

DISPARATE (ALGORITHMIC) ADVANTAGE

Yunsieg P. Kim^{*} & *Jacob Eisler*[†]

79 STANFORD LAW REVIEW ONLINE __ (forthcoming 2026)

When a hiring manager’s decisions produce a disparity, the causes are opaque, the data are noisy, and isolating the source of the disparity from “holistic judgment” is nearly impossible. When an algorithm produces a disparity, practices are specified, outputs are reproducible, and biases can be measured under controlled conditions. Yet scholars argue that algorithms make civil rights enforcement *harder*.

This consensus has it backwards. Disparate impact law is better suited to algorithms than to the humans it was designed to police. Algorithms are *consistent*: they apply the same criteria to every input, producing the cleanest statistical evidence the law has ever seen. They are *replicable*, making the business necessity defense specific, the less-exclusionary-alternative analysis tractable, and the algorithm identifiable as the particular practice that has eluded the doctrine since *Wal-Mart v. Dukes*. Algorithms are *flat*: a single process, not a web of cognition, culture, and bias, making the source of disparity easier to isolate.

Society should not fear algorithms. It should welcome them—not because they are fair, but because they make disparities easier to expose. If this advantage has not produced successful litigation, it is because the law is starved of information. Plaintiffs cannot challenge what they cannot see. The civil rights community has misdiagnosed information asymmetry as legal inadequacy, proposing new law when we really need new infrastructure—algorithmic testing, outcome reporting, and audit rights—to let the law see what it was built to expose.

^{*} Associate Professor of Law, Maurice A. Deane School of Law at Hofstra University. J.D., Yale Law School; Ph.D., University of Michigan; M.S. in Cybersecurity, New York University Tandon School of Engineering.

[†] James Edmund and Margaret Elizabeth Hennessey Corry Professor, Florida State University College of Law.

INTRODUCTION

When a human gatekeeper generates disparate outcomes, the causes are opaque, the data are noisy, and isolating the source of the disparity from “holistic judgment” is nearly impossible. When an algorithm produces a disparity, practices are specified, outputs are reproducible, and biases can be measured under controlled conditions. Yet scholars argue that algorithms make civil rights enforcement *harder*, not easier. The concerns are familiar: biased training data reproduces historical discrimination at scale;¹ opaque models defeat accountability;² and the elements of disparate impact do not map cleanly onto algorithmic systems.³ This anxiety has produced a near-consensus that existing civil rights law cannot keep up with algorithmic decisionmaking, and that new law is needed.⁴

This consensus has it backwards. Disparate impact law is better suited to algorithms than to the humans it was designed to police.⁵ Algorithms apply the same criteria to every input, producing the cleanest statistical evidence the

¹ See Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671, 671-701 (2016) (identifying mechanisms by which datamining produces discriminatory outcomes, including biased training data and proxy discrimination); Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE L.J. 2218, 2218 (2019) (“In a racially stratified world, any method of prediction will project the inequalities of the past into the future.”).

² See Brandon L. Garrett & Cynthia Rudin, *The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal Justice*, 109 CORNELL L. REV. 561, 563-64 (2024) (“‘[B]lack box’ AI designed to be non-interpretable[] mean[s] that its processes cannot be fully understood by laypeople or even by experts. . . . [In] the criminal justice system . . . black box AI poses risks to both public safety and to fundamental human and constitutional rights.”).

³ See Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857, 907 (2017) (arguing that disparate impact doctrine is “a poor fit” for algorithmic employment decisions because the statutory framework’s elements do not map cleanly onto data-driven systems); Ifeoma Ajunwa, *An Auditing Imperative for Automated Hiring Systems*, 34 HARV. J.L. & TECH. 621, 629-30 (2021) (algorithmic opacity and trade secret protection obstruct disparate impact claimants).

⁴ See, e.g., Robert Bartlett et al., *Consumer-Lending Discrimination in the FinTech Era*, 143 J. FIN. ECON. 30 (2022) (identifying racial disparities in algorithmic lending); Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249 (2008) (proposing safeguards for automated government decisionmaking).

