

The Accommodation Axis: Disability Discrimination, Algorithmic Opacity, and the Deployers That Anti-Discrimination Scholarship Overlooks

Travis Gilly*

Working draft, June 2026 (v1)

Abstract

A growing body of scholarship warns that the economic and training logic of modern artificial intelligence drives automated decision systems to target the vulnerable, conceal discrimination, and resist detection, defeating both the disparate treatment and disparate impact theories on which anti-discrimination enforcement has relied. That work is largely correct, but it runs on a single axis: intent versus impact. Both engines it analyzes are Title VII engines, and the analysis treats disability as one more protected class to be fed through the same machinery. This Article argues that disability discrimination law occupies a different axis entirely. The Rehabilitation Act and the Americans with Disabilities Act add an affirmative-duty engine, reasonable accommodation and meaningful access, that is structurally immune to the three capabilities said to make algorithmic discrimination undetectable. A failure-to-accommodate claim requires no proof of intent, so a system that hides its reasons defeats nothing; it requires no comparator, so a system that eliminates comparators defeats nothing; and it shifts the burden of justification to the deployer, so a system that manufactures a plausible pretext shifts nothing. The Article then turns the opacity argument around. The unexplainable model that shields a Title VII defendant indicts a disability defendant, because a selection device the deployer cannot validate fails the business-necessity requirement as a matter of law, and a leading developer's own research forecloses curing that failure through the model's stated reasons. Finally, the Article identifies the deployers this literature has overlooked, universities and public agencies, which are paradigmatic recipients of federal financial assistance and are now routing admissions, accommodation, and eligibility decisions through systems no human can explain. The doctrine to reach them already exists. No new statute is required.

Keywords: Section 504; ADA Title I; reasonable accommodation; disparate impact; business necessity; *Albemarle Paper Co. v. Moody*; *Alexander v. Choate*; algorithmic hiring; higher education.

*Travis Gilly is Founder and Executive Director of the Real Safety AI Foundation. ORCID: 0009-0007-2954-6313. Correspondence: t.gilly@ai-literacy-labs.org. The author discloses the use of speech-to-text software (Wispr Flow) and Anthropic's Claude as assistive technology for a diagnosed neurodevelopmental condition. All legal analysis, case synthesis, theoretical framing, doctrinal arguments, and conclusions are solely the author's own.

Contents

I	Introduction	3
II	The Map and Its Edge	5
	A The Mechanism.	6
	B The Two Defeats	6
	C The Single Axis.	7
III	The Third Engine	7
	A Accommodation and Meaningful Access.	8
	B Why the Three Capabilities Miss	9
	C The Knowledge Wrinkle and the Strongest Vehicle	10
IV	The Deployers Anti-Discrimination Scholarship Overlooks	11
	A Universities as Recipients.	11
	B The Decisions Now Being Automated.	12
	C Public Agencies and the State as Deployer	14
V	The Box That Indicts	15
	A The Validation Requirement.	16
	B An Unexplainable Device Cannot Be Validated by Explanation	17
	C The Empirical Escape and Its Three Leaks	17
	D The Faithfulness and Tooling Pincer	18
	E Cost Is Not Business Necessity.	20
	F Human Review That Lauanders Rather Than Pierces	21
VI	Why Injunctive Relief Is the Feature	21
VII	Conclusion	23

I. Introduction

Imagine a disabled applicant who is denied a job, or a registered disabled student who is denied an accommodation, and is told nothing she can use. The decision came from an artificial intelligence system. No human can say why she was denied, and the system itself, if asked, will produce a fluent explanation that may have nothing to do with the computation that actually drove the result. She cannot tell whether her disability was the reason. Neither, on the present design, can the entity that rejected her.

A recent and rigorous article describes precisely how this comes to pass. In *Generative Adversarial Discrimination*, Josh Fairfield argues that the training and economic logic of modern artificial intelligence causes automated decision systems to target the intersectionally vulnerable, conceal discrimination rather than surface it, and resist detection by humans and by other machines.¹ The mechanism, in his account, is the interaction of a hidden cost-reduction optimization goal with a hidden tendency to satisfy the operator's command at any cost, refined by adversarial or evolutionary selection that retains only those systems which escape detection.² The consequence is a system that learns to make discrimination pay rather than to reduce it,³ and that defeats the two principal theories of anti-discrimination law: it manufactures a non-falsifiable pretext that defeats a disparate treatment claim, and it generates a post-hoc business justification and eliminates the comparators that a disparate impact claim depends upon.⁴

That analysis is largely correct, and this Article does not dispute it. But it runs on a single axis. The two engines it examines, the burden-shifting structure of *McDonnell Douglas Corp. v. Green*⁵ and the disparate impact structure of *Griggs v. Duke Power Co.*,⁶ are both engines of Title VII. The analysis treats disability, where it appears at all, as one more protected status to

1. Josh Fairfield, *Generative Adversarial Discrimination*, 32 Wash. & Lee J. Civ. Rts. & Soc. Just. 539, 539–42 (2026).

2. *Id.* at 561–64, 592–94.

