Taking Algorithms to Courts: A Relational Approach to Algorithmic Accountability

Jacob Metcalf
Data & Society Research Institute, USA
jake.metcalf@datasociety.net

Emanuel Moss
Intel Labs, USA
emanuel.moss@intel.com

Ranjit Singh
Data & Society Research Institute, USA
ranjit@datasociety.net

Emnet Tafesse
Data & Society Research Institute, USA
emnet@datasociety.net

Elizabeth Anne Watkins
Intel Labs, USA
Elizabeth.Watkins@intel.com

ABSTRACT

In widely used sociological descriptions of how accountability is structured through institutions, an "actor" (e.g., the developer) is accountable to a "forum" (e.g., regulatory agencies) empowered to pass judgements on and demand changes from the actor or enforce sanctions. However, questions about structuring accountability persist: why and how is a forum compelled to keep making demands of the actor when such demands are called for? To whom is a forum accountable in the performance of its responsibilities, and how can its practices and decisions be contested? In the context of algorithmic accountability, we contend that a robust accountability regime requires a triadic relationship, wherein the forum is also accountable to another entity: the public(s). Typically, as is the case with environmental impact assessments, public(s) make demands upon the forum's judgements and procedures through the courts, thereby establishing a minimum standard of due diligence. However, core challenges relating to: (1) lack of documentation, (2) difficulties in claiming standing, and (3) struggles around admissibility of expert evidence on and achieving consensus over the workings of algorithmic systems in adversarial proceedings prevent the public from approaching the courts when faced with algorithmic harms. In this paper, we demonstrate that the courts are the primary route—and the primary roadblock—in the pursuit of redress for algorithmic harms. Courts often find algorithmic harms non-cognizable and rarely require developers to address material claims of harm. To address the core challenges of taking algorithms to court, we develop a relational approach to algorithmic accountability that emphasizes not what the actors do nor the results of their actions, but rather how interlocking relationships of accountability are constituted in a triadic relationship between actors, forums, and public(s). As is the case in other regulatory domains, we believe that impact assessments (and similar accountability documentation) can provide the grounds for contestation between these parties, but *only* when that triad is structured such that the public(s) are able to cohere around shared experiences and interests, contest the outcomes of algorithmic systems that affect their lives, and make demands upon the other parties. Where courts now find algorithmic harms non-cognizable, an impact assessment regime can potentially create procedural rights to protect substantive rights of the public(s). This would require algorithmic accountability policies currently under consideration to provide the public(s) with adequate standing in courts, and opportunities to access and contest the actor's documentation and the forum's judgments.

CCS CONCEPTS

• **Social and professional topics** → Computing/technology policy; Government technology policy; Governmental regulations.

KEYWORDS

algorithmic accountability, relationality, algorithmic impact assessment, algorithmic governance

ACM Reference Format:
Jacob Metcalf, Emanuel Moss, Ranjit Singh, Emnet Tafesse, and Elizabeth Anne Watkins. 2023. Taking Algorithms to Courts: A Relational Approach to Algorithmic Accountability. In *2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23), June 12–15, 2023, Chicago, IL, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3593013.3594092

1 INTRODUCTION

An expanding range of disciplines are concerned with how algorithmic systems may produce harm, and how such harm is identified and remedied. Whether organized around the pursuit of 'trustworthy AI' [19, 46, 50], 'data ethics' [126], 'algorithmic fairness' [31, 67, 77], or 'responsible innovation' [18, 62, 107], these practices have focused on understanding, documenting, mitigating and ultimately avoiding the dangers that algorithmic systems may present to individuals, communities, institutions, ecosystems, and society writ large.

This work, however, is not without its own set of challenges around how knowledge about algorithmic harms is produced, who is involved in adjudicating its legitimacy, and how the public gets access to it. This is evident in most public controversies centered on ethics of computing, such as whistleblower Frances Haugen's claim that Instagram's internal research showed how some of its features

FAccT '23, June 12–15, 2023, Chicago, IL, USA

ACM ISBN 979-8-4007-0192-4/23/06.
https://doi.org/10.1145/3593013.3594092

harm adolescents [51], Google's treatment of Timnit Gebru and Margaret Mitchell after their research critically examined Google's own large NLP modeling [104], and the fraught relationship between technology companies and academic researchers [66]. Amidst these controversies, multiple jurisdictions have proposed regulations and legislation that would require developers to self-study the potential consequences of the systems they build [9, 26, 47, 79, 83, 96, 123]. These first steps towards regulation are crucial; they produce conditions for broader inquiry into not only *what* is documented about these systems and *who* is mandated to disclose such documentation but also *who* is empowered to receive it and make demands upon developer(s) to change their systems.

"Algorithmic accountability" has been used as a catch-all term to encapsulate this broad set of relations. For example, Mark Bovens has conceptualized the dyadic relationship between *actors* who build systems, and the *forums* that can demand changes in such systems from actors as the primary structure of maintaining organizational accountability [14, 75, 78, 122]. Unpacking the current proposals for regulating algorithmic systems from the lens of this dyadic relationship, developers (as actors) produce documentation and disclose it, while a regulatory body (as a forum) is empowered to receive such documentation and demand changes. However, questions of accountability persist: *why and how is a forum compelled to keep making demands of the actor when such demands are called for?*

In this paper, we begin with the premise that the forum needs its own "forum." For regulators of algorithmic systems, this secondary "forum" is ultimately the public(s) with a stake in the workings of such systems as they live with their consequences. John Dewey points out that there are occasions when an otherwise "amorphous and unarticulated" [37:131] public organizes itself to express its interests in the face of problems and/or issues that affect them. Such publics are brought together relationally, through the same discursive formations that produce problems that evoke public concern [118]. Perceiving these problems, Dewey argues, often requires specialized expertise, and by acting upon them the public manifests its capacity to act as a forum to hold the government accountable. An obvious mechanism to enact such accountability is the ballot box. However, in this paper we focus on another way in which algorithmic publics—constituted around shared interests in the operation and consequences of algorithmic systems—might hold regulatory bodies (forums) accountable, and by proxy hold developers (actors) responsible: the courts.

Courts have often served as a crucial backstop to functions of the administrative state, such as assessment and documentation regimes for complex and potentially harmful systems. Given the varied conditions of jurisprudence over access to courts across countries, our core focus in this paper is on courts in the United States, except when comparisons with other jurisdictions are generative. For example, major developments or economic policies in the US are often contested on environmental grounds based on procedural rights of opponents to have access to accurate documentation and have the negative case (i.e., do not proceed with the development) heard and considered. *The primary consequence of access to courts is not that every decision is contested and adjudicated, rather it is to ensure that procedural rights (such as the right to secure due process) adequately safeguard substantive rights (such as civil rights)* [117]. Thus, the courts are a crucial site for determining the lower threshold of due diligence for algorithmic accountability. Expanding access to courts is likely to raise this threshold, however, this is easier said than done [97]. Indeed, in the emerging and pending regulatory and legislative rules, the role of courts is largely implicit, and these rules often under-specify how the citizenry will be empowered to do more than simply know about these systems; this analysis points to a potential pitfall for algorithmic accountability measures.

In this paper, we build on our prior historical research on impact assessment regimes through a sociotechnical lens [75, 78, 119]. We surveyed available materials in governance and legislation around impact assessments in the US, such as bills, federal agency guidelines, and impact statements, available critiques on such regulatory interventions from legal and sociological disciplines, and contemporary proposals around assessments for algorithmic systems, including methods for internal and independent audits and end-to-end frameworks. Based on this ongoing research, we map the challenges faced by and reflect on the role of courts in structuring accountability relations to foreground public interest in the development of algorithmic systems. We begin with exploring a brief typology of algorithmic harms and whether they can be contested in courts. Taking this typology as a point of departure, we explore (1) the relational nature of algorithmic accountability; and (2) the three core challenges (related to documentation, standing, and expertise) of taking algorithms to court. To address these challenges, we conclude with (1) elaborating on a *relational* approach to algorithmic accountability that emphasizes interlocking relationships between actor(s), forum(s), and public(s); and (2) providing a framework for understanding how governance regimes, such as the algorithmic impact assessments, can effectively structure these interlocking accountability relationships to foreground the public interest.

2 TYPOLOGY OF ALGORITHMIC HARMS AND RELATED COURT CASES

In exploring a brief typology of algorithmic harms, our purpose is not to be exhaustive. This section illustrates how algorithmic harms that seem legible and intuitive in the context of understanding the workings of algorithmic systems must be formulated in specific ways to be cognizable in courts. Our focus is on cases where the issue before the court is liability and redress when a plaintiff is faced with algorithmic harms, not on how algorithms are used in criminal trials as evidence or resource in bail decisions (although there is evidence of potential harm in such cases as well [28, 84, 91]). In the broadest sense, a courtroom is a site for reenacting a sequence of events that led to a contest of accountability requiring adjudication: "someone must be blamed, someone punished, someone rewarded for exceptional enterprise, someone, if possible, made whole" [60]. In the adversarial setting of U.S. courts, cross-examination becomes a means to establish ground truth and facticity around the sequence of events. This process becomes even more challenging when the courts work towards ascertaining consensus over what constitutes legitimate research practice and facts in any technoscientific field, including the computational sciences, before making a judgment on the sequence of events in dispute. Thus, taking algorithms to court is a challenge at every step, ranging from claiming that a harm has been done and an algorithmic system caused it to determining who should be held responsible for the workings of the system and liable for its harms. In the sub-sections that follow, we illustrate these challenges by focusing on three types of potential harms of algorithmic systems.

