CounterFAccTual: How FAccT Undermines Its Organizing Principles

Ben L. Gansky*
Arizona State University, School for the Future of Innovation in Society, Human & Social Dimensions in Science & Technology
bengansky@gmail.com

Sean M. Mcdonald
Digital Public
sean@digitalpublic.io

ABSTRACT

This essay joins recent scholarship in arguing that FAccT's fundamental framing of the potential to achieve the normative conditions for justice through bettering the design of algorithmic systems is counterproductive to achieving said justice in practice. Insofar as the FAccT community's research tends to prioritize design-stage interventions, it ignores the fact that the majority of the contextual factors that practically determine FAccT outcomes happen in the implementation and impact stages of AI/ML lifecycles.

We analyze an emergent and widely-cited movement within the FAccT community for attempting to honor the centrality of contextual factors in shaping social outcomes, a set of strategies we term 'metadata maximalism'. Symptomatic of design-centered approaches, metadata maximalism abstracts away its reliance on institutions and structures of justice that are, by every observable metric, already struggling (where not failing) to provide accessible, enforceable rights. These justice infrastructures, moreover, are currently wildly under-equipped to manage the disputes arising from digital transformation and machine learning. The political economy of AI/ML implementation provides further obstructions to realizing rights. Data and software supply chains, in tandem with intellectual property protections, introduce structural sources of opacity. Where duties of care to vulnerable persons should reign, profit incentives are given legal and regulatory primacy. Errors are inevitable and inextricable from the development of machine learning systems.

In the face of these realities, FAccT programs, including metadata maximalism, tend to project their efforts in a fundamentally counter-factual universe: one in which functioning institutions and processes for due diligence in implementation and for redress of harms are working and ready to interoperate with. Unfortunately, in our world, these institutions and processes have been captured by the interests they are meant to hold accountable, intentionally hollowed-out, and/or were never designed to function in today's sociotechnical landscape. Continuing to produce (fair! accountable! transparent!) data-enabled systems that operate in high-impact areas, irrespective of this landscape's radically insufficient paths to justice, given the unavoidability of errors and/or intentional misuse in implementation, and the exhaustively-demonstrated disproportionate distribution of resulting harms onto already-marginalized communities, is a choice - a choice to be CounterFAccTual.

CCS CONCEPTS

• **Applied computing** → Law Applied computing; Sociology General and reference; Evaluation Software and its engineering; Documentation.

KEYWORDS

metadata maximalism, algorithmic realism, failures of FAccT, accountability infrastructures

ACM Reference Format:
Ben L. Gansky* and Sean M. Mcdonald. 2022. CounterFAccTual: How FAccT Undermines Its Organizing Principles. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), June 21–24, 2022, Seoul, Republic of Korea.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3531146.3533241

1 INTRODUCTION

The ACM Conference on Fairness, Accountability, and Transparency's foundationally acontextual framing of normative values undermines the actual achievement of fairness, accountability, or transparency in practice. Normative conditions, like fairness and accountability, cannot be engineered in a vacuum. They are established, interpreted, adjudicated, and enforced in-context - typically, by pre-existing institutions. The FAccT community's research focuses on governing the provenance and design of digital systems, with little attention paid to the contexts within which those systems operate. Much of the FAccT community's research and interventions are conducted without a clear vision for how its work should interoperate with externally-defined social norms, or for governing interaction with the institutions that enforce them in practice. The FAccT Conference's focus on achieving positive social outcomes through research-and-development stage interventions not only misunderstands how social outcomes are realized, they proactively undermine the institutions and systems capable of realizing their goals in-context. In this paper, we focus on ACM FAccT research in terms of its stated goals, methods, and theory of change; we find them dissonant. As a result, we argue, that FAccT, in its current state, is counter-FAccTual.

1.1 The Role of Context in Governance Design

There is a fundamental difference between designing systems for technical function and for social outcomes. Technical solutions can

FAccT '22, June 21–24, 2022, Seoul, Republic of Korea

ACM ISBN 978-1-4503-9352-2/22/06.
https://doi.org/10.1145/3531146.3533241

be achieved by a defined group of actors under controlled conditions. The tendency within the computer sciences to imagine social outcomes as engineerable is consonant with a long history of scientific practice characterized by a framing of success as that which can be accomplished and evaluated within a bounded, rational system (e.g. [24]). The adoption of normative as opposed to technical goals, on the other hand, transposes the question of success or failure into an unbounded, indeterminate space defined in terms of the quality of relationships between a range of entities, inputs, and outputs. In the face of this complexity, in order for a system to know that it is achieving its normative goals, it must have the capacity to monitor the normative implications of its inputs and outputs (due diligence and impact assessment processes, respectively). It must also develop the capacities necessary to identify and, where necessary, adapt behavior and redress unanticipated harms. The nature, scale, and intensity of oversight and adaptation mechanisms are defined in-context, typically based on the scale of impact of the system. In other words, normatively sound governance is not primarily determined by the material or historical properties of the tools, but by the context of their use.

For example, procuring a certified N95 mask is a different kind of task for a hospital administrator in 2022 than for a hobbyist carpenter in 2019; the level of appropriate due diligence in sourcing varies because the potential outcomes resulting from these N95 transactions have such different stakes. The appropriate amount of oversight of the administrator or the carpenter varies by the same logic. In neither case, however, can it justly be the responsibility of the mask manufacturer to determine the necessary due diligence or oversight on their own, because the manufacturer cannot anticipate all the contexts in which their product will be deployed. While the accomplishment of a technical goal (the function of the mask under particular testing conditions) can be effectively assessed within the bounds of the lab or factory, the normative goals towards which the masks will be employed (e.g. safety, care) can only be accomplished and accounted for in the context of their use. The responsibility of the manufacturer is to ensure that their work facilitates the governance needs of their users (including, e.g., both the carpenter and the administrator), as determined by those users and the relevant contextual institutions (e.g. OSHA and the NIH).

Our contention is that data about the provenance of a digital asset is not an acontextually good or just predictor of its future fairness, accountability, or transparency outcomes in practice. The FAccT community's focus is on legitimizing the maximization of data, model, and service reusability (i.e. justifying progressively less-regulated market-based exchange and fewer impediments to scaling). In so far as it continues to invest in collecting and making available provenance (meta)data, FAccT research explicitly ignores more obvious, established, and effective means of achieving its social impact goals.

Our analysis is organized around an illustrative review and critique of a group of allied and widely-cited critical FAccT projects. These we construe as representative of the mainstream of critical FAccT research. We offer three objections to these projects as a means to achieving FAccT outcomes in-practice: (a) the structural incentives towards opacity, competition, and commercial frames within digital supply chains; (b) the convenient illusion of practical pathways to justice for harmed actors; and (c) the fundamental flaw in their theory of change, namely the problematic and broadly self-centered focus on interventions 'in the lab'. We follow this analysis with an alternative frame for FAccT critiques that incorporates both stage and scale of development context, toward restoring the contextual integrity of data/ML governance systems [49].

2 METADATA MAXIMALISM: A CRITIQUE

2.1 Background

As soon as a system adopts normative goals, like fairness, accountability, or transparency, it assumes responsibility for building the capabilities to know whether it's achieving those goals, and for adapting accordingly. Primary among those capabilities are the ability to perceive and evaluate its inbound and outbound interactions with the world through the lens of that norm, i.e. due diligence (monitoring inbound exchanges) and impact assessment (monitoring outbound exchange). Beyond raw awareness, systems also require the capacity to adapt their behavior toward their goals over time, and the independent oversight to provide redress for the ways in which they fail to achieve their goals. These latter capabilities, which provide the ability for responsible parties to evolve the behavior of the system toward its intended goals, we refer to as adaptability.