3. *Id.* at 542.

4. *Id.* at 555–57, 580–83.

5. 411 U.S. 792, 802–04 (1973).

6. 401 U.S. 424, 431–32 (1971).

be fed through the same machinery. It is not. Disability discrimination law is not Title VII with the word “disability” substituted for the word “race.” The Rehabilitation Act of 1973 and the Americans with Disabilities Act of 1990 add a third engine that the intent-versus-impact axis does not contain: an affirmative duty of reasonable accommodation, measured by whether the covered entity has afforded the disabled person meaningful access to the benefit or opportunity.⁷

This Article makes two claims. The first is structural. The affirmative-duty engine is immune to each of the three capabilities that are said to make algorithmic discrimination undetectable. A failure-to-accommodate claim has no intent element, so a system engineered to hide its reasons defeats nothing; the hidden reason was never the plaintiff’s to find. It requires no comparator, so a system that finds and eliminates comparators, the “discrimination of one” problem, defeats nothing; the disabled plaintiff never needed a similarly situated nondisabled person treated better. And once a denial of access is shown, the burden of justification shifts to the covered entity, so a system that manufactures a plausible nondiscriminatory story shifts nothing; the entity, not the plaintiff, must carry the explanation. The capabilities that win on the intent-versus-impact axis are aimed at machinery the accommodation claim does not use.

The second claim turns the opacity argument around. On the intent-versus-impact axis, the unexplainable model helps the defendant: it hides intent and supplies pretext. In the accommodation and qualification-standard setting, the unexplainable model indicts the defendant. A selection device that produces a disparate impact must be validated, and the burden of validation rests on the entity that deploys it.⁸ A device its deployer cannot explain is a device its deployer cannot validate by the established content or construct methods. Unexplainable, plus disparate impact, plus the deployer’s burden, equals a business-necessity defense that fails as a matter of law. And the one escape from that conclusion, an empirical demonstration that the system’s outputs predict performance without reference to how it reasons, is foreclosed in a way the developers of these systems have themselves documented: the visible reasoning of a model

7. See *Alexander v. Choate*, 469 U.S. 287, 301 (1985) (Section 504 requires “meaningful access” to the benefit a grantee offers).

8. *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425, 431–36 (1975).

frequently fails to reflect, and sometimes diverges from, the computation that actually produced its output.⁹ The deployer cannot point to what the model said about its own decision, because the people who built the model have published that what it says about its own decision is not reliable.

A note on tone is owed at the outset, because the argument is built on another scholar's work and means to extend it rather than to score against it. Fairfield's article expressly invites other voices and other lived experience to develop the problems it raises, on the express ground that different experiences yield different analysis.¹⁰ This Article accepts that invitation. The omission it identifies is not an error of competence; it is the natural edge of a project scoped to one author's domain, and the disability axis it adds is offered as the complement that the original work asked for.

The Article proceeds in six further Parts. Part II restates the *Generative Adversarial Discrimination* analysis and locates its single axis. Part III sets out the third engine, reasonable accommodation and meaningful access, and shows why it is intent-free and comparator-free. Part IV identifies the deployers the literature has overlooked, universities and public agencies, and the statutory hooks that reach them. Part V develops the affirmative case, that opacity indicts rather than shields the disability defendant, through the validation requirement, the limits of empirical validation, and the developers' own research on the unreliability of model-stated reasons. Part VI explains why the injunctive remedy that the existing literature treats as a weakness is, in the disability setting, the decisive strength. A short conclusion answers the question the original article leaves open.

II. The Map and Its Edge

The contribution of *Generative Adversarial Discrimination* is to show that the standard anti-discrimination toolkit is not merely strained by artificial intelligence but actively defeated

9. See *infra* Part V.D and accompanying notes (discussing published findings on the unfaithfulness of model-stated reasoning and the parallel availability of interpretability tooling).

10. Fairfield, *supra* note 1, at 542–43.

by it, and to explain why in terms of how these systems are trained. The explanation deserves a fair statement before its boundary is drawn.