2.1 Representational Harms

Representational harms occur when algorithmic systems reinforce the political, economic, and/or cultural subordination or denigration of individuals based on their group identity [34], which often impacts access to spaces of public discourse and the ability to represent oneself and monetize such representation in these spaces. While there is little incentive for firms to publish details on the use of the tools of algorithmic moderation, there is emerging evidence of representational harms faced by individuals and groups whose content is subject to these tools [74]. An example here is the case of *Newman v. Google LLC & YouTube LLC* [139], where Black content creators associated with the Black Lives Matter movement sued YouTube for flagging their content as inappropriate and placing them under "restricted mode." This flagging blocked minors from viewing their videos and removed the possibility of monetization. The plaintiffs claimed that YouTube's flagging of their content constitutes harm on the grounds of racial discrimination, impeding their freedom of speech enshrined in the First Amendment, and false advertising under the Lanham Act. In adjudicating this case, the court ruled that it could not consider evidence for claimed harms of racial discrimination (due to lack of evidence showing intentional discrimination), impeding their freedom of speech (due to the plaintiff's inability to prove that YouTube's conduct constitutes state action), and false advertising (due to several reasons including that YouTube's content moderation policies does not constitute commercial advertising). This case shares a family resemblance with *Divino Group LLC v Google LLC* [138], where a group of LGBTQ+ content creators raised similar concerns around YouTube's content moderation policies. These cases highlight that there is a lack of (1) transparent documentation around content moderation practices, and (2) available expertise on how decisions on restricting content are made by algorithms (and human moderators), which makes it difficult for a plaintiff to claim representational harm in courts regardless of the merits of their case.

2.2 Explanatory Harms

Explanatory harms are experienced by those who contend with automated decisions that are critical for their life or life chances, such as in hiring, criminal legal system, and medicine, and yet remain opaque and resistant to interrogation. Currently, the US court system has no template for bringing a case against the deployers of an algorithmic decision support system, and to demand an explanation about how decisions were rendered about data subjects—people who live with and are "both resources and targets" of algorithmic systems [125:2]. There is, however, such a template in the EU's GDPR: the recognition of data subjects' "right to explanation" in relation to systems which make consequential decisions about them and their life chances. The right to explanation becomes a resource to claim, contest, and ameliorate explanatory harms. While the term "right to explanation" never appears in the GDPR, its requirements enact a statutory mandate obligating firms to provide "meaningful information" on the logic behind data-driven decision-making [102]. Selbst and Powles suggest that the GDPR trends towards "strengthening data protection as a fundamental right" [102:235], creating the conditions for accountability relationships between firms, the government, and the public. Workers in the ride hailing industry tested these new relational structures when they took two ride-hailing companies to court in the Netherlands to demand that data about their earnings, work assignments, and suspensions be disclosed [100]. They won their case, with the result that Uber and its competitor Ola must make both worker data and their logic around its use transparent. Requiring firms to provide data subjects insight into and details of the process of operationalizing data-driven decision-making obligates them to create and maintain adequate documentation. Such documentation, in turn, could help the courts not only adjudicate explanatory harms, but also other types of algorithmic harms.

2.3 Allocative Harms

Allocative harms occur in the context of distributing some resource or opportunity, for example, credit or jobs [45], when algorithmic systems allocate resources to some social groups more or less favorably than others [13]. Even when allocational differences fall across protected categories like race, gender, or disability, there are instances of significant challenges that prevent redress for harms through courts [103], despite regulations against discrimination [1]. These challenges include disparate awareness such that (potential) plaintiffs may not even be aware an algorithmic system was used in making an allocation decision [92] and inscrutability of the process of making the decision [55]. It is difficult to seek redress for allocational harms unless they have a discriminatory dimension, which in turn must pass stringent tests of proving that discrimination has occurred and was intentional [106]. These challenges are often showcased in external third-party audits that showcase how hiring algorithms can be arbitrary and unjust. Along these lines, a group of German journalists audited an algorithmic hiring system which used computer vision to assess videos of job interview candidates, conducting A/B tests to see how they were "scored" for attributes like conscientiousness, and agreeableness [17]. Disconcertingly, arbitrary changes to the video, such as addition of a bookcase in the background or a change in its overall tint, led to significant changes in candidate scores.

Such audits, while critical to call attention to these systems' issues, cannot by themselves foster accountability between parties. An individual who was algorithmically excluded from the opportunity to interview for a job would not necessarily know this choice was made by an algorithmic system, let alone the criteria used. A plaintiff in such a case would need to demonstrate that they had been intentionally deprived of an opportunity based on a flawed algorithmic product. Demonstrating that the product was flawed would require expertise and careful inspection of the algorithmic model, the data on which it was trained, the accuracy metrics for its predictions, and the basis on which it makes inferences about the suitability of candidates. Often, the basis for inferences is opaque even to the developers [24, 85]; thus, both the claims that a harm has occurred because of an algorithmic inference and that it was intentional become difficult to make in courts. New York City's recent Local Law 144 [30, 36, 79], which governs the use of automated employment decision tools (AEDTs), addresses this particular gap in part by requiring all employers using an AEDT to notify jobseekers and publicly post independent bias audits for racial and gender fairness metrics. This law operationalizes relations of accountability between vendors, employers, auditors, and jobseekers, as opposed to a centralized approach wherein the regulator (as a forum) strictly limits how AEDTs may operate or how jobseekers' data may be used. Local Law 144 provides access to courts when employers fail to live up to their algorithmic transparency obligations. This stands in contrast to a 2022 settlement and consent decree between Meta and the US Department of Justice over Facebook's targeted advertising practices [39, 57], which required Meta to develop a new system for delivering housing ads to address discrimination disparities. In this case, Facebook users who may have been subject to algorithmic discrimination were also unaware of having been excluded from an opportunity, but the legal action does not foreground the accountability relationship between users, forums, and developers. While Meta paid a fine and agreed to court supervision of changes to platform governance, the result did little to improve the position of the public. In contrast to Local Law 144, the forum effectively acts as a proxy for the public's interest but does not secure additional transparency or private action rights for harmed users.

In illustrating this brief typology of harms, we have also highlighted dynamics of accountability relationships that limit opportunities to understand, contest, and/or adjudicate harm. Here, we treat accountability as a matter of how parties are brought into a relationship with and made responsible to each other in a governance regime. Under the current regime, most types of algorithmic harms claimed by plaintiffs have largely been deemed inadmissible in courts because they cannot meet standards of concretely demonstrating that a harm has occurred or that it was intentional. The current situation constraints accountability by limiting who can be held answerable for such harms, who can seek redress for injuries, and how changes to risky or harmful systems can be mandated or incentivized. The courts, thus, offer a crucial empirical site to situate and examine the relational nature of algorithmic accountability in practice and build an effective governance regime for algorithmic systems. We draw on existing literature on the relational nature of accountability in the next section to explore the role that courts can potentially play in adjudicating algorithmic harms in particular and building governance regimes for algorithmic systems in general.

3 ACCOUNTABILITY AS A RELATION

Discourse on "algorithmic accountability" often suffers from a sort of grammatical illusion. The terms "algorithmic" and "accountability" seem to alternately modify each other, allowing the discourse to slip into discussions of algorithms that are themselves accountable, that is, algorithms that hold the property of accountability. While "accountable algorithms" have been technically specified in a narrow sense [63, 70], this slippage implies that merely exposing an account of how an algorithm works satisfies a design parameter for "accountability." However, the underlying rationale of "algorithmic accountability" must be that "accountability" modifies "algorithms." That is, developers and operators should be responsive to the people who use or are otherwise affected by their systems. In the absence of effective methods by which people might demand changes to that system, "accountability" is not meaningful. Accountability resides in the *relations* between the developers, regulators, and public and its collectivities, not in the algorithmic system nor in the developer's practices alone. As we show in this section, *these accountability relations are grounded in practices of documentation and securing due process in algorithmic governance regimes.*

Exploring the relational nature of accountability must begin with engaging with the grammatical illusion that tends to persist around accountability. Prior work in this space has spanned across "two concepts of accountability" [15:948], which respectively define it either as: (1) a virtuous property of individuals [80] or systems [38, 70], or (2) a relationship between differently-positioned parties [75, 101, 122]. In short, accountability oscillates between "a virtue" and "a mechanism". In the former, accountability inheres in the personal virtue of being accountable, and studies of accountability focus on "normative issues, on standards for, and the assessment of, the actual and active behavior of public agents" [15:947]. Accountability as a virtue is more strongly associated with scholarly discourses in the United States, whereas accountability as a mechanism is more likely to frame discourses in Canadian, EU, UK, and Australian contexts [6]. In these contexts, accountability pertains to relationships between actors and forums, and is studied by focusing on whether actors "are or can be held accountable ex post facto by accountability forums" [15:948].

Accountability is an essentially contested concept in terms of its application [15, 52, 69]. While a broader review of the concept is outside the scope of this paper, the concept can be traced back to the Domesday Book of 1066 [54]. It has accumulated meanings across the domains of its application, although these meanings and the practices associated with them have also come to shape these domains [89, 108]. The flexibility and contingency of the concept suggests that it is crucial to the practices of governing algorithmic systems, even if current approaches are not yet sufficient to that task. Taking this contingency as a starting point, a relational approach to algorithmic accountability implies that we—those who want to understand how an algorithmic system affects people's lives and who is responsible when harm is done—should focus on the relations that make possible and sustain accountability for algorithmic systems. Contesting the assumption that accountability inheres in technical features of such systems, or in mere documentation to satisfy compliance requirements, *a relational approach structures accountability around conditions of possibility for publics to emerge, cohere, and assert their shared interests to regulators and developers who are obligated to listen.*

Boven articulates one possible organizational structure for such accountability as "a relationship between an actor and a forum, in which the actor has an obligation to explain and to justify his or her conduct, the forum can pose questions and pass judgment, and the actor may face consequences" [14:447]. This articulation is useful in the context of algorithmic systems [122], however, questions remain about: (1) how the forum itself might be held accountable for its power over actors, (2) what that power consists of, with respect to the operation of algorithmic systems, and (3) how the public can make demands upon the forum to justify the adequacy of its assessment process and robustness of its assessments. We contend that *a dyadic relationship between regulator (forum) and developer (actor) is inadequate for the purposes of algorithmic accountability because it cannot foreground the interests of impacted communities and individuals.* Such a dyad inevitably results in forms of legal endogeneity [41]; it has "strong potential to be undermined by the incentives and institutional logics of the private sector" [101:1].