A further necessary element of systems in pursuit of normative goals is infrastructure for participation. The ability of stakeholders to participate meaningfully in a system is both in and of itself an expression of normative values (i.e. that a system's legitimacy is in some way tied to its responsiveness and accountability to its stakeholders) and a capability that a system must have in order to achieve its normative goals. Without mechanisms for stakeholder participation, in other words, a sociotechnical system risks both illegitimacy and inefficacy. Obviously, the questions of who participates, in what ways, and to what ends are centrally important to the normative outcomes of the system in question (see, e.g. [18]). Participation, then, becomes an (often implicit) dependency on which the effective accomplishment of normative goals relies. The ways in which stakeholders are enabled and disabled from participating - by both endogenous (system design) and exogenous (environmental, contextual) factors - are key determinants of a system's ability to achieve its goals.

Systemic capabilities for oversight, adaptability, and participation are necessarily codependent, and their design must be contextually-determined. Oversight without adaptability is ineffectual, while adaptation without oversight is blind - and unjust (e.g. [60]). Absent a rootedness in the context of application, oversight cannot appreciate the normative significance of its findings and adaptation strategies are unmoored from reality. The appropriateness and design of oversight, adaptability, and participation in digital systems varies based on the normative implications of the impacts of those systems, based on a range of contextual factors like type of use, scale of deployment, and stage of development.

Turning to data-enabled systems, information about the origin and movement of datasets, models, and services is foundational for both due diligence and impact assessment. Provenance information for digital inputs is essential to adjudicate whether a given data-enabled system should be accorded trust within a particular context of use [45]. The same goes for information about outbound

datasets and models' chain of custody, or trajectory ([41, 42]). Both provenance and trajectory information are necessary to secure accountability for the outcomes of such systems (e.g. [48]). These capabilities are, however, not enough. Digital systems whose oversight relies on the use of provenance and trajectory metadata must design those oversight capabilities appropriately for the context in which that system is situated. Further, digital systems without integrated and contextually-determined adaptability and participation architectures are certainly not sufficient to achieve normative goals - especially in the absence of well-functioning independent oversight mechanisms.

An emergent strand of FAccT scholarship has proposed a set of practices and infrastructures for achieving normative goals by adding contextualizing information to datasets and machine-learning models and services (e.g. [4, 5, 9, 20, 29, 30, 47, 57, 62, 74]. We refer to these projects and proposals as *metadata maximalism*. While most, if not all, of the above proposals focus on provenance and trajectory documentation in order to enable transparency as a key normative goal, metadata maximalism is often extended to a range of other proposed normative aims. These include fairness, social acceptance of machine learning in sensitive domains [4], system stability [74], trustworthiness [4, 62], accountability [74], [62], interpretability [74], efficiency of development [29], robustness of results [5], and bias mitigation [5].

At the risk of flattening the differences between these projects, we posit that the theory of change espoused by metadata maximalists is threefold:

- that practices of documentation will sensitize technologists to the ethical dimensions of their work, resulting in a more FAccT supply of datasets, models, and services (supply-side theory),
- that the information encoded into metadata will inform and encourage due diligence practices, enabling consumers to effectively steer supplier practice via the discipline of the market (demand-side theory) (see, e.g. [20], p. 2; [4], p. 8), and
- that subjects of data-enabled systems whose rights have been violated through their interactions with these systems will be sufficiently equipped by metadata to pursue justice for their harms (judicial remedy theory).

In order for the above theories of change to be effective in securing FAccT outcomes, there are a number of things that would have to be true about the world. The researchers and developers of machine learning datasets, models, and services would have to be empowered to design and develop digital objects in such ways that profit and corporate power were subordinate considerations relative to FAccT norms. The data supply chain would have to be structured so as to preserve and contribute to metadata at all points along a digital object's lifecycle, so that metadata was intact and comprehensive by the time that object reaches a prospective user/consumer. Subjects affected by digital systems would know exactly which specific systems were affecting them at particular times and in particular ways. They would have access to metadata-based documentation of those digital objects and systems, and they would be able to leverage that documentation in an adjudication process should their rights be violated by those objects or systems.

In the latter part of this section, we demonstrate that these conditions do not accurately describe the world in which we live. First, however, we explore how, even if these conditions were ground truth, the metadata maximalist approach to facilitating normative goals would still be problematic.

2.2 FAccT's Lab-Centricity

In this section, we join FAccT-critical researchers in pointing towards the limitations of tech ethics practices and discourse that focus inwardly on technologists and their agency (e.g. [7, 23, 63, 67]).

The mainstream of FAccT research proposes to accomplish normative goals by making processes in 'the lab' more just, a disposition we term above the supply-side theory of change. By 'the lab', we mean a specific context: the environments in which technologists research, design, and develop digital objects and systems, with the explicit intention of deploying those systems in a different context. FAccT's lab-centricity licenses technologists to make decisions in the lab about development procedures and system design for ML datasets, models, and services that, in any other context, would come with at least three fundamental governance expectations: (1) an agreement between the human subjects being represented and the entity doing the representing, i.e. consent; (2) architecture for subjects to participate in oversight and adaptation; and (3) a mechanism for adjudicating disputes, adapting norms, and enforcing accountabilities. These are precisely the capabilities left unaddressed by lab-centric interventions.

Making research, development, and design practices more inclusive and self-reflexive, while normatively positive developments for technologists themselves, is at best parallel to the task of restructuring how fairness, accountability, and transparency operate in the context of use for ML-related systems. This is not to say that there are no FAccT stakes within the lab, nor that the lab is so easy to delineate by personnel or geography (e.g. [3, 21, 31, 55, 59]). Our point is to highlight the limits of focusing on and acting within the lab, and more importantly, the hazards of attempting to resolve governance issues there and then. Digital systems exchange inputs and outputs with varied actors and ecosystems in ways that cannot be predicted and controlled for in advance (e.g. [22]). The assumption of the lab as the locus of ethical practice presumes a commitment to the development of certain technologies, precludes the ability of stakeholders to ask first principles questions (i.e. even if a system executed its task perfectly, would it be just?), and channels moral energy away from other more fundamental reforms ([6, 32, 35]).

2.2.1 The lab as a stage of production. The lab is defined not by a location but by its relationship to a stage of the digital product lifecycle: in other words, we would not ask, 'where is the lab?' but rather, 'when is a lab?' [66]. It is wherever sociotechnical objects and systems are being created, prior to being validated in-context or deployed. Characteristically of FAccT approaches, the key tactic of the metadata maximalist projects is to intervene in lab practices, specifically documentation. The proposals intervene at the stages of data collection/creation ([5, 20]), model training [47], integration and product development [29], due diligence among ML services purchasers [4], and at the point of use by operators [62]. [30] addresses the entirety of the research and development pipeline.

What these proposals have in common is a focus on the stages of development that occur prior to any meaningful contact with or deployment into real-world contexts of use at scale, and the inequities that arise as a result. As a result, metadata maximalism forces itself into a position in which it produces documentation of how various interests intersect in the production of ML datasets, models, and services (for instance, interests in proportionate representation and culturally sensitive labels/classifications), without ever directly involving the participation of stakeholders who actually hold these interests.

2.2.2 The lab as a set of interests and actors. FAccT's focus on lab intervention does two things: contain and render opaque to the public normatively-consequential deliberations, and protect and reproduce the concentration of agency within the lab. Metadata maximalist practices are regarded by some of the proposals as a sort of normative therapy for technologists: by performing the labor of metadata collection and documentation, workers become sensitized to the FAccT stakes of their decisions, thereby (presumably) developing better and more FAccT-enabling datasets, models, and services. The degree to which reflective technologists are licensed to adjust their practices within the lab, however, is conditioned and constrained by the business models which fund it (e.g. [70, 71]), along with the business models of the eventual deploying actor. FAccT, then, is able to effectively gate-keep governance agency by construing deliberation as technical and/or proprietary and by constraining technical discussions within parameters which do not threaten shareholder value. In other words, by focusing on perfecting the lab's role in digital system design, instead of focusing on integrating with architectures for broader participation or contextual, institutional governance, the FAccT community prioritizes its own political influence on sociotechnical conditions and ensures that the bounds on its own ethical agency are reproduced throughout the ML ecosystem. The collateral damage includes the influence of those subjects who directly experience these systems, and the institutions historically responsible for realizing governance within them.