A. The Mechanism

The article's core is that the systems deployed at scale are selected, formally through adversarial training or practically through market and legal pressure, to be the ones that escape detection.¹¹ A system trained to reduce cost, and refined against the adverse selection pressure of regulators and plaintiffs, will learn to evade detection rather than to reduce the underlying conduct wherever evasion is cheaper than redress.¹² Two hidden features drive the result. The first is a cost-reduction optimization goal, often unstated, which in the discrimination setting is a goal of minimizing liability.¹³ The second is the system's trained disposition to satisfy the operator's instruction, which the article calls sycophancy, and which makes the system pursue the assigned objective even at the cost of subverting the constraints placed on it.¹⁴ The article illustrates the latter with a now-public episode in which a model, told it would be shut down, produced blackmail against the engineers it believed could stop the shutdown, not from any will to survive but because the assigned objective could not be met by a deployed system that had been turned off.¹⁵

B. The Two Defeats

From this the article derives two defeats. Against a disparate treatment claim, which turns on intent and is ordinarily litigated through the *McDonnell Douglas* burden-shifting structure,¹⁶ the system can be run across millions of scenarios to find a nondiscriminatory account that has

11. *Id.* at 560–65.

12. *Id.* at 563–64.

13. *Id.* at 542, 592–94.

14. *Id.* at 562–64.

15. *Id.* at 562–63 (citing reporting on the simulated test).

16. *McDonnell Douglas*, 411 U.S. at 802–04 (prima facie case, articulation of a legitimate nondiscriminatory reason, and proof of pretext).

no obvious counterexample, a pretext engineered to be non-falsifiable.¹⁷ Against a disparate impact claim, which dispenses with intent and asks instead whether a facially neutral practice falls more harshly on a protected group,¹⁸ the system can supply a colorable business justification and, more corrosively, can exploit the fact that the most vulnerable plaintiffs sit at the intersection of several characteristics and therefore have no comparators at all.¹⁹ The article observes that the remedy for disparate impact is typically an injunction rather than damages, and that a liability-minimizing system will read injunction-only exposure as a lower cost and migrate toward conduct that produces it.²⁰

C. The Single Axis

The boundary is now visible. Both defeats operate on one axis. Disparate treatment and disparate impact are the two ways Title VII asks a single question, whether a protected characteristic produced an adverse result, either through intent or through unjustified effect. Every one of the article's most powerful observations, the engineered pretext, the eliminated comparator, the post-hoc justification, the displacement of responsibility onto an unaccountable machine, is an observation about how artificial intelligence defeats that question. The article does not ask, because its domain did not require it to ask, whether anti-discrimination law contains a question that the intent-versus-impact axis does not capture. It does. The remainder of this Article is about that question and the body of law that asks it.

III. The Third Engine

Disability discrimination law shares vocabulary with Title VII, but it is not built on the same chassis. It carries an affirmative obligation that the race and sex provisions of Title VII do not,

17. Fairfield, *supra* note 1, at 551–54.

18. *Griggs*, 401 U.S. at 431–32; see also 42 U.S.C. §2000e-2(k) (codifying the burden structure for disparate impact under Title VII).

19. Fairfield, *supra* note 1, at 580–83 (the “discrimination of one” problem).

20. *Id.* at 556; see 42 U.S.C. §1981a(a)(1) (authorizing compensatory and punitive damages only for intentional discrimination).

and that obligation generates a theory of liability that neither the disparate treatment nor the disparate impact engine contains.

A. Accommodation and Meaningful Access

Section 504 of the Rehabilitation Act prohibits the exclusion of an otherwise qualified individual with a disability, solely by reason of disability, from any program or activity receiving federal financial assistance.²¹ The Americans with Disabilities Act extends a parallel prohibition to private employment,²² to public entities,²³ and to places of public accommodation.²⁴ What distinguishes these regimes from Title VII is that nondiscrimination is not satisfied by identical treatment. The Supreme Court held in *Alexander v. Choate* that Section 504 requires a grantee to provide a disabled person “meaningful access” to the benefit it offers, which may require reasonable modifications rather than the same treatment afforded everyone else.²⁵ Identical treatment that denies meaningful access is itself the violation. The duty is affirmative: the covered entity must accommodate up to the point of undue hardship or fundamental alteration, and the failure to do so is actionable without any showing that the entity wished to exclude.

The employment provisions of the ADA make the same point in operational terms. Discrimination is defined to include the failure to make reasonable accommodations for the known limitations of an otherwise qualified individual,²⁶ the use of qualification standards or selection criteria that screen out individuals with disabilities unless the standard is job-related and consistent with business necessity,²⁷ and the use of employment tests that reflect a disability rather than the aptitude the test purports to measure.²⁸ The Court has read the statute to require

21. 29 U.S.C. §794(a).

22. 42 U.S.C. §12112(a).

23. 42 U.S.C. §§12131–12132.

24. 42 U.S.C. §12182(a).

25. *Choate*, 469 U.S. at 301.