Extending the forum-actor dyad, we suggest a third entity, the public and its collectivities implicated by algorithmic systems, ought to be structurally positioned to hold the forum accountable [41]. When actors are themselves tasked with implementing oversight practices demanded by a forum (a common regulatory structure), regulatory goals inevitably tend to shift toward the actor's priorities [101]. In our proposed triadic model, the public serves as a counter-weighting third entity. While this weight is sometimes brought to bear through so-called 'court of public opinion' as represented in online and journalistic debates, or (less commonly) in community fora, for the purposes of this paper we discuss the public interest as most readily pursued through courts and the establishment of legal precedents requiring an accounting of public interest. Such accounting requires more opportunities to litigate the lived consequences of algorithmic systems.

Chief among these opportunities are practices of *documentation* and *recourse*. The ability of data subjects to secure due process of recourse when faced with algorithmic harms is firmly grounded in their access to documentation on the workings of an algorithmic system and its consequences. Such documentation establishes a route for data subjects and publics to contest the lived impacts of algorithmic systems through litigation. Recourse and transparent documentation are closely linked in jurisprudence, and contestation over outcomes is necessary for mediating accountability [64]. By providing data subjects access to an avenue to contest the harms they suffer, the courts can become: (1) a forum to provide redress for subjects' injuries; and (2) a backstop that ensures the forum is in turn also accountable for the scope of harms it holds the actors (developers) accountable for. Indeed, one of the more potent consequences of governance regimes like algorithmic impact assessments [4, 75, 78, 95], requiring developers to study and report on the consequences of their systems, is to create points at which the public can contest the adequacy and accuracy of the developer's claims. However, the persistent lack of these opportunities has increasingly become the core challenge of taking algorithms to court [97]. While we address lack of access to documentation as a standalone challenge, we have divided lack of recourse into two separate challenges of claiming standing and performing expertise in the next section.

4 THREE CORE CHALLENGES OF TAKING ALGORITHMS TO COURT

In regulatory terms, accountability exists between parties that can exert power and make demands upon each other; such demands include making an appeal, asking for permission, requiring redress when faced with harm, staking a claim to be heard, etc. These relations of accountability cannot exist when this capacity to exert power is skewed towards some parties at the expense of others. In the challenges we outline below, we explore how these power relations are often stacked against the public in a triadic actors-forum-public relationship. Collectively, they make a case for reimagining relations and arrangements of social and political power that can enable impacted people and communities to demand changes and have a voice in how algorithmic systems impact their lives.

4.1 Lack of Documentation and Incentivized Ignorance

Algorithmic systems are predominantly developed by private companies and are often hidden behind intellectual property and trade secrets protections [27]. These protections contribute to their opacity [24, 120], and also present a barrier for external, third-party actors [93] who wish to pursue accountability by calling attention to their workings. Furthermore, although internal documentation practices have grown recently [53, 76, 94], developers' concerns around potential harms may not align with the harms that concern the public interest [75, 78]. The result is that accountability for potential harm depends on what developers reveal about their systems, or what can be gleaned from their outputs.

Thus, while documentation practices are a key component of enacting accountability, they remain vulnerable: the developer necessarily must do most or all the documenting, enabling them to choose which features or consequences of the system to document. Given the tight relationship between transparency and jurisprudence [64], this raises two problems: (1) absence or selective disclosure of documentation that leaves the public with little recourse; and (2) limited availability of technical documentation on features that may only be of interest to developers and do not provide enough information to understand the relationship between algorithmic harms and system design. Under the current governance regime, documentation and transparency practices are more closely aligned with developers' interests in demonstrating compliance and facilitating business relationships than with the public interest in making demands about when and how systems ought to be deployed.

There are also perverse incentives for companies to remain ignorant of, or to obfuscate, the potential harms their systems may produce, as they can be held liable for failing to act on known or foreseeable dangers their products pose to the public. This same dynamic was central to the difficulty plaintiffs had in demonstrating the harms of tobacco products for smokers: tobacco companies knew about the health risks of tobacco, concealed that they knew, and commissioned skewed scientific research that enabled them to argue otherwise to influence public opinion and reduce plaintiffs' access to legal recourse [2]. Similarly, the whistleblower-facilitated release of the "Facebook Papers" confirms long-held suspicions that Facebook has conducted internal research, which indicates that the affordances of Instagram's design may produce emotional harm for younger users and concealed that they knew about these harms [65]. While commentators have addressed these revelations as a "Big Tobacco moment" for Facebook [124], researchers have also raised concerns around the chilling effect this may have on researching the lived consequences of social media [10, 87, 99].

Technology companies are often "leery of investigating the ramifications of their profit-seeking strategies" [112]. This form of agnotology [16, 90] can act as a shield, and has long played a role in attempts of corporate actors to sidestep responsibility for their actions by claiming ignorance of the consequences [44, 56]. However, there are cases when actors remain answerable for harms regardless of any claims of ignorance [80], particularly under strict liability torts [115]. Claims of ignorance as a defense only persist when societal norms allow them to do so.

Regulatory demands for documentation mediate between publics that benefit from access to information and recalcitrant corporate actors who do not independently volunteer such information. Such regulatory obligations may not necessarily be followed perfectly and can produce problematic trade-offs between truth and legitimacy [68], but they change how ignorance and knowledge are incentivized within particular governance regimes and shift organizational culture toward practices that comport with accountability [101]. One of the principal aims of impact assessment is "to get the people who build systems to think methodically about the details and potential impacts of a complex project before its implementation, and therefore head off risks before they become too costly to correct" [101:6]. Crucially, regulation engenders capacity to produce *meaningful* changes to the internal practices of organization, beyond the *mere* performance of regulatory rituals that do not necessarily satisfy its substantive goals.

4.2 Struggles in Claiming Standing and Incognizable Harms

"Standing" is a key concept for access to courts; it means to "stand before the court." For a plaintiff to have "standing," that is, for a court to hear their case, plaintiffs seeking redress must demonstrate a reasonable connection between another party's action or inaction and material harm to their interests. In the United States, standing in a federal court (where most computational harms are likely to be adjudicated[1]) is determined by interpretations of Article III of the Constitution, which stipulates that the ability to approach federal courts depends on establishing: 1) injury-in-fact: the plaintiff is a harmed party, with a claim to a cognizable harm; 2) causality: a link between the defendant's action or omission and the harm; 3) redressability: the claimed harm can be redressed by some action that the court is empowered to order. Harm is broadly understood as a wrongful impairment or setback of a person's, entity's, or society's interests, where an interest is any outcome in which one has a stake [49, 86]. Philosophically, there are many interests which might be wrongfully impaired—essentially any valued aspect of life, and/or life chances to achieve a desired state. However, only a relatively narrow range of interests constitute legally actionable harm.

Despite this limited range, there have been judicial innovations in questions of standing. For example, contending with forces of industrialization, questions of standing have become broader over time as the courts displaced "the individual victim [. . .] by a new kind of claimant — the class member, or the statistical victim [— in adjudicating . . .] the potential for injury to populations of indeterminate size and composition" [59:38]. Rule 23 of the Federal Rules of Civil Procedure has played a crucial role in formalizing 'class action' lawsuits since the mid-1960s and has facilitated the emergence of the 'statistical victim,' represented as a class with common features and interests [48]. These rules have allowed the courts to tackle "aggregated and probabilistic harm that seems always to lurk in modern corporate and industrial systems, no matter what steps have been taken for their control" [59:39]. While these measures allowed the courts to handle harm at the scale of formalized victim populations, they also transformed them into administrative agencies that defined the standards for eligibility and recovery for claimants [59]. While provisions of engaging with plaintiffs as statistical victims is the way forward, how to aggregate probabilistic harm remains an open question [114].

State actors are generally immune from lawsuits filed by US citizens, which makes it particularly hard to litigate their performance as a forum. Particularly, suing the federal government and its agencies for administrative procedural violations prior to material harm being done—such as conducting an inadequate impact assessment or not providing public hearings and input for major policy changes—is constrained by the Administrative Procedure Act of 1946 [127]. Currently, standing in federal courts for claims of injury arising out of a government agency's actions or its efforts to regulate another party or lack thereof is largely governed by the standard set in *Lujan v. Defenders of Wildlife* [140]. The Supreme Court ruled in this case that standing to claim such an injury requires demonstrating "an invasion of a legally protected interest" that meets two additional criteria: (1) it is "concrete and particularized"; and (2) it is "actual or imminent," not "conjectural or hypothetical" [140]. *Lujan* is especially important to question of standing as it relates to computational harms because it specifies who can challenge the actions of state actors not only as a forum (when they regulate or fail to regulate developers), but also as actors (when they develop algorithmic systems). A consequence of *Lujan* was to narrow the criteria under which plaintiffs could seek redress, requiring them to demonstrate that a state actor's (in)actions harmed them (or will harm them) in a manner prohibited by statute and specific to their own interests.

In addition to *Lujan*, standing for computational harms is also controlled by the rulings in *Clapper v. Amnesty International* in 2013 [129], a suit against the federal government claiming surveillance harms, and *Spokeo, Inc. v. Robbins* in 2016 [130], a class-action suit against a personal-records aggregator that published inaccurate information about individuals. The results of these cases established a standard that "fear of harm" and "bare procedural violations" do not rise to the level of "injury-in-fact" necessary to be granted standing. Navigating these criteria to claim standing has proven challenging for plaintiffs. Errors, breaches, and even intentional abuse may not result in cognizable harms that the courts see as "actual or imminent" and "concrete and particularized" [91].