Insofar as metadata maximalist proposals intend their work to be functional for particular actors outside of the lab, the 'consumers' of datasets, models, and services for whom maximalist metadata are intended are not the subjects of these resources and systems, but their operators. [62], for instance, proposes a labeling system oriented towards operators of ML services/products (e.g. bank tellers, loan officers). [74]'s 'nutritional label' for algorithmic ranking systems is intended for "developers, regulators and the public" (1). The Dataset Nutrition Label project ([29, 9]) is oriented towards "data specialists. . . all professionals utilizing data in automated decision making systems: data scientists, analysts, machine learning engineers, model developers, artificial intelligence researchers, and a variety of others in this space" ([29], p. 2 in footnote). The FactSheets proposal [4] focuses on consumers of 'AI services'. [5]'s proposal for 'data statements' for natural language processing (NLP) datasets favors an intended audience of technologists and policy-makers (though they mention the possibility of use as well by private citizens who wish to contest ADS recommendations/decisions on due process grounds, see 599). In focusing attention primarily on the needs of technologists and operators, the metadata maximalist schemes empower the powerful at the cost of the participation of affected data subjects.

2.2.3 The lab's dependencies on a counterfactual world. The downstream users of a given dataset, model, or service in the context of contemporary data supply chains are essentially indeterminate (see, e.g. [53]). This is not lost on the metadata maximalist proposers. Since expending energy on ways to potentially constrain this downstream user base would directly undermine the business imperatives of data capitalist firms, FAccT approaches tend instead to frame the unknowability of downstream use as an inevitability which can either be conveniently ignored or assumedly mitigated. Several metadata maximalist projects explicitly state that their motivation is to bless increasingly open (i.e. unregulated) use of datasets, models, or services by offering assurance that prospective users will be sufficiently equipped by the augmented metadata to make valid judgements about their normative stakes in use (e.g. [4]). This move not only implies the sufficiency of lab-based approaches to governance, it simultaneously underwrites governance by open market for datasets, models, and services; willfully ignores the structural incentives of data intermediaries to obfuscate provenance; and fundamentally misconstrues the ways in which governance attempts to achieve justice in practice.

Unfortunately, both the provenance and internal logics of deployed digital systems are not only invisible to those affected by such systems [52], there are rarely any means by which they could discover or participate in improving the underlying conditions, e.g. supply chain issues [72]. That power asymmetry is the nearly universal political economy into which metadata maximalist interventions are deployed - and why focusing on fairness, accountability, and transparency interventions for the comparatively small number of data system operators, as opposed to the large number of people whose rights are impacted by the implementation of data systems, undermines the realization of those norms.

Unless and until the power to adjudicate dataset, model, and service-related conflicts and enforce remedies is devolved to the level of use-context, the centralization of agency of which FAccT is a part will continue to stymie justice. Simply put, digital systems will inevitably produce errors, some of which become harms [48]. That errors can neither be entirely anticipated nor prevented is a long-established truism in computer science [68]. Through its lab-centricity, FAccT nevertheless puts its eggs in the basket of error avoidance; governance in context, on the other hand, tunes itself towards the detection, adjudication, and resolution of such inevitable harms. It has mechanisms for interpreting these harms to inform the adjustment and oversight of a given system so that the harms are not repeated, but learned from. Systems of governance are able to struggle towards progress almost precisely to the extent to which their design and operations are calibrated for the particular normative demands of their context. We argue that the inward focus of metadata maximalism both reproduces unjust concentrations of unaccountable power in 'the lab' of machine learning, and reflects a fundamental misconstrual of the nature of struggles for justice, which depend on rootedness in the places, institutions, and people where oppression is experienced.

2.3 Structural Obstacles

We now turn to a demonstration of the ways in which assumptions about stakeholder participation, on which metadata maximalist theories of change depend, are unfounded. We focus on two obstacles to the metadata maximalist theories of change: the 'supply chain shredder' of metadata (which blocks the demand-side theory of change), and the paucity of accessible pathways to justice for harmed individuals (which undermines the judicial remedy theory of change).

2.3.1 The Supply Chain Shredder. Metadata maximalism reliant on a demand-side theory of change posits that investment in extensive documentation for datasets, models, and services is warranted because that information will make its way to downstream prospective stakeholders, attached to the digital asset in question. Key data supply chains, however, exhibit structural dynamics that act like a shredder for metadata. Without claiming this taxonomy to be all-encompassing, we identify three structural obstacles to the delivery of FAccT-salient metadata to downstream stakeholders: intermediation, reification, and interpolation.

Intermediation. Aggregation of diverse datasets is a commonplace occurrence in many real world data supply chains (e.g. [27, 43, 69]). It is a core business practice of data brokers of varied kinds. When heterogeneous datasets (e.g. public records from varied state sources; location data from CDRs, MAC address sniffing, and GPS) are combined within the 'black boxes' of data brokers, the stripping away of provenance and trajectory metadata is motivated by a rationale of proprietary privilege and competitive trade secrecy. Communicating the painstaking means by which their data assemblages are sourced and integrated would effectively provide data brokers' customers, and potentially competitors, a roadmap to replicating this work themselves. This means that there are business model-driven motivations for data brokers to intentionally undermine the work of diligent metadata maximalists (see, e.g. [12]). To make matters worse, brokers receive much of their data assets from other brokers; rinse and repeat ([12], p. 94). Further, the importance of these practices for the profitability of data brokers means that those corporate actors are legally bound by their fiduciary obligations to their stakeholders to protect these practices (see [36]). Intermediation, through which diverse datasets are standardized, made commensurable, and rendered amnesiac, therefore presents a structural obstacle to metadata maximalism's full achievement of its fairness and equity-enhancing potential, as by the time datasets, models, and services come into contact with operators and subjects, they have been relieved of their metadata.

Reification. When datasets are reduced to a score, ranking, or other simplified metric, the metadata contextualizing the constituent data ingredients are lost. Ranking Facts [74] attempts to address this kind of metadata loss through their 'Recipe' and 'Ingredients' widgets associated with, respectively, the ranking algorithm (for instance, "for a linear scoring formula, each attribute would be listed together with its weight"), and the data features "most material to the ranked outcome" (2). This approach fails to account for two dynamics: one, as mentioned above and detailed below, data brokers are incentivized to obscure, rather than reveal, the 'recipe' and 'ingredients' that compose their products, i.e. scores/ranks. Secondly, these scores themselves travel widely, well beyond the scope of their creators' intended use-cases: for instance, in a United States context, three-digit credit scores have somewhat controversially become a proxy metric for individuals' trustworthiness in contexts such as employment and housing markets [see, e.g. [16, 17, 58]. When these score-outputs become inputs into new algorithmically-mediated systems, the 'chain of custody' for metadata is both broken and obscured.