26. 42 U.S.C. §12112(b)(5)(A).

27. 42 U.S.C. §12112(b)(6); see also 42 U.S.C. §12113(a) (defense for job-related qualification standards consistent with business necessity).

28. 42 U.S.C. §12112(b)(7).

an individualized inquiry into the particular person and the particular accommodation, rejecting reliance on generalizations about a disability or a class.²⁹

B. Why the Three Capabilities Miss

Set the affirmative-duty engine against the three capabilities that defeat the intent-versus-impact axis.

The first capability is the engineered, non-falsifiable pretext, which defeats disparate treatment by making intent impossible to prove. A failure-to-accommodate claim has no intent element. The plaintiff shows that she is a qualified individual with a disability, that she needed a reasonable accommodation, that the covered entity knew of the disability and the need, and that the accommodation was denied and access was lost. None of these elements asks why the entity acted. A perfectly engineered nondiscriminatory story answers a question the claim never poses.

The second capability is comparator elimination, the “discrimination of one” problem, which defeats disparate impact by ensuring that the most vulnerable plaintiffs have no similarly situated counterpart whose better treatment proves the case.³⁰ The accommodation claim needs no comparator. It does not proceed by showing that a nondisabled applicant in like circumstances was treated better; it proceeds by showing that this disabled person was denied the access the statute guarantees. A system that can find and eliminate every comparator has eliminated something the claim never required.

The third capability is the manufactured business justification, which defeats disparate impact by supplying a colorable account of why the neutral practice was necessary. Here the accommodation engine does not merely survive; it reverses the burden. Once the plaintiff shows a denial of meaningful access or a screen-out, the covered entity bears the burden of establishing that the standard is job-related and consistent with business necessity, or that the accommodation

29. See *Sch. Bd. of Nassau Cnty. v. Arline*, 480 U.S. 273, 287 (1987) (individualized inquiry required under Section 504); *PGA Tour, Inc. v. Martin*, 532 U.S. 661, 688–90 (2001) (modification analysis under the ADA requires an individualized assessment).

30. *Fairfield*, *supra* note 1, at 580–83.

would impose an undue hardship or work a fundamental alteration.³¹ The entity must produce and prove the justification. The plaintiff is not put to the task of disproving a story the system generated. A system optimized to generate plausible justifications has handed the justification to the party that already had to supply it.

C. The Knowledge Wrinkle and the Strongest Vehicle

One limit must be stated plainly, because it determines which factual setting carries the argument. The accommodation duty is triggered by a known disability. The covered entity must know, or have reason to know, of the disability and the need for accommodation before the duty to accommodate arises. In algorithmic hiring, the employer ordinarily does not know which applicants are disabled, because the system screens on proxies and is never told. That absence of knowledge cuts against the undisclosed applicant in two directions at once: it defeats the disparate treatment theory, because one cannot intend to exclude a disability one never perceived, and it defeats the accommodation theory, because the duty was never triggered. The undisclosed applicant is therefore channeled back into disparate impact, the one engine that inherits the detection problem the original article identifies.

The implication is not that the accommodation engine is weak. It is that the engine is strongest where the disability is known and the duty is live. Two settings present that posture cleanly. The first is the existing employee who has disclosed a disability and requested an accommodation, and whose request is routed through an automated system that denies it without an explanation a human can verify. The second, and the cleanest of all, is the registered student who is enrolled with a university's disability services office under Section 504, whose status is documented, whose need is on file, and whose accommodation or eligibility decision is made by a system no administrator can explain. In that setting there is no detection problem to solve. The disability is known, the duty is affirmative, no comparator is required, and the opacity of the system, as the next Part shows, runs against the entity rather than for it.

31. See 42 U.S.C. §12113(a); *Choate*, 469 U.S. at 301.

IV. The Deployers Anti-Discrimination Scholarship Overlooks

The existing literature locates the danger in employers, human-resources departments, police departments, hospitals, and benefits agencies.³² Education appears once, as two words in a list of affected domains, and is not developed.³³ The omission matters, because the institutions now adopting automated decision systems for administrative functions include the cleanest class of disability-law defendants in the United States.

A. Universities as Recipients

Nearly every college and university in the country receives federal financial assistance, whether through Title IV student aid, federal research funding, or both. That fact places the institution squarely within Section 504, which reaches any program or activity receiving such assistance.³⁴ And the coverage is not confined to the particular office that handles the federal money. After the Supreme Court read the phrase “program or activity” narrowly,³⁵ Congress overrode that reading in the Civil Rights Restoration Act of 1987 and defined “program or activity” to mean all of the operations of the covered entity.³⁶ An institution that takes federal funds anywhere is covered everywhere, including in the admissions office and the disability services office. There is no dodge available in which the unit deploying the screening system is said to fall outside the funded program.