Along similar lines, privacy law experts have pointed to *Lujan* [140], *Clapper* [129], *Spokeo* [130], and a host of other cases as a thicket of complicated and sometimes contradictory legal precedents, standards, and taxonomies that limit plaintiffs' capacity to reach standing or redress in privacy cases. For example, Citron and Solove [105] argue that the requirement for harms to be "visceral and vested" (their terms coined to encompass the range of precedents) to be cognizable establishes a burden that many privacy

[1]While the bulk of computational harms are likely to be adjudicated in federal courts, numerous computational harms have been adjudicated in states' courts, where laws like the California Consumer Privacy Act of 2020 [131] and the Illinois Biometric Information Privacy Act [71] provide standing for specific computational harms that take place within their jurisdictions.

harms cannot meet. Privacy harms—which share a family resemblance with algorithmic harms—can be diffuse, speculative, and downstream, yet standing requires material events that may not yet have happened (e.g., identity thieves combining multiple stolen databases) and injudiciously limits plaintiffs' options for recourse. Instead, Citron and Solove [105] advocate for a theory of computational privacy harms that could frame the anxiety caused by privacy breaches as a material harm, and so capture downstream risk in a manner similar to environmental harms. This remains challenging because it would imply that legislators will need to statutorily recognize new categories of rights, particularly around information technologies. As Romberg [98] has argued, court precedents on standing make such statutory recognition difficult to implement, despite the promise that plaintiff's claims of violation of such rights may serve as an effective resource to establish standing.

The issue of standing, however, is not simply a matter of reframing the process of approaching the court, it is also a matter of whether any claim of injury can withstand the scrutiny of adversarial proceedings. Waldman, for example, has observed that since privacy has often been treated as a contextual phenomenon, especially in scholarship [8, 81], it remains "open to attack as ambiguous" [116:696] in courts. Given their family resemblance, algorithmic harms are also approached as deeply contextual, which also makes them open to such attacks—the very scale and stochasticity of algorithmic harms that affect many different people makes harms appear causally vague for individuals. Thus, adversarial proceedings bring us to the second part of the challenge when it comes to recourse: cross-examinations during court proceedings.

4.3 Performing Expertise and Consensus in Court Proceedings

Adversarial proceedings in contests over harms caused by developments in technoscientific fields, including computational sciences, rely heavily on expert testimony. Performing expertise and providing testimony to establish ground truth and facticity around the matter of dispute is a key component of cross-examination. It provides the groundwork of defining standards and techniques to measure the contested harms as claimed by the plaintiffs. For example, in challenges to environmental impact assessment (EIA) documents produced by developers and approved by the Environmental Protection Agency (EPA), plaintiffs can take EPA to court by challenging: (1) the scope of the assessment, by arguing that it ignored significant impacts that affect them [134, 135]; or (2) the adequacy of the assessment, by arguing that the methods used to evaluate an impact inadequately assess the severity of a potential harm [133, 137]; or (3) the decision to permit the project, by arguing that the permitting agency ought to have recommended an alternative, less harmful, alternative design be chosen instead [132, 135–137]. Making any of these arguments requires expertise on environmental concerns that the plaintiffs must have access to and bring to bear upon their claims. The kind of challenges raised by plaintiffs also showcase a peculiar aspect of environmental litigations that they are often decided not by ascertaining ground truth of causality and consequences, but rather through cross-examination over whether "scientific assessment procedures were properly followed" [60].

During cross-examinations, credibility of the witness is crucial for evaluating claims. Even before their testimony is presented, the judge must decide in advance whether an expert witness has the right credentials to testify. Judges inevitably have come to play the role of gatekeepers [58, 60] in making such judgements. Although such gatekeeping has had its own exclusionary effects and has tended to favor corporate defendants [42, 43], it remains crucial in the process of establishing credibility of expert claims. The Daubert standard (instituted by the U.S. Supreme Court in *Daubert v. Merell Dow Pharmaceuticals* in 1993 [128]) provides the foundational heuristics for evaluating expert claims according to a set of criteria: "(1) whether the theory or technique underlying the evidence has been tested and is falsifiable; (2) whether it has been peer reviewed; (3) the technique's error rate, if known; and (4) general acceptance" [58:63] of the technique in the technoscientific field to which it belongs. The standards of peer review and general acceptance highlight that the courts have paid attention to the social process of building consensus over standards, and technoscientific facts.

Debates over validity of technical as well as qualitative methods used to measure and document the efficacy and consequences of algorithmic systems are currently ongoing [78:45-46, for example, maps the debates over suitability of different methods to measure the impact of the Alleghany Family Screening Tool]. Standards of peer review and general acceptance, amid such contestations, can become a significant barrier to admissibility of any kind of expert evidence contesting or even relying on algorithmic systems. The problem is further exacerbated by the lack of available documentation on the workings and consequences of algorithmic systems. For example, in *Newman v. Google LLC & YouTube LLC* [139] discussed to illustrate representational harms above, the plaintiffs argued that the inappropriate flagging of their content by YouTube (which is owned by Google's parent company Alphabet) constitutes harm on the grounds of racial discrimination. Among their evidence was the argument that, "in December of 2020, Google fired Timnit Gebru [. . .], the co-leader of Google's Ethical A.I. team, because Gebru complained about [Google's] 'biased filtering and blocking tools"' [139]. In bringing this event to the Court's attention, the plaintiffs were relying on Gebru's expertise in evaluating bias and discrimination in NLP algorithms used at Google as evidence of racial bias in YouTube's content moderation algorithms. In rejecting this evidence, the court reasoned that it "has no way of knowing if the filtering and blocking tools in question were used only at Google, or if they were also used at YouTube" [139]. The court reasoned that while algorithms used by Google and YouTube may be owned by the same parent company and share a family resemblance, they cannot be demonstrated to utilize the same code or moderation rules. Under current regulatory conditions, YouTube has few incentives and is under no obligation to provide documentation on its content moderation algorithms or assess the consequences of their moderation policies and algorithmic sorting processes in a manner that public(s) can contest. Regardless of the merits of plaintiffs' claims, this case shows how lack of available documentation and recognized expertise makes it harder for public(s) to cohere around potential harms and demand redress.

5 DISCUSSION: COGNIZABILITY OF ALGORITHMIC HARMS

In this paper, we have emphasized the role that access to courts can play in establishing the lower thresholds of due diligence for developers, in large part because courts are simultaneously the primary route *and* the primary roadblock for individuals and groups seeking redress for algorithmic harms. We outlined a set of three core challenges relating to lack of documentation, difficulties in claiming standing, and struggles around admissibility of expert evidence on and achieving consensus over the workings of algorithmic systems in adversarial proceedings. Resolving these challenges, however, is only one (necessary) part of the puzzle and is secondary to the primary goal of providing greater opportunities for publics to cohere around shared interests and make demands for changes to algorithmic systems. In looking for these opportunities, we build on a *relational* approach to algorithmic accountability that emphasizes not what the actors do nor the results of their actions, but *rather how interlocking relationships of accountability are constituted in a triadic relationship between actors, forums, and public(s).* The consequence of this account is that any successful algorithmic accountability regulations will likely not rely on rigidly specified controls of technical systems, but will rather shape accountability relationships such that publics can cohere around shared interests and harms, and thereby make actionable demands upon actors and fora.

5.1 Multi-Relational Algorithmic Accountability

One of the central challenges of building robust algorithmic accountability regimes is governing the middle ground that algorithmic systems occupy between individual and collective frameworks. This creates confusion around where to look in mapping consequences of these systems. Machine learning models "learn" about populations to render predictive decisions on individuals. The techno-economic foundations of machine learning are best understood as fundamentally relational: features about one person shape the fates of similarly situated persons, for better and for worse [114]. These relationalities produce both utility and risks of machine learning. *The origin of algorithmic harms is necessarily located in the collectivities captured in large datasets.* However, in an adversarial court system, each instance of harm is often most readily identifiable for individuals, as plaintiffs afforded rights within existing governance regimes.

For example, consider a facial recognition system used in policing. If it is developed with racially-biased data, it could lead to higher rates of false arrest for minority demographic groups. So, the origin of harm is in the historical relations underpinning the available facial data to train the system. But the only available remedy is often through demonstrating a particularized harm: an individual's right to recompense for a false or abusive arrest. There is no legal pathway for a remedy that could demand a wholesale rejection of the biased training data and/or the historical relations in which it is embedded.[2] This causes a short circuit, so to speak, where the only available methods for adjudicating and addressing harms to groups (and to individuals by virtue of their group membership) are centered on the rights of individuals (regardless of their group membership). An effective algorithmic accountability regime needs to bridge this gap.

Given the strict focus of U.S. laws on individual harms, this intertwining of individual and collective outcomes produces a "sociality problem" [114:8] in terms of how collective harms can be scaled down to an individual or conversely, individual harms can be scaled up to represent a problem for the public and its collectivities to engage with. This problem of scaling up and down illustrates how individual instances of algorithmic harm can only be identified and articulated out of an emerging taxonomy of collective harms. Therefore, accounts of individual harm need to include how this contingent and evolving taxonomy is practically developed and legitimized. The standards, methods, and techniques to evaluate the impact of algorithmic systems and produce such a taxonomy are still under development. Their legitimacy in identifying and evaluating harms, however, is dependent on the nature of relations they embed.

Current proposals for algorithmic accountability have included model cards, datasheets, and a range of first-, second-, and third-party audits [53, 76, 94]. These proposals depend, to varying degrees, on the prerogatives of those using an accountability tool or conducting an audit; a developer may audit their own product to verify compliance with some narrow requirement or an adversarial third-party may interrogate an algorithm to demonstrate it fails to meet a specific set of criteria the auditors are interested in. There are different relations at play in each form of audit. In a first-party audit, developers are mostly accountable to themselves. In a second-party audit, an auditor is brought into a relationship with the developer. Depending on how this relationship is organized, the auditor can potentially provide additional mechanisms for developers to internally hold themselves accountable. It is only in the context of an external third-party audit that the public(s) come into the field of relations, but only as an audience to what often becomes a public debate between the developer and the auditor over legitimacy of the audit results. While such audits are critically important, they do not by themselves enact a legal obligation with the necessary impetus that public(s) can employ to compel developers to change their systems [NYC's Local Law 144 provides an example of such an obligation 30,36]. Even the consequences of headline-making audits [22, 23] have ultimately depended on firms taking voluntary action, motivated by either goodwill or an interest in maintaining reputation [25] and public trust that had been threatened by the results of an audit [40].

