game theory's utility, however, critiques from inside and outside of the game theory community grew over time. These critiques mirror several critiques of contemporary FTs, including formal privacy: that they measured the wrong thing, or ultimately failed to capture the real-world impacts that people cared about. Beyond the critiques themselves, it is worth examining the varying responses to these critiques. As the discrepancy between formal and empirical became apparent, different groups took different approaches. One approach involved ignoring this complexity and treating formal models as heuristic rather than reflective of reality. An alternative approach involved making models more complicated to reflect the complexity of the empirical world. Most notably, Herbert Simon introduced the idea of bounded rationality, relaxing the more assumptions of strictly rational actors in favor of a model that tried to capture the limits of people's decision-making capacity. Building on this, Daniel Kahneman and Amos Tversky brought insights from psychology to try to capture the ways that cognitive biases might make people likely to deviate from rationality in practice. While their descriptive findings undermined the empirical accuracy of formal frameworks, Kahneman and Tversky maintained that, normatively, the rational ideal was something to strive for. By highlighting "gulf between the normative and the descriptive"[33], their work on cognitive biases bifurcated the science of description from the ethics of normative decision-making. Yet, other factions eschewed formal modeling entirely, focusing on actionable predictions rather than generalizable theories. This emphasis on management, rather than understanding, [70] reflects a move toward instrumentalism in service of efficient administration.

Game theory tried to create a closed world. While this closure ultimately proved to be fragile, it still had lasting consequences in scoping ethical questions and policy interventions. Notably, game theory offered the foundations for a related but distinct FT that has since been adopted in administrative and everyday contexts. Auction theory and mechanism design, developed in the 1960s and 1970s, have seen more modern, recent wins in scientific consensus and practical adoption (and closure). The Federal Communications Commission, an American administrative agency, ran an auction for electromagnetic spectrum licenses in 1994 in consult with economists in this field; this practice has since been continually renewed and produces billions of dollars in revenue [37, e.g.]. The adoption of these auctions explicitly called upon legitimacy via efficiency (including such monetary rewards; see [75]); however, scholars dispute whether it is efficiency and academic rigor, or instead ethical closure and profiteering, that drove these decisions [60, 76].

5 FTS AS RHETORICAL TECHNOLOGIES

The two use cases above reflect different time points in the research, development, translation, and adoption of FTs. But what they share in common is a set of relationships to the values embedded in U.S. administrative state governance. To investigate these relationships, we reframe the conversation around FTs by expanding our field of view. We claim FTs are not only technological artifacts, but *rhetorical* artifacts: they present information in such a way where its form, as much as its function, persuades stakeholders to prefer its adoption. This reframing does not detract from the technological promise of FTs in helping to solve difficult policy problems, but instead helps contextualize the interplay between *what* FTs contribute to policy discussions and *how* they contribute. By conceptualizing FTs as rhetorical tools, we highlight their performativity [8, 12]: their ability not only to reflect reality, but also to construct it. In this case, FTs meaningfully shape policy by framing political participation around a limited set of technical decisions.

5.1 Rhetorical Technologies

Viewing FTs as rhetorical illuminates the way that FTs are deployed to enable certain decisions, while constraining others. We focus on three rhetorical strategies shared by FTs illustrated in our analysis. First is the rhetoric of *impossibility*, which ties technological impossibility to policy infeasibility. Administrators here make claims that equate the scope of legitimate policy interventions with those legible to FTs. Next is the rhetoric of *formal quantification*, which centers formal, probabilistic, and general risks over others with more contested, context-specific interpretations. Last is the rhetoric of *control*, which emphasizes FT implementation decisions and parameters as the locus of policy negotiation precisely *because* they allow for ex ante control of some policy outcomes. Administrators implicitly assume that policy problems should be centered around FT implementation and parameter choices because they deterministically create cause-effect relationships between policy decisions and outcomes.

First, the impossibility results that motivate FT development and adoption affect the scopes in which policy interventions are deemed legitimate by administrators. In attempting to cull to broad space of possible policy interventions into something more manageable, FTs offer a seductive solution: FT researchers and developers can work on the "science" of policy mechanisms, and simultaneously, policymakers need only decide how to best deploy FTs. In our two examples, the science of policy mechanisms means the spaces of privacy-enhancing technologies with equivalent formal information leakage guarantees, or game theoretic mechanisms with the same formal equilibria properties. Administrators then have an apparent solution to the "paradox of scientific authority" [15].