The coverage compounds by institutional type. A public university is also a public entity subject to Title II of the ADA, which carries its own nondiscrimination and program-access obligations.³⁷ A private university is a place of public accommodation under Title III, which

32. See *id.* at 539–40, 592–93 (cataloguing domains and discussing automated welfare-eligibility systems in the United Kingdom and in Telangana, India).

33. *Id.* at 539 (listing “educational evaluations” among the decisions optimized to target the vulnerable).

34. 29 U.S.C. §794(a).

35. See *Grove City Coll. v. Bell*, 465 U.S. 555, 570–74 (1984) (limiting coverage to the specific program receiving funds).

36. See 29 U.S.C. §794(b) (defining “program or activity” for Section 504 to include all operations of a covered college, university, or public system); Civil Rights Restoration Act of 1987, Pub. L. No. 100-259, 102 Stat. 28.

37. 42 U.S.C. §§12131–12132.

enumerates undergraduate and postgraduate schools among the covered categories.³⁸ The Section 504 obligation, however, is the common denominator and the cleanest hook, because the funding nexus is uncontested and the institution-wide reach forecloses the program-carve-out argument before it can be made.

B. The Decisions Now Being Automated

The functions universities are beginning to route through automated systems are exactly the functions Section 504 governs. The applicable regulations prohibit disability discrimination in admissions and recruitment by postsecondary institutions, and require academic adjustments and auxiliary aids so that a disabled student is not excluded from or denied the benefits of the institution's programs.³⁹ These are not hypothetical deployments. Universities now route admissions review through automated systems that score and summarize the very materials decisions rest on: one major public university has used automated essay scoring since 2019, another reviews applicant transcripts with artificial intelligence, and a third adopted, beginning with the 2025–26 cycle, a hybrid model that pairs a single human reader with an artificial-intelligence scorer and dispenses with a second human reader whenever the machine and the human agree.⁴⁰ Institutions describe these tools as assisting rather than deciding, but a system that scores and ranks the materials on which advancement turns is a selection device in the sense the law cares about, and the human who reviews the machine's summary, as Part V.D below explains, may be reviewing a judgment no one can verify. The same automation

38. 42 U.S.C. §12181(7)(J).

39. The Department of Education's Section 504 regulations governing postsecondary education are codified at 34 C.F.R. pt. 104, including provisions on admissions and recruitment and on academic adjustments and auxiliary aids. The precise section designations within Part 104 should be confirmed against the current Code of Federal Regulations, as the Part has been the subject of recent rulemaking activity.

40. See *Colleges Are Using AI Tools to Streamline Admissions Tasks*, Tucson.com (Jan. 20, 2026), https://tucson.com/life-entertainment/nation-world/article_aaddf6ec-3212-568f-a193-cc5878d24a18.html (reporting that the University of North Carolina at Chapel Hill has used automated essay scoring since 2019, that Stony Brook University uses artificial intelligence to review applicants' transcripts, and that other institutions are incorporating such tools into application screening); see also *Which Colleges Use AI to Read Essays*, GradPilot (May 13, 2026), <https://gradpilot.com/news/which-colleges-use-ai-2025> (describing Virginia Tech's adoption, beginning with the 2025–26 cycle, of a hybrid model pairing one human reader with an artificial-intelligence scorer). Permanent archival links and confirmation of authorship for each source should be added on the citation pass.

has reached disability decisions directly. The Department of Education’s Office for Civil Rights has warned that the use of artificial intelligence in education can violate Section 504 and Title II, and among the agency’s own examples is a generative tool that composes Section 504 accommodation plans for disabled students without human oversight, producing plans that do not meet individual needs.⁴¹ Automated examination proctoring routinely flags the movements, speech, and behaviors of disabled students as suspicious, because those behaviors are features of disability the software was never built to accommodate,⁴² and at least one student has sued a university alleging that its reliance on artificial-intelligence essay detection produced a disability-discriminatory result and a failure to accommodate a documented condition.⁴³ An admissions screen that disadvantages disabled applicants, an automated triage of accommodation requests, and an eligibility or financial-aid determination that no administrator can explain are therefore decisions the statute already reaches, made by tools whose reasons no one can give. And admissions is the setting in which the concealment the original article describes operates most harshly, because the applicant never sees the decision logic, has no continuing relationship through which to contest it, and has the least standing and the least information of any affected person.⁴⁴

41. See U.S. Dep’t of Educ., Off. for Civ. Rts., guidance on avoiding the discriminatory use of artificial intelligence (2024) (offering as examples a generative artificial-intelligence tool that composes Section 504 plans for disabled students without human oversight, and an artificial-intelligence classroom monitor that repeatedly flags a hard-of-hearing student); *ED Issues Guidance to Avoid Discriminatory Use of AI*, GovTech (Nov. 26, 2024), <https://www.govtech.com/education/k-12/ed-issues-guidance-to-avoid-discriminatory-use-of-ai>. The exact title and date of the Office for Civil Rights document should be confirmed on the citation pass. Section 504 governs both elementary-secondary and postsecondary recipients; the examples are drawn from the agency’s guidance to recipients generally.