⁵ This Essay synthesizes two research agendas. One of us has argued that U.S. consumer protection suffers from a structural detection gap: the legal system’s picture of harm is drawn from the subset of injuries that become visible enough to litigate, and this conspicuous subset is mistaken for the whole. Yunsieg P. Kim, *Incognito Consumer Harm*, 104 WASH. U. L. REV. __ (forthcoming). The other has developed a constitutional framework establishing disparate impact’s permissibility under contemporary equal protection doctrine. Jacob Eisler, *Racial Gateways*, 79 ALA. L. REV. __ (forthcoming). These frameworks suggest that the doctrine is structurally sound but operationally starved of information.

law has ever seen. Their outputs are deterministic and reproducible, making it possible to test a gatekeeper’s justification for a practice and to propose concrete, less exclusionary alternatives. And because an algorithm is a single, flat process rather than a web of cognition, culture, and bias, it is identifiable as the particular practice that statutes require plaintiffs to specify.⁶

Even the most commonly cited challenge—that algorithms are “black boxes”⁷—gets the problem backwards: human cognition is the *original* black box, and disparate impact was built to police opaque decisions through outcomes, not mechanisms.⁸ If this structural advantage has not produced successful litigation, the reason is not legal but informational: plaintiffs cannot challenge what they cannot see. But the civil rights community has misdiagnosed information asymmetry as legal inadequacy. Algorithms may be preferable for the pursuit of equity not because they are fair, but because they make disparities easier to expose.

The wider status of disparate impact remains contested,⁹ but all agree that the doctrine is useful for smoking out discriminatory intent that would otherwise escape proof. Whether it also serves a broader structural-correction

⁶ 42 U.S.C. § 2000e-2(k)(1)(A)(i) (requiring complainants to demonstrate “a particular employment practice” causing a disparate impact); *see* *Wal-Mart Stores, Inc. v. Dukes*, 564 U.S. 338, 355-60 (2011) (plaintiffs failed to identify a sufficiently uniform “practice” to establish commonality across 1.5 million claims); *see also* *Tex. Dep’t of Hous. & Cmty. Affs. v. Inclusive Cmty. Project*, 576 U.S. 519, 535 (2015) [*Inclusive Communities*] (extending disparate-impact liability to the Fair Housing Act by analogy to Title VII).

⁷ *See, e.g.,* Landyn Wm. Rookard, *The Common Threats of Artificial Intelligence and Privatization*, 12 *TEX. A&M L. REV.* 831, 870 (2025) (“[A]dvanced machine learning algorithms may be black boxes by their nature, making it inherently impossible to trace their logic from the inputs to their outputs.”); Benedict Sheehy & Yee-Fui Ng, *The Challenges of AI Decision-Making in Government and Administrative Law: A Proposal for Regulatory Design*, 57 *IND. L. REV.* 665, 672 (2024) (“[T]he inability to explain AI decision-making is a significant problem when it comes to giving reasons.”).

⁸ *See* *Griggs v. Duke Power Co.*, 401 U.S. 424, 431-32 (1971) (holding that Title VII “proscribes not only overt discrimination but also practices that are fair in form, but discriminatory in operation” and evaluating the employer’s practices by their consequences, not their internal rationale). The greater ease of identifying algorithmic disparate impact is especially strong if the “core” purpose is taken to be “smoking out” hidden or unconscious discriminatory intent. *See infra* note 9 and accompanying text.

⁹ *See Inclusive Communities*, 576 U.S. at 540 (disparate-impact liability “plays a role in uncovering discriminatory intent” by allowing plaintiffs to reach biases that evade disparate-treatment classification); *cf. Ricci v. DeStefano*, 557 U.S. 557, 594-95 (2009) (Scalia, J., concurring) (questioning whether disparate-impact provisions are consistent with the Equal Protection Clause); *Washington v. Davis*, 426 U.S. 229, 239-42 (1976) (discriminatory purpose required under the Equal Protection Clause). For a discussion of the broader debate, *see* Samuel R. Bagenstos, *Disparate Impact and the Role of Classification and Motivation in Equal Protection Law After Inclusive Communities*, 101 *CORNELL L. REV.* 1115, 1136 (2016).

function is a question this Essay need not reach. Its claim holds on the common ground: disparate impact is easier to detect when the decisionmaker is an algorithm, because algorithms are easier to test under controlled conditions.

This Essay proceeds in two parts. Part I identifies three properties of algorithmic decisionmaking—consistency, replicability, and flatness—that make algorithms more amenable to disparate impact analysis than the human judgment that disparate impact was originally built to police. Part II explains why this legal advantage has not been realized in practice: plaintiffs lack the data necessary to invoke a framework that is structurally sound but operationally blind. The priority is not new law for the algorithmic age, but new infrastructure to activate the law we already have.