By creating a clear division between the consensus-driven, scientific, and supposedly objective space where FTs are developed and the contestation-driven, subjective policy space where FTs are implemented, agencies help to insulate their autonomy against external intervention by clearly designating where decision-making should occur, but that limits who participates in that decision-making and how by prioritizing those with the ability to understand, contextualize, and argue for or against FT implementation details. As a rhetorical strategy, the division provided by FTs is highly persuasive to administrators because it adds technocratic weight to the policy alternative scoping process with seemingly minimal modification. When administrators are considering policy alternatives not immediately reconcilable with a particular FT, they're offered two paths of least resistance: either modify or translate desired policy alternatives to make them compatible with FTs, or exclude them as technologically infeasible. The ex ante nature of formal possibility and impossibility suggests that, because one cannot make formal claims about particular policy outcomes, nothing of value is lost in this scoping exercise. Instead, comparisons between policy alternatives on the same formal axes should theoretically enable greater construct validity when comparing FT implementations.

Next, the logic of FTs values specific approaches to quantifying harm: formal risks, i.e. probabilistic risks over randomness inherent to mechanism implementation, are deemed more scientific based on how successfully they abstract away the background assumptions needed for making formal claims, especially when such assumptions may be difficult or impossible to verify in advance. For example, FP methods typically rely on minimal to no assumptions about a hypothetical adversary, which is seen as valuable because such assumptions involve subjective exercises of justification. Therefore for FTs, the fewer assumptions needed, the more general, universal, or scientific a particular formal claim can be. For example, formal privacy guarantees require weak or non-existent assumptions about what adversarial data users know about individual data subjects. Similarly, auction mechanisms require weak or non-existent assumptions about the kinds of information asymmetries that market participants may have. In doing so, FTs clearly delineate between risks that are precisely, mechanistically controlled versus those that must necessarily be entangled with unverifiable assumptions or unknowable future conditions.

Administrators tasked with decision-making under uncertainty see this strategy to risk management as attractive because FTs draw our attention to risks that reduce the administrative burden needed to make high-uncertainty decisions. This reflects a general trend in how the administrative state increasingly makes decisions not about the probabilities of harmful events, but the more possibilities of these events at all [7]. FTs generality provides assurances that are *independent* of these possibilities, meaning the low-probability but high-risk events of interest are provably covered in this new regime. In both our examples, FTs intentionally operate without making assumptions about unknowable background information asymmetry conditions, be they about adversarial prior knowledge for formal privacy or market signals for game-theoretic auctions. This creates the conditions for quantitative risk management that needn't concern itself with possibilities beyond its scope or control.

Lastly, FT ex-ante guarantees are both technically and rhetorically causal, a property lacking from an overwhelming majority of technical claims made about contested ethical values. Because FTs construct their own cause-effect calculus, the mere presence of avenues for control helps to justify prioritizing the resulting metrics that get controlled, regardless of how relevant such metrics are to policy.[1] Privacy loss as a mathematical construct, for example, can be precisely bounded through careful data publishing mechanisms unlike other technologies attempts to limit disclosure risks. However, privacy loss is only one of many metrics to measure data processing's privacy harms, and the connections between privacy loss and more outcome-oriented privacy metrics is much more ambiguous. It is precisely the ability to ex-ante bound privacy loss that motivates making such a metric the locus of negotiation for policy decisions, even if it requires significant contextual interpretation.

Administrators see value in prioritizing these ex ante metrics over others because they can transparently demonstrate the consequences of their interventions and their deterministic impact on policy under uncertainty. Just as with formal quantification, FTs make the managerial task of demonstrating impact trivial by drawing attention towards metrics with precise avenues for control. Contrast this with other policy interventions that require intensive program evaluations, like social benefits programs, for which their quantitative effectiveness can be more readily contested.