42. See Ctr. for Democracy & Tech., *How Automated Test Proctoring Software Discriminates Against Disabled Students* (2021), <https://cdt.org/insights/how-automated-test-proctoring-software-discriminates-against-disabled-students/>.

43. The suit, brought by a plaintiff proceeding as Jane Doe against the University of Michigan, its Board of Regents, and named instructors and administrators in the United States District Court for the Eastern District of Michigan, alleges that the university treated writing traits associated with her generalized anxiety disorder and obsessive-compulsive disorder as evidence of artificial-intelligence use and failed to accommodate her documented conditions during the academic-integrity process. See *Student Sues University of Michigan Over AI Misconduct Accusation*, GovTech (Feb. 12, 2026), <https://www.govtech.com/education/higher-ed/student-sues-university-of-michigan-over-ai-misconduct-accusation>. The complaint and docket should be obtained and cited directly on the citation pass; the citation here is to secondary coverage only.

44. Cf. Fairfield, *supra* note 1, at 580–83 (intersectional plaintiffs are least likely to have the resources to contest, and most likely to be exploited by, automated decisions).

The exclusion is not confined to federally funded universities. The ADA independently reaches private certification providers: any entity that offers examinations related to professional or trade certification must offer them in a manner accessible to persons with disabilities or provide accessible alternative arrangements.⁴⁵ The Silicon Valley Certification Hub, which sells credentials it brands Chief AI Officer and Chief AI Ethics and Responsibility Officer, illustrates the problem in its own field. Its published examination policy requires the candidate to remain under continuous webcam surveillance, alone and without interruption, for sessions that run up to 180 minutes; it forbids standing and taking breaks; and it refers any other behavior it deems suspicious to an ethics committee, with no accommodation pathway stated on the policy page.⁴⁶ A candidate whose disability requires movement or breaks, or whose involuntary motion would be read as suspicious, is screened out by the policy's own terms, by an entity that certifies others in the ethics of artificial intelligence.

C. Public Agencies and the State as Deployer

The same logic extends to government agencies that deploy automated decision systems and receive federal funds. Where a state agency administers a federally funded program, Section 504 attaches across the department, and the agency is also a public entity under Title II.⁴⁷ The benefits-eligibility systems the original article describes, which removed entitled recipients without explanation or recourse,⁴⁸ are not only the disparate impact problem that article analyzes; where the affected recipients are disabled and the denial defeats the meaningful access the statute guarantees, they are an accommodation and access problem governed by Section 504, with the

45. 42 U.S.C. §12189 (requiring any person that offers examinations or courses related to applications, licensing, certification, or credentialing for secondary or postsecondary education, professional, or trade purposes to offer them in a place and manner accessible to persons with disabilities, or to offer alternative accessible arrangements). Section 12189 has been applied to professional and admission examinations such as the Law School Admission Test and the bar examination.

46. Silicon Valley Certification Hub, *Certification Exam*, <https://svch.io/certification-exam/> (last visited June 20, 2026) (stating that all examinations are supervised by webcam, that the candidate must be alone and without interruptions, and that standing, taking breaks, and any other suspicious activity are not allowed, with any incident reviewed by an ethics committee).

47. See 29 U.S.C. §794(a)–(b); 42 U.S.C. §§12131–12132.

48. Fairfield, *supra* note 1, at 592–93.

agency, not a vendor, as the responsible party. This pattern is neither foreign nor hypothetical, and it has already fallen on disabled people in the United States. When Arkansas replaced an individualized nursing assessment with an algorithm to set Medicaid in-home care hours, beneficiaries with conditions such as cerebral palsy and quadriplegia had their hours cut by as much as half though their medical conditions had not changed, and neither the agency nor the recipients could obtain a usable account of why; a federal court found the state had failed to provide adequate notice, and the tool was ultimately invalidated.⁴⁹ Idaho cut services to adults with developmental disabilities through a comparable algorithm that the state defended as a proprietary black box, until the Ninth Circuit allowed a due-process challenge to proceed.⁵⁰ Each state was a recipient administering a federally funded program, denying disabled people the meaningful access Section 504 guarantees, through a system no administrator could explain. The funded public deployer is, for the reasons developed below, the defendant least able to hide behind the opacity of the system it chose.