I. THREE PROPERTIES

Algorithms have three properties that make them better suited to disparate impact analysis than human decisionmakers. Algorithms are *consistent*: they apply the same criteria to every input. They are *replicable*: their outputs can be reproduced and tested under controlled conditions. They are *flat*: an algorithm is a single process, not a layered system of cognition, culture, and bias. Each of these properties resolves a problem that has long undermined disparate impact claims against humans.

A. Consistency

An algorithm applies the same criteria, weighted the same way, to every input. A human decisionmaker does not. A hiring manager’s evaluation of two identical resumes will differ depending on the time of day, the order in which they are reviewed, and the manager’s mood, fatigue, and implicit associations. These variables are uncontrolled and unobservable. Algorithms’ behavior, in contrast, is “consistent: given the same inputs, they should reliably produce the same outputs without significant variation.”¹⁰ Even large language models like ChatGPT, which introduce controlled randomness, draw from a fixed distribution: individual outputs may vary, but aggregate patterns across a population remain stable and testable.¹¹

¹⁰ Rebecca Crootof, Margot E. Kaminski & W. Nicholson Price II, *Humans in the Loop*, 76 VAND. L. REV. 429, 463-64 (2023).

¹¹ See, e.g., Jiafu An et al., *Measuring Gender and Racial Biases in Large Language Models: Intersectional Evidence from Automated Resume Evaluation*, 4 PNAS NEXUS (2025) (using five LLMs including ChatGPT and Claude to score ~361,000 resumes with randomized identities and finding statistically significant racial and gender disparities stable across 500 attempts).

This consistency resolves the statistical proof problem that has long plagued disparate impact claims.¹² When human decisionmaking is inconsistent, as it often is, the signal is noisy and confounded. It is difficult to demonstrate that a pattern of outcomes reflects a single practice rather than a collection of idiosyncratic decisions. When an algorithm is consistent, as it always is, the signal is clean: run the model on a population, disaggregate the outputs by race, and perform a standard significance test.¹³ Algorithms produce the cleanest evidentiary environment that disparate impact has ever encountered.¹⁴

But clean statistical evidence is only the first step. The disparate impact framework also requires an identifiable practice, a defensible justification, and testable alternatives.¹⁵ Algorithmic replicability delivers all three.

B. Replicability

Disparate impact requires plaintiffs to identify a “particular employment practice” that causes the challenged disparity. That requirement has been the doctrine’s most persistent obstacle. *Wal-Mart Stores, Inc. v. Dukes* held that a policy of delegating discretion to individual managers was not a sufficiently identifiable “practice.”¹⁶ The problem was that the alleged discrimination was dispersed: thousands of managers made subjective calls in thousands of different ways, producing a pattern too diffuse to attribute to any single practice. With human decisionmakers, the challenged “practice” is a cognitive process that cannot be isolated, specified, or re-run.

An algorithm eliminates this problem. An algorithm *is* a particular employment practice: a specified, centralized process that applies defined

¹² See *Griggs*, 401 U.S. at 431-32; *Inclusive Communities*, 576 U.S. at 540-42.

¹³ See Bartlett et al., *supra* note 4, at 34-45 (Latino and Black borrowers paid approximately 5-8 basis points more on home-purchase loans).

¹⁴ See *Racial Gateways*, *supra* note 5 (developing an efficient-cause framework establishing disparate impact’s permissibility under contemporary equal protection doctrine); *Louisiana v. Callais*, No. 24-109 (U.S., re-argued Oct. 2025) (decision pending); *Students for Fair Admissions, Inc. v. President & Fellows of Harvard Coll.*, 600 U.S. 181 (2023) (finding race-conscious admissions programs violate the Equal Protection Clause).

¹⁵ 42 U.S.C. § 2000e-2(k)(1)(A)(i) (requiring identification of a particular employment practice and imposing on defendants the burden of demonstrating job-relatedness and business necessity); *id.* § 2000e-2(k)(1)(A)(ii) (complainants may prevail by demonstrating an alternative employment practice that respondent refuses to adopt).