5.2 Beyond Law, Beyond Quantification

Prior work has discussed the rhetorical power of law, technology, science, and quantification. A rhetorical study of technology makes visible the reconfiguration that both technology and rhetoric create– creating new connections that reflect, enable, and constrain various forms of agency [53]. As technologies that formalize ethical principles, FTs act first and foremost as rhetorical devices that enact closure over open ethical questions. Both law and quantification have historically played similar rhetorical roles, supporting closure in different ways. FTs go beyond both law and quantification to combine formalizing effects of both. As rhetorical tools, FTs synthesize the strengths of quantification and law to enable closure of open policy problems and contested, value-laden decisions. This rhetorical strength is helpful for advancing bureaucratic goals, aligning with traditional pursuits of legitimacy via non-arbitrariness, efficiency, and accountability

Law enacts closure by creating certainty, with foreseeable consequences for specific events [41]. Similarly FTs create closure by creating policy instruments where inputs map predictably onto outputs. Simultaneously, FTs borrow their

[1] A contemporary of von Neumann and Morgenstern offered a critique of the self-limiting rhetoric of FTs. In correspondence from Herbert Simon to Morgenstern, Simon characterizes the threat of the aesthetic, formal choices of game theory to drive their empirical relevance. In 1945 in response to their seminal work on the *Theory of Games and Economic Behavior*, Simon wrote [excerpted in 52, p.260]:

> "I am not certain that I understand Professor von Neumann's position on this point, but I got the distinct impression from discussion with him that his preference for the definition you used was based largely on aesthetic and formal grounds. Being a social scientist rather than a mathematician, I am not quite willing for (sic) the formal theory to lead the facts around by the nose to quite this extent."

force from the rhetoric of quantification. In their work on the rhetoric of statistics in political discourse, Brendan Lawson and Matthew Lovatt argue that statistical arguments function by eliding the circumstances of their production [51]. By omitting the assumptions and people involved in their design, statistics appear to be factual representations of the natural world, rather than human constructions. We observe this phenomenon in FTs, where abstraction allows administrators to elide the messy work of justifying their assumptions.

Across both the formal privacy and game theory cases, we observe the same pattern of closure. Returning to the closure as consensus formation (recalling the STS concept where we see closure through rhetorical closure and problem definition), we observe that in both cases, FTs 1, leverage the rhetoric of optimization and non-arbitrariness to resolve deliberation and 2, redefine the policy problem to fit within the scope of the FT's parameterization. Moreover, through closure-via-exclusion, we can see that FTs, in fact, constrain modes of participation in policy-making by limiting outside intervention. However, this closure is more fragile than it may seem. We argue that FTs enact a form of closure based not via democratic consensus, but instead via problem redefinition and rhetorical exclusion. Yet, even when the limits of FTs are exposed, these technologies can have durable effects on policymaking approaches. As we observed in the game theory case, technologies not only have the power to change social relations but also naturalize these new social relations such that the process by which the technology was created and adopted disappear [5].

Like law, FTs create closure over open ethical questions. This makes both tools appealing for policymaking, allowing decisions to be made in the face of uncertainty or ambiguity. Despite their similarities, however, law and FTs are not interchangeable. Attending to the differences between these two closure mechanisms is instructive for two reasons. First, we can identify differences in how closure is achieved and, second, we can distinguish between the policy consequences of using law or FTs to intervene in ethical dilemmas.

In liberal democracies, the authority of law and of regulation are legitimated through processes that introduce checks and balances. These checks and balances include democratic elections, judicial oversight, and public participation, for example through the notice-and-comment system. These legitimation tactics emphasize the values of democratic pluralism and create closure primarily through identifying or building consensus across a wide variety of interest groups [69]. Moreover, via the legal system, law is legitimated through argumentation and justification, rather than through appeals to objectivity [71]. However, this strategy is reversed in the case of FTs. Unlike the law, the authority of FTs, is not often subject to deliberation or to checks and balances, and must therefore be legitimated by other means.

We argue that FTs, unlike law and regulation, circumvent the process of consensus building and instead enact closure primarily via rhetorical exclusion. FTs move the consensus-building component of closure to the scientific community and outside of the ambit of policy discussion. This creates closure by preserving the autonomy of technologies and the agency officials involved in its adoption to interpret the meaning of ambiguous ethical questions. For example, the Census Bureau advocated for DP by equating this particular definition of privacy with the confidentiality requirements laid forth in Title 13 of the U.S. Code. Similarly, the technical impossibility results that motivated FP development are being used to advocate for technocratically scoping the definitions of personal information in omnibus privacy laws [6]. These interpretations use the perceived objectivity of scientific consensus to stand in for consensus among the broader set of stakeholders implicated in the policy process. However, the scientific paradigm for consensus-making, unlike political paradigms for consensus-making, is not built on compromise and negotiation but on expert judgment and closed institutions like peer review. Moreover, scientific consensus cannot meaningfully stand in for consensus on the ethical issues implied by policy choices. Yet, the meaning of privacy is heavily contested [56]. Despite widespread contestation, DP rhetorically takes consensus about the meaning of privacy as a given by appealing to consensus within a small scientific community. Consequently, FTs close off contestation about translation and operationalization decisions to the larger public and stakeholders that would traditionally be involved in the policy process. Instead, participation and contestation for this larger group is limited to a narrow set of implementation decisions.