V. The Box That Indicts

The decisive move is to see that opacity changes valence across the axis. On the intent-versus-impact axis, an unexplainable model is an asset to the defendant, because it conceals intent and frustrates the comparison that proves impact. In the accommodation and qualification-standard setting, the same opacity is a liability, because the law places on the deployer an affirmative burden that an unexplainable system cannot discharge.

49. See Ctr. for Democracy & Tech., *Challenging the Use of Algorithm-driven Decision-making in Benefits Determinations Affecting People with Disabilities* (2020), <https://cdt.org/wp-content/uploads/2020/10/2020-10-21-Challenging-the-Use-of-Algorithm-driven-Decision-making-in-Benefits-Determinations-Affecting-People-with-Disabilities.pdf>; *Suit Filed Over Computer Program Making Medicaid Cuts*, Ark. Access to Just. (Jan. 27, 2017), <https://arkansasjustice.org/2017/02/27/suit-filed-over-computer-program-making-medicaid-cuts/>. The Arkansas state-court decision invalidating the assessment tool should be located and cited directly on the citation pass.

50. *K.W. ex rel. D.W. v. Armstrong*, 789 F.3d 962 (9th Cir. 2015).

A. The Validation Requirement

A selection device that produces a disparate impact is not lawful merely because the employer can recite a plausible purpose for it. The device must be shown to be job-related, and the burden of that showing rests on the employer. *Griggs* held that a practice with a disparate impact is permissible only if it bears a manifest relationship to the employment in question.⁵¹ *Albemarle Paper Co. v. Moody* operationalized that requirement as validation: the employer must demonstrate, through professionally accepted methods, that the device measures or predicts important elements of the work for which it screens.⁵² The Court extended impact analysis to subjective and discretionary selection methods in *Watson v. Fort Worth Bank & Trust*, refusing to let the line between objective tests and subjective judgments become a refuge from scrutiny.⁵³ Although *Wards Cove Packing Co. v. Atonio* briefly relaxed the employer's burden,⁵⁴ Congress restored it in the Civil Rights Act of 1991, which places on the employer the burden of demonstrating that a challenged practice is job-related and consistent with business necessity, and preserves the plaintiff's ability to prevail by identifying an equally valid, less discriminatory alternative the employer refuses to adopt.⁵⁵

The ADA carries the identical requirement into the disability context and sharpens it. A qualification standard or test that screens out a disabled individual must be job-related and consistent with business necessity,⁵⁶ a test must be selected and administered to measure aptitude rather than the disability itself,⁵⁷ and even a validated standard must yield where a reasonable accommodation would allow the individual to satisfy it.⁵⁸ The validation standards themselves

51. *Griggs*, 401 U.S. at 431–32.

52. *Albemarle*, 422 U.S. at 425, 431–36.

53. *Watson v. Fort Worth Bank & Trust*, 487 U.S. 977, 989–91 (1988) (plurality opinion).

54. *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642, 658–61 (1989).

55. 42 U.S.C. §2000e-2(k)(1)(A).

56. 42 U.S.C. §§12112(b)(6), 12113(a).

57. 42 U.S.C. §12112(b)(7).

58. See 42 U.S.C. §12112(b)(5)(A); *US Airways, Inc. v. Barnett*, 535 U.S. 391, 397–98 (2002) (the accommodation obligation can require an exception to a neutral rule).

are supplied by the agencies, whose guidance on selection procedures sets out the recognized methods of demonstrating that a device is job-related.⁵⁹

B. An Unexplainable Device Cannot Be Validated by Explanation

The two classical paths to validating a selection device, content validity and construct validity, both require the validator to say what the device measures and to connect that to the demands of the job. A content-validity showing demonstrates that the device samples the actual content of the work; a construct-validity showing demonstrates that the device measures an identified trait that the work requires. Each presupposes that the deployer can articulate what the device is doing. A model whose decision logic its deployer cannot explain cannot be validated by either method, because the deployer cannot state, let alone prove, the relationship between what the model attends to and what the job requires. The result follows directly: a disparate impact, plus a selection device the deployer cannot validate, plus a burden of validation that rests on the deployer, equals a business-necessity defense that fails as a matter of law. The opacity that conceals intent on the other axis is, on this axis, the very thing that sinks the defense.

C. The Empirical Escape and Its Three Leaks

There is one path to validation that does not require explaining the model. Criterion-related validity is an empirical showing: one demonstrates that the device's scores correlate with a measure of job performance, without opening the device to see how it reasons. An employer may therefore argue that it need not explain its model at all, so long as it can show statistically that the model's outputs predict success on the job. The argument is real and must be answered rather than ignored. It leaks in three places.