¹⁶ 564 U.S. at 355-57 (2011). The Court acknowledged that delegated discretion can constitute a cognizable employment practice, *see* *Watson v. Fort Worth Bank & Tr.*, 487 U.S. 977, 990 (1988), but held that plaintiffs failed to identify a common mode of exercising discretion sufficient for Rule 23(a)(2) commonality.

criteria to every input. Because its outputs are deterministic, the algorithm can be run on a population, re-run with altered inputs, and shown to be the same process producing materially identical results. In *Mobley v. Workday, Inc.*, a federal court described an AI screening system as a “unified policy” sufficient to support a nationwide collective action.¹⁷ The algorithm was exactly the kind of identifiable practice that was missing in *Wal-Mart*.

Replicability also makes the subsequent stages of the burden-shifting framework tractable. Once a plaintiff establishes a prima facie case, the burden shifts to the defendant to demonstrate that the challenged practice is job-related and consistent with business necessity.¹⁸ When the practice is a human manager’s “holistic judgment,” the defense is vague and difficult to challenge with specificity. When the practice is an algorithmic model, the defendant must explain why these particular features, weighted in this particular way, producing these particular decision boundaries, are necessary. The specificity of the model compels specificity in the defense. When the model is an AI system whose internal reasoning the defendant itself cannot fully explain, the burden becomes heavier still.¹⁹ A company that deploys a model it cannot interpret will struggle to carry that burden.

The same logic applies to the statutory requirement that plaintiffs identify a less exclusionary alternative. With a human decisionmaker, that requirement is practically meaningless: one cannot propose a less exclusionary “alternative” to a cognitive process because one cannot redesign how a person thinks. With an algorithm, concrete alternatives are routine—remove a feature correlated with race, retrain on a more representative dataset, or substitute a simpler model.²⁰ Replicability means that each alternative can be tested and the difference measured.²¹ The doctrine’s

¹⁷ No. 3:23-cv-00770 (N.D. Cal. July 12, 2024) (allowing disparate impact claims to proceed against Workday under an agent liability theory); *id.* (N.D. Cal. May 16, 2025) (granting preliminary certification and describing Workday’s AI recommendation system as a “unified policy” applicable to all putative collective members).

¹⁸ 42 U.S.C. § 2000e-2(k)(1)(A)(i); see *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425-36 (1975) (operationalizing the business necessity standard by validating employment tests against job performance).

¹⁹ See Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 *FORDHAM L. REV.* 1085, 1094-1100 (2018) (distinguishing inscrutability from nonintuitiveness as sources of the “black box” problem); Jenna Burrell, *How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms*, 3 *BIG DATA & SOC’Y*, Jan.-June 2016, at 1, 1-5 (model opacity is distinct from corporate secrecy or technical illiteracy).

²⁰ See An et al., *supra* note 11.

²¹ 42 U.S.C. § 2000e-2(k)(1)(A)(ii); see Emily Black, John Logan Koepke, Pauline T. Kim, Solon Barocas & Mingwei Hsu, *Less Discriminatory Algorithms*, 113 *GEO. L.J.* 53, 62-

burden-shifting framework, originally designed for subjective human judgment, becomes *more* tractable when the decisionmaker is a machine. What remains is the most fundamental objection, and the one most frequently directed at AI: that the machine is a black box.

C. Flatness

The black box objection has dominated the scholarly response to algorithmic disparate impact.²² The objection may appear to be the strongest for AI systems. Deep learning models can contain billions of parameters, and even their developers cannot fully explain the reasoning process that produces a given output.²³ If we cannot see inside the model, we cannot hold it accountable. But human cognition is the *original* black box: opaque, unreplicable, and resistant to inspection. Disparate impact was designed from the beginning to police decisionmaking processes whose internal workings could not be observed.²⁴

The reason that the human black box is deeper than the algorithmic one is not simply opacity. It is that human decisionmakers are layered systems. A hiring manager’s decision reflects the interaction of explicit criteria, implicit bias, institutional culture, interpersonal dynamics, emotional states, and cognitive heuristics: multiple overlapping motivational systems, each capable of producing exclusionary outcomes independently. These layers make it extraordinarily difficult to isolate which aspect of the decisionmaker’s process produced the disparity. By comparison, algorithms are flat: one model, one set of inputs, one set of weights, and one output, with no overlapping motivational systems whose independent contributions must be disentangled. Even an AI neural network with billions of parameters is

71 (2024) (model multiplicity—the existence of many comparably accurate models for a given prediction task—means that less discriminatory alternatives will often be available and testable).