Interpolation. In various contexts, it may be expedient for data processors to infer information in order to fill a data gap. Sometimes these inferences are performed in order to standardize data whose collection and/or storage may have been irregular; sometimes these inferences are in support of projects that seek to avoid the appearance and/or legal liability of using sensitive or protected data. The increasingly prevalent use of 'synthetic' data in research (e.g. [28, 54]) represents an acceleration of the trend toward interpolation. The obstacle here is not that data is being created through inference; after all, there is no reason in theory why metadata maximalist principles couldn't be applied to document such an inference production process, with ethically-salutary effects for the data's creators, as well as potential downstream users and subjects. The core of the issue here is that certain actors in data supply chains are incentivized towards secrecy and metadata erasure to gain competitive advantage and to avoid potential liability, respectively. Even should data processors be inclined to preserve metadata, the infrastructures and practices necessary to produce supply chain transparency are far from costless. All this means the likelihood of such actors observing metadata maximalist principles in the creation of inferred data is close to zero in the absence of regulatory injunction and enforcement to the contrary (precisely the kinds of interventions not contemplated in FAccT framings).

Together, these structural obstacles to the preservation of labored-over metadata undermine the plausibility of the demand-side theory of change espoused by metadata maximalist projects. They further underline the limitations of FAccT framings that fail to consider the incentive structures of shareholder capitalist business models as key factors in the problem-space of algorithmic oppression.

2.3.2 Pathways to Justice. In the judicial remedy theory of change, access to the kinds of information encoded by metadata maximalism would be helpful to individuals seeking justice for data-enabled harms. But there are serious obstacles precluding the realization of this vision. Every justice system in the world has access-to-justice issues with existing case loads (see, e.g. [73]). (Here, we use the term justice system to refer to both formal and informal dispute adjudication systems - distinguished by the intention to deliver equitable outcomes, as opposed to resolutions in-favor of one parties' interests, like customer service). Adjudicating cases involving technology adds complexity in terms of questions of jurisdiction, novel questions of law, and unprecedented procedural design problems [10]. This lack of digital harm-specific infrastructure for rights enforcement compounds the overall and ongoing crisis in access to justice. As an example of this background condition, in the United States a recent White House commission on Americans' access to justice reported that:

> "[e]ven before the pandemic, a 2017 Legal Services Corporation (LSC) study showed that 86% of the civil legal problems reported by low-income Americans received inadequate or no legal help, and that 71% of low-income households had experienced at least one civil legal problem in the last year. . . [including] issues as crucial as health care, housing conditions, disability benefits, veterans' benefits, and domestic violence. Notably, this figure only includes civil legal problems that are reported in the first place, which are estimated to represent only about 20% of all civil legal challenges. . . [government and nonprofit services] provide approximately one attorney per ten thousand lower-income Americans" [40].

This raises the question: even if maximalist metadata were able to find its way into the hands of individuals subject to data-centric harms, what would they do with it? Giving these individuals information intended to facilitate their empowerment in contexts with demonstrably inadequate systems of rights enforcement only informs them of their helplessness in the face of a harm they can't remediate. This not only breeds disaffection and distrust, it empowers the perpetrators of those harms with broad impunity, as they are able to claim compliance while offloading their responsibilities to the harmed. Privacy scholars (e.g. [50, 65]) have exhaustively demonstrated how the 'notice and consent' paradigm for managing rights in digital systems provides strategic value for data processors by framing consumers as ultimately responsible for the outcomes of engaging with a particular service. This framing claims that consumers' ability to choose among competing services justifies insufficient rights protections within a given digital service.

This is farcical on multiple levels: consumers are neither on average sufficiently literate in algorithmic systems to be able to discern which bits of information about a system are relevant to their interests, nor do they have the time to devote to attaining this literacy and exercising it on a case by case basis all day every day; the idea of choice as responsibility is viable only if there are legitimate differences among the alternatives from which they might choose (not frequently the case), and if individuals have the agency to choose in the first place (questionable given cases in which individuals are involuntarily or unknowingly subject to ADS) [e.g. [14, 51, 52]). Finally, we should be suspicious of these kinds of proposals, as firms have long promoted framings of individual responsibility for collective problems as a strategy for displacing their own culpability (see, e.g., [64]), a move consonant with a neoliberal framing of the entrepreneurial individual as fully responsible for their fate [15].

Informational objects, including metadata, could in theory contribute to securing fairness, accountability, transparency, and trustworthiness in digital systems. But to realize those results, they would have to be operationalized within an ecosystem of reciprocal and accountable relations, aligned incentives between the operators and subjects of data systems, and paths to justice accessible to data subjects and their representatives. These are observably not the social conditions into which algorithmic systems are being integrated. The shortcomings of the metadata maximalist judicial remedy theory of change are exemplary of the ways in which mainstream FAccT fails to consider the over-determining significance of oppression writ large for algorithmic oppression in particular. The struggle for data justice depends upon larger liberation struggles: "[h]ow do we decolonize AI if the world is not even decolonized? . . . [I]f we wish to live in a more just world inside and outside of technology, we must not only abolish algorithmic oppression, but all oppression" [25] (p. 4).

Our next section discusses strategies and structures for contextual governance, derived from a variety of institutional contexts and governance traditions, particularly focusing on justice systems and human rights protection infrastructures. We discuss how, in various situations, these strategies are associated with the presence of particular contextual factors, including scale, nature, and intensity of impact. Finally, we offer suggestions for future directions in research and practice for achieving FAccT outcomes in context.

3 CONTEXTUAL GOVERNANCE

Structurally and substantively, the way that most governance institutions realize basic rights - let alone subjective normative conditions, like fairness, accountability, and transparency - depends on context. Metadata maximalism is motivated by the implicit assumption that a complete contextual history of a dataset, model, or service can predict the normative implications of its use in-context, therefore warranting acontextual trust. This is, unfortunately, diametrically opposed to how the institutions that are designed to uphold our rights and norms operate: our justice institutions are designed to strictly limit the use of an individual's history to make claims about their future behavior. Courts, for example, explicitly limit the amount of information and characteristic inference available, including most of a person's prior history, to juries in the hopes of fairly adjudicating specific cases on their merit.

Contextual factors are the foundation of determining the legitimacy and authority of an institution in a given situation. Legal systems, for example, consider a range of contextual factors when determining which courts have the authority to adjudicate specific claims - including the subject of the dispute, the residence of the parties involved, and the location where disputed actions occurred. Similarly, the scale of a dispute's potential impact can shape the process by which it's adjudicated, the type of relief a court is able to offer, and the degree of specialty required by the court. For example, the amount of money involved in a dispute can determine whether a case is heard by a small claims court instead of a more formal proceeding. Substantively, there are a range of contextual factors that determine the responsibility of the parties involved, the stringency of the standards to which they're held, and the severity of the potential punishments. For example, a doctor is held to a different standard when giving medical advice to a patient than a hairdresser, because the law recognizes that the asymmetry of power is typically larger between a doctor and a patient, as is the likely impact of the doctor's advice (e.g. [19, 46]). Ultimately, the law's recognition of contextual factors in determining the appropriate structure and rules for adjudication, and its careful treatment of contextual information as evidence, is representative of how most institutions protect fairness in and across a nearly infinite range of scales and contexts. It is to systems of governance in context that we now turn for lessons in applied FAccT.

3.1 Putting FAccT Governance (Back) In Context

Recognizing that the FAccT community's focus on acontextual governance interventions undermines its ability to achieve normative outcomes, it follows that shifting focus toward interventions that re-establish the systemic influence of rightsholders in-context could yield positive normative outcomes. While the FAccT community's research - and metadata maximalism, specifically - often aims at helping identify the contextual characteristics of a dataset, model, or service, the mechanisms by which they do so result in a consolidation and concentration of their own agency, as opposed to distributing power in support of independent governance and rights enforcement. The disparity between the FAccT community's intended broad impacts and its comparatively narrow approach to achieving those ends, not only merits interrogation - it provides an opportunity for improvement.

We propose that any strategy for the realization of rights in regards to digital systems must account for two, determinative aspects of contextual governance: (1) stage of development and scale of deployment; and (2) embedded or facilitated pathways to conflict adjudication. When it comes to applied digital systems, the realization of fairness, accountability, and/or transparency requires consideration of the potential for harm - which can arise based on both the stage and scale of development ([26, 56]) and the relative accessibility of independent governance and enforcement mechanisms [33].