Formalizing technologies extend the legacy of quantification in establishing legitimacy. Similarly to other forms of quantification, formalization leverages the perceived objectivity of numbers in order to limit external contestation [65]. However, the literature on quantification has largely overlooked the different forms that quantification can take. Formalization represents a particular strand of quantification, which has consequences beyond other quantitative technologies. Specifically, FTs distinguish themselves from other forms of quantification through their emphasis on ex ante guarantees rather than empirical measurement. These features shape the process of closure by shaping where consensus occurs.

Prior work in STS shows that scientific expertise often does not produce consensus but instead becomes an object of debate in itself [44]. However, because FTs are not tied to empirics, they are not easily contested on the grounds of evidence. When game theoretic models failed to line up with empirical evidence, it was unclear how to diagnose the problems, with many defenders of game theory arguing that solving these discrepancies was only a matter of creating more complex models.

FTs can still be susceptible to controversy. For example, the Census Bureau's adoption of differential privacy became an object of significant controversy, in part because of value-laden issues like the political role of census data and the meaning of privacy [17, 58]. However, such contestations are rendered out of scope by the rhetorical closure that formalization enacts. Through its capacity to produce closure, FTs force debates to be on their own terms. Even when controversies emerge in the broader public, FTs align squarely with broader administrative trends toward quantitative demonstrations of legitimacy. The adoption of FTs amidst public concern reaffirms that performances of legitimacy are often for and by the state, rather than designed to foster public trust [13]. Specifically, the process of closure enacted by FTs closes off contestation to a small group of experts in order to preserve agency autonomy above democratic accountability.

In doing so, FTs instrumentalize bureaucratic goals over other purported values, emphasizing efficiency as the key to legitimacy. FTs make it easier to justify prioritizing efficiency by rhetorically restricting the scope of decision-making to a single parameterization in an optimization problem. This parameterization often places efficiency front and center. For example, FP mechanisms are deemed "state of the art" by technologists for their ability to maximize data utility at a fixed privacy level, while auction mechanisms are deemed optimal when they maximize price discoverability under fixed participation constraints. This emphasis on efficiency is aligned between bureaucrats and the technologists or scientists often responsible for FT design [14, 16, 72]. However, overlooks a more pluralistic and democratic notion of what values ought to be at the core of political decision-making.

6 CONCLUSION

In this paper, we show that FTs invoke ethical closure by constraining ethics with the force of technology. Specifically, FTs bypass the democratic consensus-making process by limiting the scope of contestation and those who are able to contest to more closely align with expert-driven scientific consensus-making. Relying on scientific consensus to stand in for democratic consensus forecloses the possibility of negotiation which is key to resolving ethical ambiguities in a pluralistic society. Extending beyond law and beyond mere quantification, FTs leverage the rhetorical power of both to constrain the space of legitimate policy participation and intervention. The rhetoric used to justify FTs has consequences for how policy decisions are scoped (i.e. which alternatives are legitimate or not) and how policy alternatives are evaluated (i.e. how to assess which alternatives meet particular ethical end goals). As John Lynch and

William J. Kinsella write, "Rhetorical studies of technology (or technologies) ideally encourage us to identify and reflect upon the moments of decision in technological development writ large. They ask us to examine the choices that we have made during the creation and dissemination of any given technology, which we can hopefully revise or redesign." [53] By reflecting on the rhetorical power of FTs, we recover the hidden political decisions embedded within the work of formalization. Through the process of formalization, policymakers crystallize new relationships between the technical and the social, between science and policy. Viewing FTs as fundamentally rhetorical matters because the line between science and policy is inescapably blurry; we should not sustain illusions that FTs can perform this bifurcation for us [15]. In order to strive for holistic policy evaluation, we must sometimes confront evidence that resists formalization (whether formalization in general or the formalization *du jour*), even when such evidence threatens the legitimacy of FTs. Through productive contestation around ethical values, administrators might endeavor to create closure through democratic consensus-building rather than rhetorical exclusion.