²² See Garrett & Rudin, *supra* note 2; Rookard, *supra* note 7; Sheehy & Ng, *supra* note 7; Ashley Deeks, *The Judicial Demand for Explainable Artificial Intelligence*, 119 COLUM. L. REV. 1829, 1829 (2019) (“A recurrent concern about machine learning algorithms is that they operate as ‘black boxes.’”).

²³ See Tomek Korbak et al., *Chain of Thought Monitorability: A New and Fragile Opportunity for AI Safety 2* (2025), https://tomekkorbak.com/cot-monitorability-is-a-fragile-opportunity/cot_monitoring.pdf (“[t]he opacity of advanced AI agents underlies many of their potential risks” and even chain-of-thought reasoning traces—the most advanced method for interpreting AI reasoning—are “incomplete representations” of the model’s actual reasoning process) (authored by researchers representing OpenAI, Google, and Anthropic, among others).

²⁴ See *Griggs*, 401 U.S. at 432 (“Congress directed the thrust of the Act to the consequences of employment practices, not simply the motivation.”).

computationally complex, but motivationally flat—a single process with defined inputs and a defined output, not a web of cognitive and cultural systems.

Disparate impact is an outcomes-based doctrine. The prima facie case requires showing *that* a practice produces disparate results, not *how* it does so internally.²⁵ That showing is straightforward even when the model is a complete black box: run it on a dataset, disaggregate the outputs by race, and measure the disparity.²⁶ The model’s internal opacity is irrelevant to this analysis because the analysis operates entirely on inputs and outputs. This is true whether the model is a simple algorithm or ChatGPT.²⁷ Moreover, at the business necessity stage, AI opacity does not help the defendant. A company deploying a model whose reasoning even its own engineers cannot explain faces a steeper burden to demonstrate that the practice is job-related. The doctrine thus creates incentives for interpretability, which is the very outcome that the explainable AI community has been pursuing, now backed by legal force.²⁸

This structural simplicity also addresses the judicial anxieties driving the Supreme Court’s narrowing of effects-based liability. *Inclusive Communities* warned that disparate impact must not become a means to “second-guess” valid governmental and private priorities.²⁹ *Brnovich v. Democratic National Committee* narrowly interpreted Section 2 of the Voting Rights Act’s results test in a manner reflecting broader discomfort with effects-based claims overriding legitimate state choices.³⁰ Both cases reflect anxiety about liability regimes that second-guess human discretion. That concern lacks force when the challenged practice is not a judgment call, but a model. There is no human discretion to chill. There is only a system to audit.

As more decisions migrate from humans to algorithms, the factual

²⁵ See *Griggs*, 401 U.S. at 431-32; Selbst & Barocas, *supra* note 19, at 1094-1100. The distinction also maps onto the prima facie analysis: even Selbst and Barocas’s category of “inscrutability” does not defeat the prima facie case.

²⁶ See An et al., *supra* note 11.

²⁷ See *id.*

²⁸ See Selbst & Barocas, *supra* note 19, at 1113-24 (arguing that the demand for explainability in machine learning reflects legitimate normative concerns about autonomy and accountability, but that the appropriate response varies by context).

²⁹ *Inclusive Communities*, 576 U.S. at 540 (FHA “is not an instrument to force housing authorities to reorder their priorities”); *id.* at 544 (disparate impact liability must not “displace valid governmental and private priorities”).

³⁰ 594 U.S. 647, 668-72 (2021) (adopting a multifactor test for Section 2 of the VRA that significantly raised the bar for effects-based challenges to voting regulations). The Court distinguished Section 2’s results test from Title VII-style disparate impact, *id.* at 672-74.

predicates that trouble the Court will gradually disappear. Subjective discretion gives way to specified criteria. The impossibility of proposing concrete alternatives gives way to model multiplicity. The risk of de facto quotas gives way to measurable, auditable outputs. The law that the Court is narrowing for human contexts may be most easily applied to algorithmic ones. The retrenchment of disparate impact doctrine is, paradoxically, clearing the ground for exactly the kind of claims that even the most ardent disparate impact skeptics should accept.³¹

II. STARVED OF INFORMATION

The argument thus far still has an obvious weakness. If disparate impact is structurally better suited to algorithms than to humans, why has no wave of successful algorithmic disparate impact litigation materialized? The answer is not legal but informational: the law is starved of the information it needs.

A. *Two Barriers*

Consider what information a plaintiff must have in hand to bring a disparate impact claim against an algorithmic system, and what information is systematically unavailable.