In making a case for multi-relational algorithmic accountability, we argue that every standard to evaluate the impact of algorithmic systems and their harms produces a field of relations between involved parties [11 provides an in-depth analysis of relational ethics in machine learning]. The stronger these relations are and the more balanced the distribution of power between parties in implementing such a standard, the more robust the algorithmic accountability regime that the standard would produce. It is in

[2]For example, in 2021 the victim of false arrest by the Detroit Police Department, Robert Williams, with the support of the American Civil Liberties Union, filed suit demanding recompense and policy changes about the use of facial recognition by Detroit law enforcement [3, 109].[3, 109]. His ability to file this suit was dependent on him being individually injured by it, not on his relational status to a history of biases in the production and use of training data.

this context that we can see how key legal cases such as *Lujan* [140] act as a standard that shapes *who* can approach the court for algorithmic harms today. Future legal cases around emergent concerns of algorithmic harms can potentially provide an empirical site for what a FAccT-informed brand of legal and policy analysis could look like as it engages with how the court's judgment may come to shape the shifting role of algorithmic systems in everyday life and practice. Given the importance of legal precedents, such judgments are not simply a resolution of the matter at hand, but also a resource by which fairer, more reliable, and more defensible judgments can be rendered over time. We believe that the goal of a FAccT-informed analysis of court cases would not be to critique the emergence of new legal doctrines (this task is better handled by scholars of law), but to explore how such doctrines and the relations of accountability they engender may shape, extend, and undermine the possibilities of FAccT design and practice.

In the present moment, adversarial audits have become, and will remain, a critical need because the public(s) has struggled to occupy its needful place and stake as an interested party in court proceedings. On one hand, the rise of adversarial audits may be a symptom of the vacuum of relations that sustain algorithmic accountability. Algorithmic impact assessment regimes, on the other hand, can create the conditions of possibility for the public(s) to be in relation with both industry and government.

5.2 The Governance Regime of Algorithmic Impact Assessments

As legislation and regulatory efforts increasingly consider requiring developers to conduct assessments of their systems, a relational approach to algorithmic accountability provides some guidance around how to foreground the public interest. As discussed above, there are always many possible publics, and it is only through a contingent process that these publics cohere together to speak for their interests [37]. Establishing the right to contest the outcomes of algorithmic systems will be integral to this process [64]. Even without regulatory mandates, there are examples of publics cohering around shared interests in localized algorithmic accountability, such as community-sourced labeling of police violence data in Chicago [33] or tenants of Atlantic Towers in Brooklyn, New York, organizing against facial recognition systems [82]. At times, these publics are temporary and tactically opportune, other times they are robust and sustained; such publics cannot be pre-figured and may cohere around concerns with emergent technology not yet envisioned.

The central task of governing algorithmic systems in public interest is determining what type of algorithmic transparency regimes would provide public(s) with the requisite knowledge to contest the harms they identify. Efforts to cohere publics around algorithmic systems have drawn on regulation and case law for environmental harms [101], public nuisance [7], and regulation of food, drugs, and cosmetics [111] as suitable precedents. This diversity of analogous governance regimes is held together by the common feature of grappling with uncertainty around establishing causes and consequences.

Along these lines, algorithmic impact assessments (AIAs) are increasingly invoked as an appropriate mechanism when there are diffuse, probabilistic, and potentially novel forms of harms and many intersecting interests implicated in a proposed development or a deployed system [101]. Inferring from other domains of impact assessments, AIAs have the potential to partially address the challenges of building capacity for algorithmic accountability [4, 4, 75, 78, 95]. *Where courts now refuse standing for algorithmic harms for not being adequately "concrete" and "particular", or for not demonstrating "injury-in-fact," an AIA regime may create procedural rights that can protect the substantive rights that courts have struggled to recognize.* Making algorithmic harms more cognizable to courts shifts the field of relations to be considered in evaluating impact of a system, fosters norms that developers attend more closely to its possible harms [87] and provides a backstop for impacted public(s) to seek due process and remedies.

One of the central lessons of impact assessment regimes—especially, environmental impact assessment (EIA)—is that in matters of foregrounding public interests, the *how* of structuring accountability relations is more important than the content of the reporting/documentation process. Under the National Environmental Protection Act of 1969 (NEPA), developers under certain regulatory conditions are required to conduct EIAs prior to receiving a permit to proceed with their undertaking. EIAs are expected to thoroughly document a wide range of environmental impacts from the project, including impacts to water quality, wildlife habitat, air quality, cultural resources, and soil quality. In challenging EIAs, plaintiffs contest projects, not by suing the private company responsible for them, but by suing the agency responsible for accepting the EIA. This is a crucial point, as a plaintiff's standing to bring a suit against a private company is much more limited by the requirement that they demonstrate more direct, material harm. An individual or organization does have standing, however, to bring a suit under NEPA against a government agency "if he or she is adversely affected by an agency action" [12], such as accepting an impact assessment, under the Administrative Procedure Act [127].

The relatively recent evolution of EIAs to include environmental justice concerns [88, 113] illustrates the importance of litigating impact assessments. A long history of harmful developments in already environmentally degraded locations populated by disadvantaged communities [20], a growing public sentiment about environmental racism issues [21], and lawsuits claiming that environmental racism violated the Administrative Procedure Act (APA) [127], led to an executive order mandating inclusion of environmental justice concerns in evaluating permit requests [121]. By leveraging procedural rights to challenge the completeness of impact assessments that lack an account of environmental justice, a "public" was able to establish a proxy substantive right to have their interests considered, feeding a cycle in which novel or newly-understood harms were studied and then integrated into EIA [110]. Developers appealed rulings of regulatory agencies, regulatory agencies tested the boundaries of their authority, community advocates sued regulatory agencies, and new public interest organizations formed to pursue strategic litigation—the accountability relations structured by NEPA enabled the public(s) to emerge, cohere and made demands even if they did not win every ruling.

While AIAs have been implemented only in limited ways (see the use of AIAs in Canada as a self-assessment tool for developers [72]), forthcoming regulatory proposals have been reported as calling for more robust forms of algorithmic impact assessment, although the precise names of the assessment process differ [61]. These more robust forms of AIA would require developers as actors to document the expected impacts of such systems, and submit that documentation itself, or a report summarizing the assessment, to a government agency tasked with acting as a forum. This agency would: (1) mandate significant public consultation with stakeholders who might be affected by the system; (2) require developers to address harmful impacts that could be ameliorated by changes to the design or deployment of the system; and (3) make aspects of AIA documentation publicly available. As proposed, such regulation satisfies a need for greater understanding of how algorithmic systems produce harmful impacts. Crucially, by locating responsibility for overseeing AIAs within a federal agency, regulatory approaches have the opportunity—if drafted appropriately—to make algorithmic harms more cognizable to the courts. This also creates conditions for the public, as intermediated by the courts, to be better positioned in holding not only the forum, but also the developers accountable. For example, recently proposed additions to the business and professions code, relating to artificial intelligence in the California Legislature, stipulates conditions under which a person is authorized to bring a civil suit against a developer or deployer of an automated decision tool when faced with algorithmic discrimination [9].

Before we conclude, it is important to note that we impute a self-organizing capacity to public(s), which may not manifest in response to diverse and diffused forms of algorithmic harm that disproportionately impact systemically disadvantaged and minoritized communities. Furthermore, we present regulations centered on impact assessments as establishing a set of relations under the assumption that they are only a piece of the larger jigsaw puzzle of algorithmic accountability. Contending with algorithmic harms requires concerted efforts on multiple fronts ranging from better technical design to sustained advocacy efforts for public(s) to cohere around and demand redress for algorithmic harms.

6 CONCLUSION

Transparency and documentation can be an appealing solution to a wide range of accountability challenges in the machine learning industry and are familiar governance methods for developers and regulators alike. However, they are not a panacea [5, 5, 73]. As we have argued, the structure of accountability relations is critical to foregrounding the public interest in regulations for algorithmic systems. However, many of the proposed regulations currently under discussion globally, and including some already deployed, tend to provide little opportunity for public input.

Regulatory structures that create a dyadic relationship between a regulatory agency and developers for reporting and accepting/approving algorithmic impact assessments are at high risk of the regulator becoming fully dependent upon the developer for defining and measuring impacts. It is inevitable that developers will be primarily responsible for assessing their own systems. The question at hand is whether they alone are permitted to choose the metrics by which their systems are assessed or if the public can exert pressure on the thoroughness and adequacy of the assessments. For example, in the proposed EU AI regulations [32], the public regulators define risk tiers, prohibit certain applications deemed contrary to human rights, and require conformance audits for higher risk systems prior to deployment. However, there are few clear opportunities for impacted publics to contest the terms and outcomes of those audits [35]. Similarly, in the Algorithmic Accountability Act of 2019 [29], and its ongoing revisions [61], rely heavily on the Federal Trade Commission to oversee AIA reporting, but underspecify how the public can access and contest such reporting.

Such regulations should prioritize establishing a triadic developer-regulator-public accountability relation that provides clear footholds for the public to cohere around shared interests and contest the impacts of algorithmic systems. We suggest the following interventions in supporting the role of public(s):

- *Build feedback loops between regulators and scientific standards bodies.* Standards bodies, such as NIST in the US or the EU Standards Hub, and IEEE and ISO, are an effective route for researchers and advocacy groups to contribute to the formation of measurement practices that developers will use. Regulators are rarely equipped to determine on their own what process developers should follow, and standards bodies are well placed to receive feedback from the public and developers alike to find viable metrics and methods.
- *Specify public rights to review impact assessment documentation.* At a minimum the public should have access to documentation about the purpose and limits of systems, what features it was trained on, criteria used to measure algorithmic fairness, and information necessary to judge the gap between developer's claims about what the system does and its real-world performance.
- *Provide meaningful access to researchers.* Regulators should facilitate researchers seeking to understand broadly how the AI/ML industry is meeting accountability obligations, including providing portals for access to impact assessments and aggregated trends.