ACKNOWLEDGMENTS

Earlier versions of this work were presented at the 2024 Privacy and Public Policy Conference and the 2024-2025 Rackham Interdisciplinary Workshop on Science, Technology, and Society. We thank the participants of both events for helpful feedback.

REFERENCES

[1] 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 US Census Bureau's Adoption of Differential Privacy. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 1150–1162.
[2] John M Abowd. 2018. Staring-down the database reconstruction theorem. In *Joint statistical meetings, Vancouver, BC*, Vol. 234.
[3] John M Abowd, Tamara Adams, Robert Ashmead, David Darais, Sourya Dey, Simson L Garfinkel, Nathan Goldschlag, Daniel Kifer, Philip Leclerc, Ethan Lew, et al. 2023. *The 2010 Census Confidentiality Protections Failed, Here's How and Why.* Technical Report. National Bureau of Economic Research.
[4] Bruce Ackerman. 1980. *Social justice in the liberal state.* Vol. 401. Yale University Press.
[5] Madeleine Akrich. 1992. The de-scription of technical objects. *Shaping technology/building society. Studies in sociotechnical change* (1992), 205–224.
[6] Micah Altman, Aloni Cohen, Kobbi Nissim, and Alexandra Wood. 2021. What a hybrid legal-technical analysis teaches us about privacy regulation: The case of singling out. *BUJ Sci. & Tech. L.* 27 (2021), 1.
[7] Louise Amoore. 2013. *The politics of possibility: Risk and security beyond probability.* Duke University Press.
[8] John Langshaw Austin. 1975. *How to do things with words.* Harvard university press.
[9] Kenneth A Bamberger. 2008. Normative canons in the review of administrative policymaking. *Yale LJ* 118 (2008), 64.
[10] Kenneth A Bamberger. 2009. Technologies of compliance: Risk and regulation in a digital age. *Tex. L. Rev.* 88 (2009), 669.
[11] Kenneth A Bamberger and Deirdre K Mulligan. 2010. Privacy on the Books and on the Ground. *Stan. L. Rev.* 63 (2010), 247.
[12] Karen Barad. 2003. Posthumanist performativity: Toward an understanding of how matter comes to matter. *Signs: Journal of women in culture and society* 28, 3 (2003), 801–831.
[13] Rodney Barker. 1990. *Political legitimacy and the state.* Oxford University Press.
[14] Elizabeth Popp Berman. 2022. *Thinking like an economist: How efficiency replaced equality in US public policy.* Princeton University Press.
[15] Wiebe E Bijker, Roland Bal, and Ruud Hendriks. 2009. *The paradox of scientific authority: The role of scientific advice in democracies.* MIT press.
[16] Abeba Birhane, Pratyusha Kalluri, Dallas Card, William Agnew, Ravit Dotan, and Michelle Bao. 2022. The values encoded in machine learning research. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 173–184.
[17] danah boyd and Jayshree Sarathy. 2022. Differential perspectives: Epistemic disconnects surrounding the US Census Bureau's use of differential privacy. *Harvard Data Science Review (Forthcoming)* (2022).
[18] Lisa Schultz Bressman. 2003. Beyond accountability: Arbitrariness and legitimacy in the administrative state. *NYUL rev.* 78 (2003), 461.
[19] Vannevar Bush. 1945/2020. Science–The Endless Frontier: A Report to the President on a Program for Postwar Scientific Research (75th anniversary edition). *Washington DC, National Science Foundation* (1945/2020). https://nsf-gov-resources.nsf.gov/2023-04/EndlessFrontier75th_w.pdf
[20] Daryl E Chubin. 1985. Open science and closed science: Tradeoffs in a democracy. *Science, Technology, & Human Values* 10, 2 (1985), 73–80.
[21] Harry M Collins. 1981. Stages in the empirical programme of relativism. , 3–10 pages.

[22] A Feder Cooper and Ellen Abrams. 2021. Emergent unfairness in algorithmic fairness-accuracy trade-off research. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. 46–54.

[23] Ronald Cramer, Ivan Bjerre Damgård, et al. 2015. *Secure multiparty computation*. Cambridge University Press.

[24] Rachel Cummings, Varun Gupta, Dhamma Kimpara, and Jamie Morgenstern. 2019. On the compatibility of privacy and fairness. In *Adjunct publication of the 27th conference on user modeling, adaptation and personalization*. 309–315.