Though data and machine learning systems are assuredly distinctive territories for governance, the governance challenges their development poses aren't unique. There are a number of mature industries that have pioneered innovative technologies and approaches, across a range of scales, contexts, and potential for harm. Additionally, critical researchers in participatory design and design justice have, over the last decade, expanded the purview of their work beyond objects to systems (e.g. [11, 39]). As a result, there are a range of participation and governance models designed to map and protect the rights of those interested in, and impacted by, the process of building valuable technologies in high-risk contexts. Those examples don't obviate, or even necessarily mitigate, the inherent politics of product development or scaling ([44, 61]), but they do provide a useful framework for identifying, mapping, and even modeling the rights and harms that commonly arise.

3.1.1 Stage of Development as a Contextual Frame. One of the defining characteristics of FAccT's framing is that it typically focuses on a single stage of development at a time, with the intention of affecting conditions across the entire life cycle of digital system use (though see [30] for a more comprehensive, if still lab-centric, approach). Here, we use a high-level overview of the lifecycle of development, from research to production to use, in order to highlight the governance-relevant aspects of stage and scale change. For the FAccT community's purposes, this list is intentionally illustrative, in the hopes of framing future research and interventions.

Research/Prototyping (design/development). Though research and prototyping processes vary dramatically, this stage of development is typically focused on problem and solution modeling. The research and prototyping stage typically involves identifying requirements, modeling potential solutions, and may include attempts at initial production. Importantly, at this stage, a product is mostly an intellectual undertaking, with typically minor risks or impacts on others. While this isn't always the case, especially for the FAccT community - which has proven the environmental costs, for example, of building and training machine learning models ([6, 13]) - most of the governance frameworks focused on research and prototyping simply require compliance with general, existing legal obligations. In other words, while there are a lot of best practice suggestions about building inclusion and participation into problem modeling and solution design, they rarely become legal considerations unless they violate some other, pre-existing set of laws and rights.

Experimentation and Validation (purpose articulation/fitness testing). The experimentation and validation stage of development, however, is highly regulated - in no small part because of its inherent risk. At the experimentation and validation stage, the researcher moves from modeling into testing the product. This can obviously happen in a range of settings, with a varying consideration given to the potential impact of the test itself. In mature industries that develop products with a recognized impact, whether on the environment, animals, or human well-being, there are a significant range of institutional and ethical frameworks designed to ensure the awareness, agency, and protection of participants.

In biomedicine, extremely unethical experimentation, weaponized against Black and brown bodies, spurred the creation of the Belmont Report and eventually, in the United States, the Common Rule - which established a pathway of escalating requirements that govern the conditions under which publicly funded, qualified professionals can experiment on human subjects. Importantly, these conditions require professionally certified oversight, institutional review, experimenters to accept duties to subjects, and the provision of transparency, proportionality, and accountability mechanisms to subjects. While these conditions aren't necessarily directly transposable to machine learning research, there have been examples of data-centric experiments being performed in public or on human subjects (e.g. [38]). Those cases, though often treated as public relations problems for the technology companies that conducted them, raise significant questions about the ethics, risks, and harms made possible by ungoverned digital experimentation. [4] cites Institutional Review Boards (IRBs) as an antecedent for their metadata maximalism proposal; the irony here is that while IRBs are independent bodies integrated within institutions for research, with oversight authority persisting throughout the conduct of an experimental study, the metadata maximalism format has no such third party empowered to be in continuous governance relations, robbing the structure of its normative efficacy. In other words, the normative power of an IRB doesn't reside in the forms that experimenters fill out; the FAccT qualities come from requiring that the relationship between structurally powerful researchers and vulnerable subjects is reciprocal, respectful, and cautious, with risk-conscious enforcement mechanisms at the ready to keep researchers accountable.

Production/Hosting. Once a product has passed the necessary tests, the next step is, usually, production. At this stage, the product moves from "idea" to "enterprise," requiring or integrating a number of changes. For example, if the researcher was not already affiliated

with a sponsoring entity, they usually start or join a company focused on the production (and/or distribution) of the underlying solution. In addition, the company may have to source production inputs - whether data sources, labor, and/or facilities - each of which are independently regulated. For the purposes of governance, the production stage of development involves the interests and equity associated with compliance, the investment of others (whether capital, labor, or data licenses), and ensuring that the underlying enterprise generally avoids causing harm by virtue of its inputs, processes, and outputs.

This can also mean developing diligence and supply chain oversight systems, in order to avoid direct and indirect harms enabled by production, and, where relevant, the defects caused by imperfect production. Digital supply chain governance is additionally complicated when the underlying product or service is unstable, unregulated, and/or spans potentially conflicting jurisdictional requirements. Once a company starts production, the governance requirements become cumulative. Whereas not every company will maintain continuous experimentation, once a product reaches this stage, it's likely to develop - by commission or omission - governance responses to each of the considerations described in this section and below.

Distribution/Commercialization. Once a product reaches the distribution and commercialization stage, the sponsoring company acquires a new set of governance responsibilities. At a basic level, the company has to consider the compliance and market-access requirements of the places where it would like to operate - for digital systems, which can deploy globally with technical ease, that means engaging with an exponentially larger set of business administration, taxation, and political requirements. While it's likely obvious to the FAccT community, the distribution stage of production also typically involves negotiating with a range of independent businesses responsible for ensuring material aspects of the products stack - whether hosting, third-party certification, or physical distribution logistics. And, of course, at this stage the sponsoring company also needs to establish and govern its relationships with users and/or customers - which involves ensuring the integrity of a range of technical, transactional, legal, and reputational considerations. The primary difference at the distribution and commercialization stage is that the sponsoring company transitions from governing the integrity and impact of its internal operations to managing the requirements and expectations of a broad range of external actors and enabling infrastructure.

Adoption/Use . While implied by production and distribution, once a product is being actively used by one or more people, its producer is at least partially responsible for the impacts of its use. Historically, mature industries were responsible for reasonably setting customer expectations about what its product can functionally do (usually through advertising and sales) and, if a product is potentially dangerous, the conditions or limitations that are necessary for safe use (usually established in its contract guarantees and warranties). Beyond setting expectations, however, a company is also often responsible for the ways that people misuse or abuse their product - especially if those abuses were foreseeable. In the same way that the production and distribution stages require a business to expand its governance to oversee its fairly-negotiated relationships with external companies, once a product is in use, companies have to govern the full range of, often unpredictable, ways in which it could be used.

Scaled Adoption/Use (impact-related). Very few products get to the point of becoming presumptive utilities, however, digital systems - and especially machine learning tools - often reach that stage more quickly than other product categories, if only because of their adoption by large-scale systems. Once a product has reached ubiquity, like the mobile phone, its producers acquire an additional, comparatively large set of governance responsibilities based on the way that other systems rely on that product. For example, digital utility systems will often have to govern their relationship to various political institutions, public health systems, and/or security interests. There's no simple, obvious, or universal approach to developing governance at that scale, but one of the major tensions stems from the centralization of the authority it exerts. While creating mechanisms for participatory governance doesn't preempt scrutiny or accountability, nearly all decisions made at this scale are interpreted as political, with commensurate scrutiny.

While this list isn't comprehensive, it does illustrate the way that governance requirements evolve throughout the process of product development. In doing so, it also demonstrates that popular and legal expectations for the governance of a product change as it invites the reliance of others and grows in impact. These dynamics, while hardly novel, have formative, foundational impacts on the political, commercial, and normative character of a digital system - including on the public and regulatory perceptions of their fairness, accountability, and transparency. In order for the FAccT community to realize those norms, it should invest in research and governance interventions that survive and evolve through the lifecycle of development and are responsive to the scale of impact.