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers, and members of the RAW session at Data & Society Research Institute and Siegel Fellows Research Seminar for their invaluable feedback and suggestions. Parts of the paper were presented in a workshop at RightsCon 2022 organized in collaboration with Borhane Blili-Hamelin, Gina Helfrich, Jillian Powers, and Vanja Skoric. Research for this paper was supported by the National Science Foundation Award #1704425, Digital Life Initiative at Cornell Tech, and Siegel Family Foundation.

REFERENCES

[1] 42 U.S.C. §2000e et seq. 1964. *Civil Rights Act of 1964*. Retrieved July 9, 2021 from https://www.govinfo.gov/content/pkg/STATUTE-78/pdf/STATUTE-78-Pg241.pdf#page$=$1
[2] Mohamed Abdalla and Moustafa Abdalla. 2021. The Grey Hoodie Project: Big Tobacco, Big Tech, and the Threat on Academic Integrity. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 287–297. https://doi.org/10.1145/3461702.3462563
[3] ACLU. 2021. Michigan Father Sues Detroit Police Department for Wrongful Arrest Based on Faulty Facial Recognition Technology. Retrieved January 2, 2022 from https://www.aclu.org/press-releases/michigan-father-sues-detroit-police-department-wrongful-arrest-based-faulty-facial
[4] Ada Lovelace Institute and Open Government Institute. 2021. *Algorithmic accountability for the public sector.* Retrieved from https://

www.opengovpartnership.org/documents/ algorithmic-accountability-public-sector/

[5] Mike Ananny and Kate Crawford. 2018. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society* 20, 3: 973–989. https://doi.org/10.1177/1461444816676645

[6] Peter Aucoin and Ralph Heintzman. 2000. The Dialectics of Accountability for Performance in Public Management Reform. *International Review of Administrative Sciences* 66, 1: 45–55. https://doi.org/10.1177/0020852300661005

[7] Jack M. Balkin. 2017. The Three Laws of Robotics in the Age of Big Data. *Faculty Scholarship Series* 5159. Retrieved from https://digitalcommons.law.yale.edu/fss_papers/5159

[8] Kenneth A. Bamberger and Deirdre K. Mulligan. 2012. PIA requirements and privacy decision-making in US government agencies. In *Privacy Impact Assessment*. Springer, Dordrecht, 225–250. Retrieved from https://link.springer.com/chapter/10.1007/978-94-007-2543-0_10

[9] Rebecca Bauer-Kahan. 2023. *California AB-331 Automated decision tools*. Retrieved May 3, 2023 from https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id$=$202320240AB331

[10] Susan Benesch. 2021. Nobody Can See Into Facebook. *The Atlantic*. Retrieved November 5, 2021 from https://www.theatlantic.com/ideas/archive/2021/10/facebook-oversight-data-independent-research/620557/

[11] Abeba Birhane. 2021. Algorithmic injustice: a relational ethics approach. *Patterns* 2, 2: 100205. https://doi.org/10.1016/j.patter.2021.100205

[12] Ilyssa Birnbach. 2017. Newly Imposed Limitations on Citizens' Right to Sue for Standing in a Procedural Rights Case. *Fordham Environmental Law Review* 9, 2: 311–336.

[13] Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. 2020. Language (Technology) is Power: A Critical Survey of "Bias" in NLP. *arXiv:2005.14050 [cs]*. Retrieved November 12, 2021 from http://arxiv.org/abs/2005.14050

[14] Mark Bovens. 2007. Analysing and Assessing Accountability: A Conceptual Framework. *European Law Journal* 13, 4: 447–468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

[15] Mark Bovens. 2010. Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism. *West European Politics* 33, 5: 946–967. https://doi.org/10.1080/01402382.2010.486119

[16] danah boyd. 2019. Agnotology and Epistemological Fragmentation. *Points / Medium*. Retrieved January 7, 2022 from https://points.datasociety.net/agnotology-and-epistemological-fragmentation-56aa3c509c6b

[17] BR24. 2021. Objective or Biased: On the questionable use of Artificial Intelligence for job applications. *BR*. Retrieved January 10, 2022 from https://interaktiv.br.de/ki-bewerbung/en/

[18] Miles Brundage. 2016. Artificial Intelligence and Responsible Innovation. In *Fundamental Issues of Artificial Intelligence*, Vincent C. Müller (ed.). Springer International Publishing, Oxford, UK, 543–554. https://doi.org/10.1007/978-3-319-26485-1

[19] Miles Brundage, Shahar Avin, Jasmine Wang, Haydn Belfield, Gretchen Krueger, Gillian Hadfield, Heidy Khlaaf, Jingying Yang, Helen Toner, Ruth Fong, Tegan Maharaj, Pang Wei Koh, Sara Hooker, Jade Leung, Andrew Trask, Emma Bluemke, Jonathan Lebensold, Cullen O'Keefe, Mark Koren, Théo Ryffel, J. B. Rubinovitz, Tamay Besiroglu, Federica Carugati, Jack Clark, Peter Eckersley, Sarah de Haas, Maritza Johnson, Ben Laurie, Alex Ingerman, Igor Krawczuk, Amanda Askell, Rosario Cammarota, Andrew Lohn, David Krueger, Charlotte Stix, Peter Henderson, Logan Graham, Carina Prunkl, Bianca Martin, Elizabeth Seger, Noa Zilberman, Seán Ó hÉigeartaigh, Frens Kroeger, Girish Sastry, Rebecca Kagan, Adrian Weller, Brian Tse, Elizabeth Barnes, Allan Dafoe, Paul Scharre, Ariel Herbert-Voss, Martijn Rasser, Shagun Sodhani, Carrick Flynn, Thomas Krendl Gilbert, Lisa Dyer, Saif Khan, Yoshua Bengio, and Markus Anderljung. 2020. Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims. *arXiv:2004.07213 [cs]*. Retrieved December 29, 2020 from http://arxiv.org/abs/2004.07213

[20] Robert D Bullard. 1999. Dismantling Environmental Racism in the USA. *Local Environment* 4, 1: 25.

[21] Robert D. Bullard. 1999. Anatomy of Environmental Racism and the Environmental Justice Movement. In *Confronting Environmental Racism: Voices From the Grassroots*, Robert D. Bullard (ed.). South End Press.

[22] Joy Buolamwini. 2019. Response: Racial and Gender bias in Amazon Rekognition — Commercial AI System for Analyzing Faces. *Medium*. Retrieved October 5, 2020 from https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced

[23] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of Machine Learning Research*, 1–15. Retrieved from http://proceedings.mlr.press/v81/buolamwini18a.html

[24] 24. Jenna Burrell. 2016. How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms. *Big Data & Society* 3, 1: 1–12. https://doi.org/10.1177/2053951715622512

[25] E. Madalina Busuioc and Martin Lodge. 2016. The Reputational Basis of Public Accountability. *Governance* 29, 2: 247–263. https://doi.org/10.1111/gove.12161

[26] François Candelon, Rodolphe Charme di Carlo, Midas De Bondt, and Theodoros Evgeniou. 2021. AI Regulation Is Coming. *Harvard Business Review*. Retrieved December 23, 2021 from https://hbr.org/2021/09/ai-regulation-is-coming

[27] Anupam Chander. 2013. How Law Made Silicon Valley. *Emory Law Journal* 63: 639–694.

[28] Angèle Christin, Alex Rosenblat, and danah boyd. 2015. Courts and Predictive Algorithms. Retrieved from https://datasociety.net/library/data-civil-rights-courts-and-predictive-algorithms/

[29] Yvette D. Clarke. 2019. *H.R.2231 - 116th Congress (2019-2020): Algorithmic Accountability Act of 2019*. Retrieved October 5, 2020 from https://www.congress.gov/bill/116th-congress/house-bill/2231

[30] Committee on Technology. 2021. File #: Int 1894-2020: Automated employment decision tools. *The New York City Council*. Retrieved January 30, 2023 from https://legistar.council.nyc.gov/LegislationDetail.aspx?ID$=$4344524&GUID$=$B051915D-A9AC-451E-81F8-6596032FA3F9&Options$=$Advanced&Search

[31] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. 2017. Algorithmic Decision Making and the Cost of Fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 797–806. https://doi.org/10.1145/3097983.3098095

[32] Council of Europe and European Parliament. 2021. Regulation on European Approach for Artificial Intelligence Laying Down a Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. Retrieved from https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence

[33] CPDP Team. Citizens Police Data Project. *Invisible Institute*. Retrieved January 25, 2023 from https://invisible.institute/police-data

[34] Kate Crawford. 2017. Keynote: The Trouble with Bias. Retrieved January 19, 2022 from https://www.youtube.com/watch?v$=$fMym_BKWQzk

[35] Data & Society Research Institute and European Center for Non-for-Profit Law. 2021. *Recommendation for Incorporating Human Rights Into AI Impact Assessments*. Retrieved January 18, 2022 from https://datasociety.net/wp-content/uploads/2021/11/HUDIERA-Full-Paper_FINAL.pdf

[36] DCWP. 2023. NYC Rules. *NYC Rules*. Retrieved January 30, 2023 from https://rules.cityofnewyork.us/rule/automated-employment-decision-tools-updated/

[37] John Dewey. 1927. *The Public and Its Problems*. H. Holt and Company, New York.