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

[26] Bibiana Duarte-Abadía, Rutgerd Boelens, and Emerson Buitrago. 2023. Neoliberal commensuration and new enclosures of the commons: mining and market–environmentalism governmentalities. *Territory, Politics, Governance* 11, 7 (2023), 1480–1500.

[27] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*. Springer, 265–284.

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

[29] Cynthia Dwork and Aaron Roth. 2014. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science* 9, 3–4 (2014), 211–407.

[30] Lauren B Edelman. 2020. *Working law: Courts, corporations, and symbolic civil rights*. University of Chicago Press.

[31] Paul N Edwards. 1996. *The closed world: Computers and the politics of discourse in Cold War America*. MIT press.

[32] Paul Erickson. 2019. *The world the game theorists made*. University of Chicago Press.

[33] Paul Erickson, Judy L. Klein, Lorraine Daston, Rebecca Lemov, Thomas Sturm, and Michael D. Gordin. 2013. *How Reason Almost Lost Its Mind*. University of Chicago Press, Chicago. https://doi.org/doi:10.7208/9780226046778

[34] Wendy Nelson Espeland and Michael Sauder. 2007. Rankings and reactivity: How public measures recreate social worlds. *American journal of sociology* 113, 1 (2007), 1–40.

[35] Wendy Nelson Espeland and Mitchell L Stevens. 1998. Commensuration as a social process. *Annual review of sociology* 24, 1 (1998), 313–343.

[36] Wendy Nelson Espeland and Mitchell L Stevens. 2008. A sociology of quantification. *European Journal of Sociology/Archives européennes de sociologie* 49, 3 (2008), 401–436.

[37] Patricia Moloney Figliola and Jill C Gallagher. 2023. The Federal Communications Commission's Spectrum Auction Authority: History and Options for Reinstatement. *Congressional Research Service (CRS) Reports and Issue Briefs* (2023), NA–NA.

[38] Walter Bryce Gallie. 1956. Essentially contested concepts. In *Proceedings of the Aristotelian Society*.

[39] Michael Hammond. 2010. What is an affordance and can it help us understand the use of ICT in education? *Education and Information Technologies* 15 (2010), 205–217.

[40] Ursula Hébert-Johnson, Michael Kim, Omer Reingold, and Guy Rothblum. 2018. Multicalibration: Calibration for the (computationally-identifiable) masses. In *International Conference on Machine Learning*. PMLR, 1939–1948.

[41] Mireille Hildebrandt. 2020. Closure: On Ethics, Code, and Law. In *Law for Computer Scientists and Other Folk*. Oxford University Press. https://doi.org/10.1093/oso/9780198860877.003.0011 arXiv:https://academic.oup.com/book/0/chapter/288379040/chapter-pdf/57455832/oso-9780198860877-chapter-11.pdf

[42] Mireille Hildebrandt. 2021. *Machine We Trust. Perspectives on Dependable AI*. MIT Press, Chapter The Issue of Bias. The Framing Powers of Machine Learning.

[43] Ron S Jarmin, John M Abowd, Robert Ashmead, Ryan Cumings-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 of the United States of America* 120, 43 (2023), e2220558120.

[44] Sheila Jasanoff. 1986. *Risk management and political culture*. Vol. 12. Russell Sage Foundation.

[45] Carlee Joe-Wong, Soumya Sen, Tian Lan, and Mung Chiang. 2013. Multiresource allocation: Fairness–efficiency tradeoffs in a unifying framework. *IEEE/ACM Transactions on Networking* 21, 6 (2013), 1785–1798.

[46] Sallie Ann Keller and John M Abowd. 2023. Database reconstruction does compromise confidentiality. *Proceedings of the National Academy of Sciences* 120, 12 (2023), e2300976120.

[47] Christopher T Kenny, Shiro Kuriwaki, Cory McCartan, Evan TR 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 US Census. *Science advances* 7, 41 (2021), eabk3283.

[48] 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. *Harvard Data Science Review* (2022).

[49] Daniel Kifer, John M Abowd, Robert Ashmead, Ryan Cumings-Menon, Philip Leclerc, Ashwin Machanavajjhala, William Sexton, and Pavel Zhuravlev. 2022. Bayesian and frequentist semantics for common variations of differential privacy: Applications to the 2020 census. *arXiv preprint arXiv:2209.03310* (2022).