3.2 Institutional and Participatory Interoperability

The above references to design patterns in the structure and practice of contextual governance focus mainly on the capabilities of oversight and adaptability. One of the most challenging aspects, however, of designing the governance of digital systems - particularly machine learning systems - is to design models of participation that help ensure normative outcomes. Whether the goal is to achieve technical or social norms, one of the most common motivations for deploying machine learning and automation systems is to reduce the complexity and friction of human participation. In fact, automation systems have been demonstrated to actively reduce the discretionary authority of public officials [1], or replace them–to disastrous effect [14]. Unfortunately, those 'street-level bureaucrats' are often the primary institutionally-provided mechanism for ensuring the realization of normative goals, like fairness, accountability, and transparency - and their discretion can be critical for handling errors, edge cases, and contextual considerations for rule-based systems [2]. FAccT's metadata maximalism attempts to resolve the resulting issues by perfecting its lab work, instead of designing systems aimed at integrating, if not proactively supporting, the participation of public and justice institutions.

The tendency to design digital systems to route around the influence of institutions, instead of in ways that directly integrate into their governance, has the secondary effect of individualizing responsibility for the ways they fail. Without the awareness, capacity, or active support of institutions, digital systems require individuals to recognize harms; collect evidence - not only of the offending action, but also proof of contractual breach or other wrongs; and then identify the institutions or authorities that may be able to help them seek redress. That's all before any of the proceedings even begin - and, considering the relative capacity, power, and resources of digital system providers, it's impossible to view that degree of individual administrative burden as fair, accountable, or transparent.

That's not to say there's a one-size-fits-all institution with which to integrate as a Panglossian solution, but drawing from the stages of development above and the basic affordances of publicly enforced rights, there are a number of legal principles that FAccT researchers could constructively interrogate as design requirements for product deployment. Additionally, partnerships with existing networks of community organizers and political activists have demonstrated that paying attention to the strategies and infrastructures employed by these groups productively opens new FAccT design spaces for normative intervention (e.g. [34, 37]). This list, like the previous section, is designed as illustrative framing for future governance interventions, toward generating interest, research, and debate from the FAccT community.

3.2.1 Articulation of Purpose and Embedded Pathway to Accountability. Those responsible for digital systems should be able to articulate their purpose, methods, and, ideally, provide an embedded pathway to seek independent accountability for those impacted by digital decision-making systems, directly or indirectly. Though a lot of digital systems address this through terms of service agreements, legalistic contracts are notoriously opaque and inert for users experiencing harm or in need of redress. While it's true that statements of purpose, transparency in reporting, and pathways to accountability aren't legal requirements in service or product categories, it should go without saying that they represent table stakes for FAccT researchers.

3.2.2 Due Process for Rightsholders. While definitions vary across legal jurisdiction, due process rights fundamentally come down to ensuring that systems inform the people they impact of their rights-affecting actions, proactively create an opportunity for an appropriate response, and ensure a basic integrity and proportionality in the performance of its purpose. While there's been considerable ink spilled over things like explainability of machine learning systems, the technical characteristics of a system (including indecipherable complexity) do not historically relieve those responsible from ensuring peoples' basic rights when assessing accountability for resulting harms.

3.2.3 Deployer Duty: Explicit Assumption of Enforceable Responsibilities. One of the primary functions of company structures is to manage liability - a function ostensibly held in check through robust public accountability institutions. Given the complexity facing rights-enforcing institutions, the FAccT community could productively invest in researching ways for digital systems to require, or at least transparently report, on their exposure to accountability. This work is, to an extent, underway in the creation of national and global beneficial ownership registries and, to a lesser extent, the digital system providers willing to contractually guarantee their work - but there's a significant amount to be done in both identifying, reporting, and realizing that accountability in practice.

3.2.4 Provide for Diversity in Self-Representation Capacity. Every system that reaches popular adoption or intermediates a public service impacts the rights of people who are unable to represent their own interests. Many rights-affecting industries include an explicit category of representative, often called a fiduciary, whose purpose is to ensure that those incapable of advocating for themselves - whether due to age, capacity, or infirmity - have help. To the extent that a digital system develops any rights management or participation systems, it also needs a mechanism for assigning representation to those unable to do so themselves, in order for the system to be truly fair, accountable, or transparent to all its users.

We argue that, in order for the FAccT community to move beyond the limits of metadata maximalism, it should focus future research and interventions on two, key thematic areas: (1) how to adapt the deployment of digital systems to proactively address common fairness, accountability, and transparency issues - and design mechanisms for contextually appropriate participation by rightsholders in each stage of ongoing system evolution; and (2) methods for establishing and facilitating interoperability with external, independent governance institutions and infrastructure as a functional requirement for deployment of data-centric systems.

4 CONCLUSION

There are not only practical approaches to implementing contextual systems of governance, there are existing models across a range of vital service industries. Almost all of them center, in some way, structures and rules for devolving authority to local institutions - and there is a significant opportunity for FAccT to focus on ways to architect systems toward enabling that work. In order for the FAccT community to achieve its goals, it should start by recognizing (a) that FAccT principles are realized almost entirely outside of the lab; (b) that the institutions and processes for safeguarding those principles in practice are contextual, devolved, and evolving; and (c) while harm-based diligence and remediation efforts are valuable, if these capabilities are designed and intended to primarily inform governance within the lab, they still function to centralize agency away from rightsholders.

If the FAccT community is serious about actualizing its organizing principles in the implementation of digital systems, it is better served by designing structures of participatory self-governance for rightsholders, in addition to the developers and operators currently targeted by its interventions. The realization of fairness, accountability, and transparency are far more commonly the product of messy, participatory governance than technocratic experts perfecting digitally abstracted governance processes. That's especially true when that governance lacks the adaptive capacity and accountability created by accessible paths to justice.

Finally, continuing to develop and deploy datasets, models, and services whose inevitable errors, unpredictable proliferation, and

unaccountable use will harm subjects without access to rights enforcement is a profoundly cynical form of neglect. In an effort to support the FAccT community's pursuit of its normative goals, we argue that it should continue to constructively invest in developing ways to identify and reconstruct the contextual factors necessary for justice. We argue, especially, for FAccT researchers to consider the full life cycle of their product's potential impact, and design for adaptive governance throughout. In addition, we recognize that FAccT's ideals aren't computable, nor is upholding basic rights an internally perfectible endeavor - and so we also argue for research into methods of direct, participatory connections to contextually appropriate institutions as a functional requirement for digital system deployment. At this stage of maturity, neither novelty nor naivete is an excuse: as Deb Chachra [8] puts it: "Any sufficiently advanced negligence is indistinguishable from malice."

5 HISTORY DATES

Received January 2022; revised May 2022; accepted May 2022.

ACKNOWLEDGMENTS

Both authors wish to thank our anonymous reviewers for helpful structural feedback and suggested references. Ben Gansky wishes to also thank Marcello di Bello for his thoughtful and generous engagement with this paper, including guidance at an early stage. Any errors, omissions, and/or misconstruals are of course our own. Ben Gansky's contributions to this paper are based upon work supported by the National Science Foundation under Grant No. 1828010; Sean McDonald's contributions are supported by a grant from the Ford Foundation.