[38] Nicholas Diakopoulos. 2016. Accountability in Algorithmic Decision Making. *Communications of the ACM* 59, 2: 56–62. https://doi.org/10.1145/2844110

[39] DOJ. 2022. Justice Department Secures Groundbreaking Settlement Agreement with Meta Platforms, Formerly Known as Facebook, to Resolve Allegations of Discriminatory Advertising. *The United States Department of Justice: Press Releases*. Retrieved January 30, 2023 from https://www.justice.gov/opa/pr/justice-department-secures-groundbreaking-settlement-agreement-meta-platforms-formerly-known

[40] Dylan Doyle-Burke and Jessie Smith. IBM, Microsoft, and Amazon Disavow Facial Recognition Technology: What Do You Need to Know? with Deb Raji. Retrieved October 6, 2020 from https://radicalai.podbean.com/e/ibm-microsoft-and-amazon-disavow-facial-recognition-technology-what-do-you-need-to-know-with-deb-raji/

[41] Lauren B. Edelman and Shauhin A. Talesh. 2011. To Comply or Not to Comply – That Isn't the Question: How Organizations Construct the Meaning of Compliance. In *Explaining Compliance*. Edward Elgar Publishing. https://doi.org/10.4337/9780857938732.00011

[42] Gary Edmond. 2002. Legal Engineering: Contested Representations of Law, Science (and Non-science) and Society. *Social Studies of Science* 32, 3: 371–412. https://doi.org/10.1177/0306312702032003002

[43] Gary Edmond and David Mercer. 2004. Daubert and the Exclusionary Ethos: The Convergence of Corporate and Judicial Attitudes towards the Admissibility of Expert Evidence in Tort Litigation*. *Law & Policy* 26, 2: 231–257. https://doi.org/10.1111/j.0265-8240.2004.00011.x

[44] Kevin C. Elliott. 2015. Selective ignorance in environmental research. In *Routledge international handbook of ignorance studies*, Matthias Gross and Linsey McGoey (eds.). Routledge, Taylor & Francis Group, New York.

[45] Karin van Es, Daniel Everts, and Iris Muis. 2021. Gendered language and employment Web sites: How search algorithms can cause allocative harm. *First Monday*. https://doi.org/10.5210/fm.v26i8.11717

[46] European Commission. 2019. Ethics Guidelines for Trustworthy AI. Retrieved from https://ec.europa.eu/newsroom/dae/document.cfm?doc_id$=$60419

[47] European Commission. *Digital Services Act*. Retrieved May 3, 2023 from https://eur-lex.europa.eu/eli/reg/2022/2065/oj

[48] Federal Rules of Civil Procedure. Rule 23. Class Actions. *LII / Legal Information Institute*. Retrieved November 15, 2021 from https://www.law.cornell.edu/rules/frcp/rule_23

[49] Joel Feinberg. 1984. *Harm to Others*. Oxford University Press.

[50] Luciano Floridi. 2019. Establishing the rules for building trustworthy AI. *Nature Machine Intelligence* 1, 6: 261–262. https://doi.org/10.1038/s42256-019-0055-y

[51] Sheera Frenkel. 2021. Whistleblower discusses how Instagram may lead teenagers to eating disorders. *The New York Times*. Retrieved December 23, 2021 from https://www.nytimes.com/live/2021/10/05/technology/facebook-whistleblower-frances-haugen

[52] W. B. Gallie. 1955. Essentially Contested Concepts. *Proceedings of the Aristotelian Society* 56: 167–198.

[53] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2018. Datasheets for Datasets. In *Proceedings of the 5th Workshop on Fairness, Accountability, and Transparency in Machine Learning*. Retrieved July 6, 2020 from http://arxiv.org/abs/1803.09010

[54] Andy Godfrey and Keith Hooper. 1996. Accountability and decision-making in feudal England: Domesday Book revisited. *Accounting History* 1, 1: 35–54. https://doi.org/10.1177/103237329600100103

[55] James Grimmelmann and Daniel Westreich. 2016. Incomprehensible discrimination. *Calif. L. Rev. Circuit* 7: 164.

[56] Matthias Gross and Linsey McGoey (eds.). 2015. *Routledge international handbook of ignorance studies*. Routledge, Taylor & Francis Group, New York.

[57] Tracy Jan and Elizabeth Dwoskin. 2019. Facebook agrees to overhaul targeted advertising system for job, housing and loan ads after discrimination complaints. *Washington Post*. Retrieved January 30, 2023 from https://www.washingtonpost.com/business/economy/facebook-agrees-to-dismantle-targeted-advertising-system-for-job-housing-and-loan-ads-after-discrimination-complaints/2019/03/19/7dc9b5fa-4983-11e9-b79a-961983b7e0cd_story.html

[58] Sheila Jasanoff. 1995. *Science at the Bar: Law, Science, and Technology in America*. Harvard University Press, Cambridge, Massachusetts.

[59] Sheila Jasanoff. 2002. Science and the Statistical Victim: Modernizing Knowledge in Breast Implant Litigation. *Social Studies of Science* 32, 1: 37–69.

[60] Sheila Jasanoff. 2018. Science, Common Sense & Judicial Power in U.S. Courts. *Daedalus* 147, 4: 15–27. https://doi.org/10.1162/daed_a_00517

[61] Khari Johnson. 2021. The Movement to Hold AI Accountable Gains More Steam. *Wired*. Retrieved December 17, 2021 from https://www.wired.com/story/movement-hold-ai-accountable-gains-steam/

[62] Kristen Johnson, Frank Pasquale, and Jennifer Chapman. 2019. Artificial Intelligence, Machine Learning and Bias in Finance: Toward Responsible Innovation. *Fordham Law Review* 88, 2: 499–530.

[63] Severin Kacianka and Alexander Pretschner. 2021. Designing Accountable Systems. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 424–437. https://doi.org/10.1145/3442188.3445905

[64] Margot E Kaminski and Jennifer M Urban. 2021. The Right to Contest AI. *Columbia Law Review* 121, 7: 1957–2048.

[65] Cecilia Kang. 2021. Facebook Whistle-Blower Urges Lawmakers to Regulate the Company. *The New York Times*. Retrieved November 5, 2021 from https://www.nytimes.com/2021/10/05/technology/facebook-whistle-blower-hearing.html

[66] Kate Kaye. 2021. Facebook is "not a researchers-friendly space" say academics. *Digiday*. Retrieved December 23, 2021 from https://digiday.com/marketing/facebook-is-not-a-researchers-friendly-space-say-academics-encountering-roadblocks-to-analyzing-its-2020-election-ad-data/

[67] Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, and Ashesh Rambachan. 2018. Algorithmic Fairness. *AEA Papers and Proceedings* 108: 22–27. https://doi.org/10.1257/pandp.20181018

[68] Ida Koivisto. 2022. *The Transparency Paradox*. Oxford University Press, Oxford, United Kingdom.

[69] Joshua A. Kroll. 2020. Accountability in Computer Systems. In *The Oxford handbook of ethics of AI*, Markus Dirk Dubber, Frank Pasquale and Sunit Das (eds.). Oxford University Press, New York, NY, 181–196.

[70] Joshua A Kroll, Joanna Huey, Solon Barocas, Edward W Felten, Joel R Reidenberg, David G Robinson, and Harlan Yu. 2017. Accountable Algorithms. *University of Pennsylvania Law Review* 165: 633–705.

[71] Legislative Reference Bureau. 2013. 740 ILCS 14/ Biometric Information Privacy Act. *Illinois General Assembly*. Retrieved January 11, 2022 from https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID$=$3004&ChapterID$=$57

[72] Mathieu Lemay. 2021. Understanding Canada's Algorithmic Impact Assessment Tool. *Medium*. Retrieved December 17, 2021 from https://towardsdatascience.com/understanding-canadas-algorithmic-impact-assessment-tool-cd0d3c8cafab

[73] Karen EC Levy and David Merritt Johns. 2016. When open data is a Trojan Horse: The weaponization of transparency in science and governance. *Big Data & Society* 3, 1: 1–6. https://doi.org/10.1177/2053951715621568

[74] Mandy Lau. Content moderation as language policy:Connecting commercial content moderation policies, regulations, and language policy. *Working papers in Applied Linguistics and Linguistics at York*.

[75] Jacob Metcalf, Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, and Madeleine Clare Elish. 2021. Algorithmic Impact Assessments and Accountability: The Co-construction of Impacts. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (FAccT '21), 735–746. https://doi.org/10.1145/3442188.3445935

[76] 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*, 220–229. https://doi.org/10.1145/3287560.3287596

[77] 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: 141–163. https://doi.org/10.1146/annurev-statistics-042720-125902

[78] Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, Madeleine Clare Elish, and Jacob Metcalf. 2021. *Assembling Accountability: Algorithmic Impact Assessment for the Public Interest*. Data & Society Research Institute. Retrieved August 17, 2021 from https://datasociety.net/library/assembling-accountability-algorithmic-impact-assessment-for-the-public-interest/

[79] New York City Department of Consumer and Worker Protection. 2023. *Notice of Adoption for Use of Automated Employment Decisionmaking Tools*. Retrieved April 26, 2023 from https://rules.cityofnewyork.us/wp-content/uploads/2023/04/DCWP-NOA-for-Use-of-Automated-Employment-Decisionmaking-Tools-2.pdf

[80] Helen Nissenbaum. 1996. Accountability in a computerized society. *Science and Engineering Ethics* 2, 1: 25–42. https://doi.org/10.1007/BF02639315

[81] Helen Nissenbaum. 2004. Privacy as Contextual Integrity. *Washington Law Review* 79: 41.

[82] Mutale Nkonde. 2019. Automated Anti-Blackness: Facial Recognition in Brooklyn, New York. *Harvard Journal of African American Public Policy* 20: 30–36.

[83] Serena Oduro, Emanuel Moss, and Jacob Metcalf. 2022. Obligations to assess: Recent trends in AI accountability regulations. *Patterns* 3, 11: 100608. https://doi.org/10.1016/j.patter.2022.100608

[84] Francesca Palmiotto. 2021. The Black Box on Trial: The Impact of Algorithmic Opacity on Fair Trial Rights in Criminal Proceedings. In *Algorithmic Governance and Governance of Algorithms: Legal and Ethical Challenges*, Martin Ebers and Marta Cantero Gamito (eds.). Springer International Publishing, Cham, 49–70. https://doi.org/10.1007/978-3-030-50559-2_3

[85] Frank Pasquale. 2015. *The black box society: the secret algorithms that control money and information*. Harvard University Press, Cambridge.

[86] Stephen Perry. 2003. Harm, History, and Counterfactuals. *San Diego Law Review* 40, 4: 1283.