The first barrier is identification. In many employment, housing, and lending contexts, applicants do not know that an algorithm processes their applications, let alone which one produces the disparity. Even a plaintiff who suspects algorithmic screening cannot see the decision pipeline: which particular algorithm does what in a chain of intake screening, resume-parsing, and final scoring.³²

³¹ The constitutional permissibility of effects-based liability remains contested. *See Racial Gateways*, *supra* note 5. *Cf.* *Washington v. Davis*, 426 U.S. 229, 239-48 (1976) (requiring discriminatory purpose under the Equal Protection Clause while preserving statutory effects-based claims under Title VII); *Inclusive Communities*, 576 U.S. at 540, 544 (sustaining statutory disparate impact while cautioning against overreach).

³² *See Incognito Consumer Harm*, *supra* note 5 (describing the “survivorship bias” in enforcement: the legal system’s picture of harm is drawn from the subset of injuries that become visible enough to litigate). In *Mobley*, the plaintiff applied to over 100 jobs using Workday’s AI platform and was rejected by all of them, often within an hour and during non-business hours. *Mobley*, No. 3:23-cv-00770. In *iTutorGroup*, the EEOC challenged an algorithm that automatically rejected applicants based on age. *EEOC v. iTutorGroup, Inc.*, No. 1:22-cv-02565 (E.D.N.Y.) (consent decree approved Sept. 8, 2023; \$365,000 settlement). *See also* *United States v. Meta Platforms, Inc.*, No. 1:22-cv-05187 (S.D.N.Y. 2022) (\$115,054 civil penalty; DOJ’s first challenge to algorithmic bias under the FHA).

The second barrier is evidence. Even plaintiffs who know that an algorithm was used lack access to aggregate outcome data disaggregated by, say, race or gender. They see only their own outcomes, accepted or rejected, and cannot see the statistical distribution across demographic groups. Performing the analysis necessary to establish a prima facie case requires both data and technical capacity that many individuals and advocacy organizations lack.³³

These information barriers are not unique to algorithms. Plaintiffs challenging human decisionmakers also lack access to the ‘deep’ causal driver of a disparity and to aggregate outcome data. But with algorithms, each barrier is *informational*, not legal. The law could handle the case, but cases never get filed because necessary information is unavailable. Courts have created doctrinal friction such as standing limitations, trade secret objections to discovery, and uncertain vendor liability,³⁴ but these problems are secondary. The dominant constraint is that most algorithmic discrimination is never detected at all. This is entirely solvable. The same consistency, replicability, and flatness that make the doctrine work also make detection infrastructure feasible: an algorithm can be systematically tested in ways that a human mind cannot.

B. Detection-First Reforms

The civil rights community has misdiagnosed information asymmetry as legal inadequacy, proposing new law when existing law is structurally sound but operationally blind.³⁵ The priority should be detection infrastructure that makes algorithmic disparate impact observable, so that the law can operate.

The most novel intervention is tester standing for algorithmic systems. Civil rights enforcement has a long history of using testers—individuals who

³³ Many scholarly proposals for addressing algorithmic discrimination presuppose that discrimination has already been detected. *See, e.g.,* Ajunwa, *supra* note 3, at 624 (proposing mandatory auditing of automated hiring systems but presupposing that discrimination has already been detected); Mayson, *supra* note 1, at 2275 (analyzing corrective measures that require knowledge of bias). This diagnosis-cure mismatch is a central topic of *Incognito Consumer Harm*, *supra* note 5.

³⁴ *See* *TransUnion LLC v. Ramirez*, 594 U.S. 413, 431-34 (2021) (narrowing Article III standing requirements for statutory violations, with potential implications for algorithmic discrimination plaintiffs who suffered no traditional “concrete injury”). Algorithms can also produce disparities through proxy variables that do not map neatly onto traditional protected classes, further complicating the task of establishing standing and identifying the relevant “practice.”

³⁵ *See Incognito Consumer Harm*, *supra* note 5 (arguing for a “detection-first paradigm” in consumer protection and proposing regulatory reform designed to surface hidden harm).

apply for housing or work, for example, not to obtain an apartment or a job but to test for discrimination. Courts have recognized tester standing under the Fair Housing Act since *Havens Realty Corp. v. Coleman*.³⁶ Extending tester methodology to algorithmic systems would mean sending synthetic applicant profiles to detect disparate treatment or impact. Computer science literature has already proven audit methods using fictitious profiles.³⁷ It requires no change to substantive law, only an extension of testing long applied to human gatekeepers.