[50] Daniel Lee Kleinman and Mark Solovey. 1995. Hot Science/Cold War: The National Science Foundation After World War II. *Radical History Review* 1995, 63 (1995), 111–139.

[51] Brendan Lawson and Matthew Lovatt. 2021. Towards a rhetorical understanding of statistics in politics: Quantifying the National Health Service 'Winter Crisis'. *European Journal of Communication* 36, 2 (2021), 110–124.

[52] Robert Leonard. 2010. *Von Neumann, Morgenstern, and the creation of game theory: From chess to social science, 1900–1960.* Cambridge University Press.

[53] John A Lynch. 2013. The rhetoric of technology as a rhetorical technology. *Poroi* 9, 1 (2013).

[54] Andrea Mennicken and Wendy Nelson Espeland. 2019. What's new with numbers? Sociological approaches to the study of quantification. *Annual Review of Sociology* 45, 1 (2019), 223–245.

[55] Shira Mitchell, Eric Potash, Solon Barocas, Alexander D'Amour, and Kristian Lum. 2021. Algorithmic fairness: Choices, assumptions, and definitions. *Annual review of statistics and its application* 8, 1 (2021), 141–163.

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

[57] Deirdre K Mulligan, Joshua A Kroll, Nitin Kohli, and Richmond Y Wong. 2019. This thing called fairness: Disciplinary confusion realizing a value in technology. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–36.

[58] Priyanka Nanayakkara and Jessica Hullman. 2023. What's Driving Conflicts Around Differential Privacy for the U.S. Census. *IEEE Security & Privacy* 21, 5 (2023), 33–42. https://doi.org/10.1109/MSEC.2022.3202793

[59] Dorothy Nelkin. 1978. Threats and promises: Negotiating the control of research. *Daedalus* (1978), 191–209.

[60] Edward Nik-Khah and Philip Mirowski. 2019. The ghosts of Hayek in orthodox microeconomics: Markets as information processors. In *Markets*. meson press.

[61] Donald A Norman. 1999. Affordance, conventions, and design. *interactions* 6, 3 (1999), 38–43.

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

[63] Fabienne Peter. 2007. Rawls' idea of public reason and democratic legitimacy. *Politics and Ethics Review* 3, 1 (2007), 129–143.

[64] Trevor J Pinch and Wiebe E Bijker. 1984. The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social studies of science* 14, 3 (1984), 399–441.

[65] Theodore M Porter. 1995. *Trust in numbers: The pursuit of objectivity in science and public life.* Princeton University Press.

[66] RAND. accessed January 20, 2025. A Brief History of the RAND Corporation. https://www.rand.org/about/history.html

[67] John Rawls. 2020. Political liberalism. In *The New Social Theory Reader*. Routledge, 123–128.

[68] John Rawls. 2020. *A theory of justice: Revised edition.* Harvard university press.

[69] Robert F Rich. 1987. Politics, public policy-making, adn the process of reaching closure. In *Scientific controversies: Case studies in the resolution and closure of disputes in science and technology*, Hugo Tristram Engelhardt and Arthur L Caplan (Eds.). Cambridge University Press, 151–167.

[70] Joy Rohde. 2023. The Contest between Information and Uncertainty: The Cold War Origins of Computational Policy Knowledge. *Public Culture* 35, 3 (2023), 405–416.

[71] Kurt M Saunders. 1994. Law as rhetoric, rhetoric as argument. *Journal of Legal Education* 44, 4 (1994), 566–578.

[72] Morgan Klaus Scheuerman, Alex Hanna, and Emily Denton. 2021. Do datasets have politics? Disciplinary values in computer vision dataset development. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–37.

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

[74] U.S. EPA. accessed October 3, 2025. National Primary Drinking Water Regulations. https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations

[75] Gerald P Vaughan. 2003. FCC Spectrum Auctions. https://wireless.fcc.gov/releases/JV_speech_4_2_03.pdf

[76] Salomé Viljoen, Jake Goldenfein, and Lee McGuigan. 2021. Design choices: Mechanism design and platform capitalism. *Big data & society* 8, 2 (2021), 20539517211034312.

[77] Langdon Winner. 1980. Do artifacts have politics? *Daedalus* 109, 1 (1980), 121–136.

[78] Felix T Wu. 2013. Defining privacy and utility in data sets. *U. Colo. L. Rev.* 84 (2013), 1117.