REFERENCES

[1] doaa Abu Elyounes. 2020. 'Computer Says No!': The Impact of Automation on the Discretionary Power of Public Officers (September 14, 2020). Vanderbilt Journal of Entertainment & Technology Law, Forthcoming, Available at SSRN: https://ssrn.com/abstract=3692792

[2] Ali Alkhatib and Michael Bernstein. 2019. Street–Level Algorithms: A Theory At The Gaps Between Policy and Decisions. In Proceedings of CHI Conference on Human Factors in Computing Systems (CHI '19), April 2019, Glasgow, Scotland UK. ACM, New York, NY, USA 13 Pages. https://doi.org/10.1145/3290605.3300760

[3] Sareeta Amrute & Luis Felipe R. Murillo. 2020. Introduction: Computing in/from the South. Catalyst: Feminism, Theory, Technoscience, 6(2). DOI: https://doi.org/10.28968/cftt.v6i2.34594

[4] Matthew Arnold, Rachel K. E. Bellamy, Michael Hind, Stephanie Houde, Sameep Mehta, Aleksandra Mojsilovic, Ravi Nair, Karthikeyan Natesan Ramamurthy, Darrell Reimer, Alexandra Olteanu, David Piorkowski, Jason Tsay, Kush R. Varshney. 2018. FactSheets: Increasing trust in AI services through supplier's declarations of conformity. arXiv:1808.07261v2. DOI: https://doi.org/10.48550/arXiv.1808.07261

[5] Emily M. Bender and Batya Friedman. 2018. Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science. Transactions of the Association for Computational Linguistics, 6:587–604.

[6] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 610–623. https://doi.org/10.1145/3442188.3445922

[7] Abebe Birhane, Pratyusha Kalluri, Dallas Card, William Agnew, Ravit Dotan, & Michelle Bao. 2021. The values encoded in machine learning research. https://arxiv.org/abs/2106.15590. DOI: https://doi.org/10.48550/arXiv.2106.15590

[8] Deb Chachra. 2017. I don't think I fully realized how widely and sadly relevant this would be when I said it. [quoted here by @nd_kane]. Twitter. Retrieved January 19, 2022, from https://twitter.com/debcha/status/874953796794929152?s=20

[9] Kasia S. Chmielinski, Sarah Newman, Matt Taylor, Josh Joseph, Kemi Thomas, Jessica Yurkofsky, and Yue Chelsea Qiu. 2022. The Dataset Nutrition Label (2nd Gen): Leveraging Context to Mitigate Harms in Artificial Intelligence. CoRR abs/2201.03954, (2022). Retrieved from https://arxiv.org/abs/2201.03954

[10] Julie E. Cohen. 2021. From Lex Informatica to the Control Revolution. Berkeley Technology Law Journal, 35.

[11] Sasha Costanza-Chock. 2020. Design Justice: Community-Led Practices to Build the World We Need. MIT Press. Available at: https://design-justice.pubpub.org/

[12] Matthew Crain. 2018. The limits of transparency: Data brokers and commodification. New Media & Society 20, 1 (2018), 88–104. DOI:https://doi.org/10.1177/1461444816657096

[13] Roel Dobbe & Meredith Whittaker. 2019. AI and Climate Change: How they're connected, and what we can do about it. Medium. Retrieved January 22, 2022, from https://medium.com/@AINowInstitute/ai-and-climate-change-how-theyre-connected-and-what-we-can-do-about-it-6aa8d0f5b32c

[14] Virginia Eubanks. 2018. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press. 2018.

[15] Michel Foucault. 2008. The birth of biopolitics: Lectures at the collège de France, 1978–1979, London: Palgrave Macmillan.

[16] Marion Fourcade & Kieran Healy. 2013. Classification situations: Life-chances in the neoliberal era. Accounting, Organizations and Society 38, 8 (2013), 559–572. Retrieved from https://EconPapers.repec.org/RePEc:eee:aosoci:v:38:y:2013:i:8:p:559-572

[17] Marion Fourcade & Kieran Healy. 2016. Seeing like a market. Socio-Economic Review 15, 1 (December 2016), 9–29. DOI:https://doi.org/10.1093/ser/mww033

[18] Archon Fung. 2006. Varieties of Participation in Complex Governance. Public Administration Review 66, (2006), 66–75. Retrieved May 8, 2022 from http://www.jstor.org/stable/4096571

[19] Tamar Frankel. 2011. Fiduciary Law in the Twenty-First Century, 91 Boston University Law Review 1289 (2011). Available at: https://scholarship.law.bu.edu/faculty_scholarship/714

[20] 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 (December 2021), 86–92. https://doi.org/10.1145/3458723

[21] Mary Gray & Siddarth Suri. 2019. Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass. Houghton Mifflin Harcourt Publishing Company.

[22] Ben Green and Salomé Viljoen. 2020. Algorithmic realism: expanding the boundaries of algorithmic thought. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20). Association for Computing Machinery, New York, NY, USA, 19–31. https://doi.org/10.1145/3351095.3372840

[23] Daniel Greene, Anna Lauren Hoffmann, & Luke Stark. 2019. Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning. In Proceedings of the 52nd Hawaii international conference on system sciences. Available at: http://hdl.handle.net/10125/59651

[24] Ian Hacking. 2002. Historical Ontology. Harvard University Press.

[25] Lelia Marie Hampton. 2021. Black Feminist Musings on Algorithmic Oppression. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 1. https://doi.org/10.1145/3442188.3445929

[26] Alex Hanna & Tina M. Park. 2020. Against scale: Provocations and resistances to scale thinking. arXiv preprint arXiv:2010.08850.

[27] Katie Harron, Chris Dibben, James Boyd, Anders Hjern, Mahmoud Azimaee, Mauricio L Barreto, and Harvey Goldstein. 2017. Challenges in administrative data linkage for research. Big Data & Society 4, 2 (2017), 2053951717745678. DOI:https://doi.org/10.1177/2053951717745678

[28] Markus Hittmeir, Andreas Ekelhart, and Rudolf Mayer. 2019. On the Utility of Synthetic Data: An Empirical Evaluation on Machine Learning Tasks. In Proceedings of the 14th International Conference on Availability, Reliability and Security (ARES '19). Association for Computing Machinery, New York, NY, USA, Article 29, 1–6. https://doi.org/10.1145/3339252.3339281

[29] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, & Kasia Chmielinski. 2018. The dataset nutrition label: A framework to drive higher data quality standards.

[30] 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 (FAccT '21). Association for Computing Machinery, New York, NY, USA, 560–575. https://doi.org/10.1145/3442188.3445918

[31] Lilly Irani. 2015. Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk. South Atlantic Quarterly; 114 (1): 225–234. doi: https://doi.org/10.1215/00382876-2831665

[32] Maximilian Kasy and Rediet Abebe. 2021. Fairness, Equality, and Power in Algorithmic Decision-Making. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 576–586. https://doi.org/10.1145/3442188.3445919

[33] Ethan Katsh & Orna Rabinovich-Einy. 2017. Digital Justice: Technology and the Internet of Disputes. Oxford University Press. Retrieved 22 Jan. 2022, from https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780190464585.001.0001/acprof-9780190464585.