Another reform is mandatory outcome reporting. Employers, lenders, and landlords using algorithms in covered decisionmaking contexts should be required to report outcome data disaggregated by protected class. The model is EEO-1 reporting, which requires covered employers to report workforce composition data by race, ethnicity, and sex.³⁸ Extending this framework to algorithmic decisions would make statistical patterns necessary for disparate impact claims observable without requiring individual victims to detect them.

The third is algorithmic audit rights for regulators. Current transparency frameworks give individuals access to their own data but do not give regulators access to aggregate outcome distributions across protected classes. An audit right requiring disclosure of aggregate results, not source code or trade secrets, would provide the evidentiary foundation that currently does not exist.³⁹ Each of these interventions is designed not to change what

³⁶ 455 U.S. 363, 373-74 (1982) (tester given false information about housing availability had standing to sue under the FHA regardless of intent to rent). On the legality of fictitious profiles used to audit algorithmic systems, see Sandvig v. Barr, 451 F. Supp. 3d 73, 88-92 (D.D.C. 2020) (fictitious audit profiles do not violate the CFAA because mere terms-of-service violations do not trigger liability); Van Buren v. United States, 593 U.S. 374, 389-91 (2021) (narrowing “exceeds authorized access” to a gates-up-or-down inquiry, further supporting algorithmic audit legality).

³⁷ See, e.g., Joshua Asplund et al., *Auditing Race and Gender Discrimination in Online Housing Markets*, 14 *ICWSM* 24-35 (2020) (demonstrating a controlled audit methodology using fictitious profiles to expose race and gender discrimination in online housing platforms).

³⁸ See 29 C.F.R. §§ 1602.7-1602.11 (employer reporting obligations under Title VII); cf. *id.* §§ 1602.12-1602.14 (recordkeeping).

³⁹ See CAL. CODE REGS. tit. 11, §§ 7155(a)(1), 7157(a)-(e), 7222(b)(2)-(3) (2026) (requiring pre-use risk assessments for certain high-risk processing; requiring only summary submissions and executive attestations to the CPPA, without routine filing of disaggregated outcome data); Regulation (EU) 2024/1689, arts. 27(1)-(3), 113, 2024 O.J. (L 1689) 1 (requiring certain pre-deployment fundamental-rights impact assessments without periodic aggregate outcome reporting, applicable Aug. 2, 2026). These developments narrow but do not close the gap we identify. Among states, Illinois makes it a civil-rights violation in employment to use AI that “has the effect of subjecting employees to discrimination” on the basis of protected classes. Pub. Act 103-0804, §§ 5, 99 (codified at 775 ILL. COMP. STAT.

the law says, but to let the law see what it was built to expose.

CONCLUSION

The civil rights community has treated the rise of algorithmic decisionmaking as a threat. We, in contrast, suggest that it is an *opportunity*. Every human gatekeeper replaced by an algorithm is a decisionmaker whose disparate outputs become testable, auditable, and legally actionable. This is not a claim that current algorithmic systems are unbiased, but that algorithmic bias—unlike human bias—is the kind that civil rights law is best equipped to expose. If that detection infrastructure is built, the migration of high-stakes decisions from humans to algorithms would represent not a crisis for civil rights enforcement but the largest expansion of its practical capacity in decades. The question for future work is not how to resist the rise of algorithms—which, for better or worse, may already be irreversible⁴⁰—but how to design the legal and regulatory infrastructure that makes algorithms work for civil rights. The tools are available. What remains is the will to let the law see what it was built to expose.

5/2-102(L)(1)) (Jan. 1, 2026).

⁴⁰ See, e.g., Crootof, Kaminski & Price, *supra* note 10, at 432 (“Artificially intelligent algorithms are being integrated into decisionmaking processes at mind-boggling speed and scale”); Kate Crawford & Jason Schultz, *AI Systems as State Actors*, 119 COLUM. L. REV. 1941, 1942 (2019) (“Every month, more algorithmic and predictive technologies are being applied in domains such as healthcare, education, criminal justice, and beyond.”); Note, *Machine Rulemaking: Arbitrary and Capricious Review in the Age of AI*, 138 HARV. L. REV. 1821, 1821 (2025) (“[U]se of AI/ML is no longer limited to tech companies or the private sector . . . 157 [AI/ML] use cases across 64 agencies.”).