[87] Nathan Persily. 2021. Facebook hides data showing it harms users. Outside scholars need access. *Washington Post*. Retrieved November 5, 2021 from https://www.washingtonpost.com/outlook/2021/10/05/facebook-research-data-haugen-congress-regulation/

[88] Sara Pirk. 2002. Expanding Public Participation in Environmental Justice: Methods, Legislation, Litigation and Beyond. *Journal of Environmental Law and Litigation* 17, 1: 35.

[89] Michael Power. 1994. *The Audit Explosion*. Demos, London: Demos. Retrieved from https://www.demos.co.uk/files/theauditexplosion.pdf

[90] Robert Proctor and Londa L. Schiebinger (eds.). 2008. *Agnotology: the making and unmaking of ignorance*. Stanford University Press, Stanford, CA.

[91] Serena Quattrocolo. 2020. Artificial Intelligence, Computational Modelling and Criminal Proceedings: A Framework for A European Legal Discussion. Springer, Cham, Suisse.

[92] Manish Raghavan and Solon Barocas. 2019. *Challenges for mitigating bias in algorithmic hiring*. Brookings Institute, Washington, DC. Retrieved November 12, 2021 from https://www.brookings.edu/research/challenges-for-mitigating-bias-in-algorithmic-hiring/

[93] Inioluwa Deborah Raji and Joy Buolamwini. 2019. Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 429–435. https://doi.org/10.1145/3306618.3314244

[94] Inioluwa Deborah Raji, Andrew Smart, Rebecca N White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. 2020. Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing. In *Conference on Fairness, Accountability, and Transparency (FAT* '20)*, 12. Retrieved from https://doi.org/10.1145/3351095.3372873

[95] Dillon Reisman, Jason Schultz, Kate Crawford, and Meredith Whittaker. 2018. *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*. AI Now Institute, New York. Retrieved from https://ainowinstitute.org/aiareport2018.pdf

[96] Frank Rep. Pallone. 2022. *Text - H.R.8152 - 117th Congress (2021-2022): American Data Privacy and Protection Act*. Retrieved March 3, 2023 from http://www.congress.gov/

[97] Rashida Richardson, Jason M Schultz, and Vincent M Southerland. 2019. *Litigating Algorithms 2019 US Report: New Challenges to Government Use of Algorithmic Decision Systems*. AI Now Institute, New York. Retrieved from https://ainowinstitute.org/litigatingalgorithms-2019-us.pdf

[98] Jon Romberg. 2020. Trust the Process: Understanding Procedural Standing Under Spokeo. *Oklahoma Law Review* 72: 86.

[99] Zeve Sanderson, Jonathan Nagler, and Joshua A. Tucker. 2021. Academic Researchers Need Access to the Facebook Papers. *Slate*. Retrieved November 5, 2021 from https://slate.com/technology/2021/11/facebook-papers-frances-haugen-academic-research-transparency.html

[100] Paul Sawers. 2020. Uber drivers sue for data on secret profiling and automated decision-making. *VentureBeat*. Retrieved January 10, 2022 from https://venturebeat.com/2020/07/20/uber-drivers-sue-for-data-on-secret-profiling-and-automated-decision-making/

[101] Andrew D Selbst. 2021. An Institutional View of Algorithmic Impact Assessments. *Harv. J.L. & Tech.* 35: 78.

[102] Andrew D Selbst and Julia Powles. 2017. Meaningful information and the right to explanation. *International Data Privacy Law* 7, 4: 233–242. https://doi.org/10.1093/idpl/ipx022

[103] Elaine W. Shoben. 2004. Disparate Impact Theory in Employment Discrimination: What's Griggs Still Good For? What Not? *Brandeis Law Journal* 42, 3: 597–622.

[104] Tom Simonite. 2021. What Really Happened When Google Ousted Timnit Gebru. *Wired*. Retrieved January 7, 2022 from https://www.wired.com/story/google-timnit-gebru-ai-what-really-happened/

[105] Daniel J. Solove and Danielle Keats Citron. 2017. Risk and Anxiety: A Theory of Data-Breach Harms. *Texas Law Review* 96: 737.

[106] Sandra F. Sperino. 2011. Rethinking Discrimination Law. *Michigan Law Review* 110, 1: 69–125.

[107] Jack Stilgoe, Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42, 9: 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

[108] Marilyn Strathern. 2000. *Audit Cultures: Anthropological Studies in Accountability, Ethics, and the Academy*. Routledge, New York.

[109] Ryan Tate-Mosley. 2021. The new lawsuit that shows facial recognition is officially a civil rights issue. *MIT Technology Review*. Retrieved January 2, 2022 from https://www.technologyreview.com/2021/04/14/1022676/robert-williams-facial-recognition-lawsuit-aclu-detroit-police/

[110] Serge Taylor. 1984. *Making Bureaucracies Think: The Environmental Impact Statement Strategy of Administrative Reform*. Stanford University Press, Stanford, CA.

[111] Andrew Tutt. 2017. An FDA for Algorithms. *Administrative Law Review* 69, 1: 83–123. https://doi.org/10.2139/ssrn.2747994

[112] Heidi Tworek. 2019. Social Media Platforms and the Upside of Ignorance. *Centre for International Governance Innovation*. Retrieved November 5, 2021 from https://www.cigionline.org/articles/social-media-platforms-and-upside-ignorance/

[113] Uma Outka. 2006. NEPA and Environmental Justice: Integration, Implementation, and Judicial Review. *Boston College Environmental Affairs Law Review* 33, 3: 601–626.

[114] Salome Viljoen. 2020. *A Relational Theory of Data Governance*. Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.3727562

[115] John W. Wade. 1973. On the Nature of Strict Tort Liability for Products. *Mississippi Law Journal* 44, 5: 825–851.

[116] Ari Ezra Waldman. 2018. Designing without privacy. *Houston Law Review* 55, 659.

[117] Jeremy Waldron. 2011. The Rule of Law and the Importance of Procedure. *Nomos* 50: 3–31.

[118] Michael Warner. 2005. *Publics and counterpublics*. Zone Books, New York.

[119] Elizabeth Anne Watkins, Emanuel Moss, Jacob Metcalf, Ranjit Singh, and Madeleine Clare Elish. Elizabeth Anne Watkins, Emanuel Moss, Jacob Metcalf, Ranjit Singh, and Madeleine Clare Elish. 2021. Governing Algorithmic Systems with Impact Assessments: Six Observations. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (AIES '21), 1010–1022. https://doi.org/10.1145/3461702.3462580

[120] Rebecca Wexler. Rebecca Wexler. 2018. Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System. *Stan. L. Rev.* 70: 1343. https://doi.org/10.2139/ssrn.2920883

[121] White House. White House. 1994. Executive Order 12898 of February 11, 1994 | Federal Actions To Address Environmental Justice in Minority Populations and Low-Income Populations (59 FR 7629). *Federal Register* 59, 32. Retrieved from https://www.archives.gov/files/federal-register/executive-orders/pdf/12898.pdf

[122] Maranke Wieringa. Maranke Wieringa. 2020. What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 1–18. https://doi.org/10.1145/3351095.3372833

[123] Ron Wyden. Ron Wyden. *Algorithmic Accountability Act of 2022*. Retrieved February 6, 2023 from https://www.congress.gov/bill/117th-congress/house-bill/6580/text

[124] Cat Zakrzewski. Cat Zakrzewski. 2021. Facebook whistleblower's revelations could usher in tech's 'Big Tobacco moment,' lawmakers say. *Washington Post*. Retrieved November 5, 2021 from https://www.washingtonpost.com/technology/2021/10/06/facebook-whistleblower-frances-haugen-senate-unity/

[125] Malte Ziewitz and Ranjit Singh. 2021. Critical companionship: Some sensibilities for studying the lived experience of data subjects. *Big Data & Society* 8, 2: 1–13. https://doi.org/10.1177/20539517211061122

[126] Andrej Zwitter. 2014. Big Data Ethics. *Big Data & Society* 1, 2: 1–6. https://doi.org/10.1177/2053951714559253

[127] 1946. *Administrative Procedure Act*. Retrieved from https://www.law.cornell.edu/uscode/text/5/551

[128] 1993. *Daubert v. Merrell Dow Pharmaceuticals, Inc.* Supreme Court. Retrieved from https://scholar.google.com/scholar_case?case$=$827109112258472814

[129] 2013. *Clapper v. Amnesty Intern. USA*. Supreme Court. Retrieved from https://scholar.google.com/scholar_case?case$=$13987610465347808803

[130] 2016. *Spokeo, Inc. v. Robins*. Retrieved from https://scholar.google.com/scholar_case?case$=$11810453531811593153

[131] 2018. California Consumer Privacy Act. *California Legislative Information*. Retrieved from https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id$=$201720180AB375

[132] 2018. *Save Our Sound Obx, Inc. v. Nc Dept. of Transp*. Retrieved from https://scholar.google.com/scholar_case?case$=$16317094648654068553

[133] 2019. *Oregon Natural Desert Association v. Rose*. Retrieved from https://scholar.google.com/scholar_case?case$=$4325132358104036335

[134] 2019. *Diné Citizens Against Ruining Env't v. Bernhardt*. Retrieved from https://scholar.google.com/scholar_case?case$=$9745116248581207846

[135] 2019. *Conservation Congress v. United States Forest Service*. Retrieved from https://scholar.google.com/scholar_case?case$=$14603573195018924737

[136] 2019. *Protect Our Communities Foundation v. LaCounte*. Retrieved from https://scholar.google.com/scholar_case?case$=$17922905489759569030

[137] 2019. *National Parks Conservation Ass'n v. Semonite*. Retrieved from https://scholar.google.com/scholar_case?case$=$15694333209763019110

[138] 2021. *Divino Group LLC v. Google LLC*. Retrieved from https://scholar.google.com/scholar_case?case$=$9914604969716406988

[139] 2021. *Newman v. Google LLC*. Retrieved from https://scholar.google.com/scholar_case?case$=$5968246280551343353

[140] 1992. Lujan v. Defenders of Wildlife, 504 U.S. 555. *Justia Law*. Retrieved from https://supreme.justia.com/cases/federal/us/504/555/