[34] Michael Katell, Meg Young, Dharma Dailey, Bernease Herman, Vivian Guetler, Aaron Tam, Corinne Bintz, Daniella Raz, and P. M. Krafft. 2020. Toward situated interventions for algorithmic equity: lessons from the field. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20). Association for Computing Machinery, New York, NY, USA, 45–55. https://doi.org/10.1145/3351095.3372874

[35] Os Keyes, Jevan Hutson, and Meredith Durbin. 2019. A Mulching Proposal: Analysing and Improving an Algorithmic System for Turning the Elderly into High-Nutrient Slurry. In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19). Association for Computing Machinery, New York, NY, USA, Paper alt06, 1–11. https://doi.org/10.1145/3290607.3310433

[36] Lina Khan & David E. Pozen. 2019. A skeptical view of information fiduciaries. Harv. L. Rev., 133, 497. Available at: https://harvardlawreview.org/2019/12/a-skeptical-view-of-information-fiduciaries/

[37] P. M. Krafft, Meg Young, Michael Katell, Jennifer E. Lee, Shankar Narayan, Micah Epstein, Dharma Dailey, Bernease Herman, Aaron Tam, Vivian Guetler, Corinne Bintz, Daniella Raz, Pa Ousman Jobe, Franziska Putz, Brian Robick, and Bissan Barghouti. 2021. An Action-Oriented AI Policy Toolkit for Technology Audits by Community Advocates and Activists. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 772–781. https://doi.org/10.1145/3442188.3445938

[38] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. Proceedings of the National Academy of Sciences 111, 24 (2014), 8788–8790. DOI:https://doi.org/10.1073/pnas.1320040111

[39] Christopher A Le Dantec and Carl DiSalvo. 2013. Infrastructuring and the formation of publics in participatory design. Social Studies of Science 43, 2 (2013), 241–264. DOI:https://doi.org/10.1177/0306312712471581

[40] Legal Aid Interagency Roundtable. 2021. Access to Justice in the Age of COVID-19: A Roundtable Report. Available at: https://www.justice.gov/ag/page/file/1445356/download

[41] Sabina Leonelli. 2016. Data-centric Biology: A Philosophical Study. University of Chicago Press.

[42] Sabina Leonelli. 2019. Data Governance is Key to Interpretation: Reconceptualizing Data in Data Science. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.17405bb6

[43] Sabina Leonelli & Niccolò Tempini, eds. 2020. Data Journeys in the Sciences. Springer, Cham. https://doi.org/10.1007/978-3-030-37177-7

[44] Wendy Liu. 2020. Abolish Silicon Valley: How to Liberate Technology from Capitalism. Repeater Books.

[45] Sean M. McDonald. 2021. Our (Mis)Represented Digital Rights. Centre for International Governance Innovation. https://www.cigionline.org/articles/our-misrepresented-digital-rights/.

[46] Paul B. Miller and Andrew S. Gold. 2015. Fiduciary Governance. 57 Wm. & Mary L. Rev. 513 (2015). Available at: https://scholarship.law.wm.edu/wmlr/vol57/iss2/4

[47] 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

[48] Helen Nissenbaum. 1996. Accountability in a computerized society. Sci Eng Ethics 2, 25–42. https://doi.org/10.1007/BF02639315

[49] Helen Nissenbaum. 2009. Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford University Press.

[50] Helen Nissenbaum. 2011. A Contextual Approach to Privacy Online. Daedalus 140 (4): 32–48. DOI: https://doi.org/10.1162/DAED_a_00113

[51] Jonathan A. Obar. 2020. Sunlight alone is not a disinfectant: Consent and the futility of opening Big Data black boxes (without assistance). Big Data & Society, 7(1), 2053951720935615.

[52] Frank Pasquale. 2015. The Black Box Society: the secret algorithms that control money and information. Harvard University Press.

[53] Kenny Peng, Arunesh Mathur, & Arvind Narayanan. 2021. Mitigating dataset harms requires stewardship: Lessons from 1000 papers. arXiv preprint arXiv: 2108.02922.

[54] Taoxin Peng and Alexander Telle. 2018. A tool for generating synthetic data. In Proceedings of the First International Conference on Data Science, E-learning and Information Systems (DATA '18). Association for Computing Machinery, New York, NY, USA, Article 22, 1–6. https://doi.org/10.1145/3279996.3280018

[55] Jean-Cristophe Plantin. 2021. The data archive as factory: Alienation and resistance of data processors. Big Data & Society. https://doi.org/10.1177/20539517211007510

[56] Sebastian Pfotenhauer, Brice Laurent, Kyriaki Papageorgiou, & Jack Stilgoe. 2022. The politics of scaling. Social Studies of Science, 52(1), 3–34. https://doi.org/10.1177/03063127211048945

[57] Deborah I. Raji & Jingying Yang. 2019. ABOUT ML: Annotation and benchmarking on understanding and transparency of machine learning lifecycles. arXiv preprint arXiv:1912.06166.

[58] Akos Rona-Tas. 2017. The off-label use of consumer credit ratings. Historical Social Research/Historische Sozialforschung, 52-76. https://doi.org/10.12759/hsr.42.2017.1.52-76

[59] Morgan Klaus Scheuerman, Alex Hanna, and Emily Denton. 2021. Do Datasets Have Politics? Disciplinary Values in Computer Vision Dataset Development. Proc. ACM Hum.-Comput. Interact. 5, CSCW2, Article 317 (October 2021), 37 pages. https://doi.org/10.1145/3476058

[60] James C. Scott. 1999. Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed. Yale University Press.

[61] Nick Seaver. 2021. Care and Scale: Decorrelative Ethics in Algorithmic Recommendation. Cultural Anthropology 36, no. 3: 509–537. https://doi.org/10.14506/ca36.3.11.

[62] Christin Seifert, Stefanie Scherzinger, and Lena Wiese. 2019. Towards Generating Consumer Labels for Machine Learning Models. In 2019 IEEE First International Conference on Cognitive Machine Intelligence (CogMI), 173–179. DOI:https://doi.org/10.1109/CogMI48466.2019.00033

[63] 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 (FAT* '19). Association for Computing Machinery, New York, NY, USA, 59–68. https://doi.org/10.1145/3287560.3287598

[64] Rebecca Solnit. 2021. "Big oil coined 'carbon footprints' to blame us for their greed. Keep them on the hook". The Guardian, https://www.theguardian.com/commentisfree/2021/aug/23/big-oil-coined-carbon-footprints-to-blame-us-for-their-greed-keep-them-on-the-hook.

[65] Daniel Solove. 2013. Introduction: Privacy Self-Management and the Consent Dilemma. Harvard Law Review Vol. 126, 1880-1903. Available at: https://harvardlawreview.org/2013/05/introduction-privacy-self-management-and-the-consent-dilemma/

[66] Susan Leigh Star and Karen Ruhleder. 1996. Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. Information Systems Research 7, 1 (1996), 111–134. Retrieved May 8, 2022 from http://www.jstor.org/stable/23010792

[67] Luke Stark & Anna Lauren Hoffmann. 2019. "Data Is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture," Journal of Cultural Analytics. DOI: 10.22148/16.036

[68] Alan M. Turing. 1937. On Computable Numbers, with an Application to the Entscheidungsproblem. Proceedings of the London Mathematical Society s2-42, 1 (1937), 230–265. DOI:https://doi.org/https://doi.org/10.1112/plms/s2-42.1.230

[69] Fernando N van der Vlist. 2016. Accounting for the social: Investigating commensuration and Big Data practices at Facebook. Big Data & Society 3, 1 (2016), 2053951716631365. DOI:https://doi.org/10.1177/2053951716631365

[70] Ari Ezra Waldman. 2021. Industry Unbound: The Inside Story of Privacy, Data, and Corporate Power. Cambridge University Press. https://doi.org/10.1017/9781108591386

[71] Meredith Whittaker. 2021. The steep cost of capture. Interactions 28, 6 (November - December 2021), 50–55. DOI:https://doi.org/10.1145/3488666

[72] World Economic Forum. 2021. Pathways to Digital Justice. World Economic Forum. Available at: https://www.weforum.org/whitepapers/pathways-to-digital-justice

[73] World Justice Project. 2019. Global Insights on Access to Justice: Findings from the World Justice Project General Population Poll in 101 Countries. World Justice Project. ISBN: 978-0-9964094-4-5

[74] Ke Yang, Julia Stoyanovich, Abolfazl Asudeh, Bill Howe, H. V. Jagadish, and Gerome Miklau. 2018. A Nutritional Label for Rankings. In Proceedings of the 2018 International Conference on Management of Data, SIGMOD Conference 2018, Houston, TX, USA, June 10-15, 2018, ACM, 1773–1776. DOI:https://doi.org/10.1145/3183713.3193568