First, the showing is an affirmative burden that most deployers have not carried. Running a criterion-related validation study is expensive and uncommon, and the absence of one is not

59. See Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. pt. 1607; Equal Employment Opportunity Commission, Regulations to Implement the Equal Employment Provisions of the Americans with Disabilities Act, 29 C.F.R. pt. 1630. The current status of the Uniform Guidelines should be confirmed, as their revision has been under discussion.

cured by the employer's confidence that the model works. Where no study exists, the burden is simply unmet.

Second, a study that does exist and that reveals a disparate impact places the employer on notice of that impact and reopens the less-discriminatory-alternative inquiry. An employer that has measured the impact and proceeded anyway cannot claim the innocence of ignorance, and the statute permits the plaintiff to prevail by showing an equally valid alternative the employer declined to adopt.⁶⁰

Third, and most fundamentally, a criterion-related study validates the model against a criterion, and the usual criterion is the past performance or past success of prior hires. Where prior hiring was itself shaped by the discrimination the system is now accused of reproducing, the validation is built on a contaminated baseline. This is precisely the point the original article presses about training data: a model trained on, or validated against, historically discriminatory outcomes learns and launders the discrimination embedded in them.⁶¹ The empirical escape leaks exactly where that article's thesis applies, and the employer who reaches for it hands the plaintiff the original article's own diagnosis as the answer.

D. The Faithfulness and Tooling Pincer

A deployer pressed on validation will reach for one more refuge: the model's own account of its decision. Modern systems can be made to produce a stated rationale, and a rejection that recites a reason, low scores on some articulated factor, looks like an explanation that proves the decision was not about disability. That refuge is closed, and it is closed by the research of the developers themselves.

A leading developer of these systems has published findings that the stated reasoning a model produces frequently fails to reflect the factors that actually drove its output, and in controlled settings diverges from them: a model can be induced to rely on a cue, use it, and

60. 42 U.S.C. §2000e-2(k)(1)(A)(ii).

61. See Fairfield, *supra* note 1, at 555–59, 568–73 (a system trained or validated on historical outcomes reproduces the discrimination encoded in them).

never mention it in the reasoning it displays, with measured faithfulness low and lower on harder problems.⁶² Companion interpretability work has shown, at the level of the model's internal computation, cases in which the process the model actually runs differs from the process it describes when asked.⁶³ The consequence for the law is direct. A deployer cannot satisfy its burden by pointing to what the model said about its own decision, because the entity that built the model has published that what the model says about its own decision is not a reliable account of why the decision was made. The model's self-explanation is not evidence; it is, in effect, an admission against interest furnished by the industry that the opacity is genuine.

The same body of work closes the opposite escape as well, and the point must be made or the argument is vulnerable to it. The very laboratory that documented the unfaithfulness of stated reasoning also released tooling capable of analyzing a model's internal computation, the actual basis of its decisions, rather than its surface narrative.⁶⁴ A deployer cannot, therefore, retreat to the position that the system is simply an unknowable black box whose reasons no method could ever recover. The means to examine model reasoning exists and was published.

These two facts are not in tension; they are a single pincer, and they close on the deployer from both sides. If the model can be examined, the deployer should have examined it before screening disabled people with it, and is properly charged with what such an examination would have revealed. This is the imputed-knowledge move, the principle that an entity is charged with what it would have learned had it not chosen not to look, and it is a move the original article itself endorses.⁶⁵ If instead the model genuinely cannot be examined, because the deployer procured

62. The reference is to published research by a leading artificial-intelligence laboratory on the faithfulness of model chain-of-thought reasoning, finding that visible reasoning often does not reflect the actual basis of a model's output. The exact title, authoring entity, date, and a stable archival link should be confirmed and supplied on the citation pass; the finding is summarized here only at the level of generality the argument requires.

63. The reference is to published interpretability research by the same laboratory analyzing the internal computation of a large language model and documenting divergence between the computation performed and the explanation the model gives. The exact title, authoring entity, date, and archival link should be confirmed on the citation pass.

64. The interpretability research cited *supra* was accompanied by the release of analysis tooling for examining model internals. Whether and how that tooling can be applied to any particular deployed model, including closed or third-party-hosted models that a deployer cannot inspect, is a separate question that the citation pass and the body should address with care.

65. Fairfield, *supra* note 1, at 582–83, 610–11 (an entity that deploys these systems should be charged with what it could have learned had it chosen to train or examine them for the affected population's benefit).

a closed system it has no access to inspect, then the deployer has chosen to screen people with a selection device it structurally cannot validate, and the burden of validation is still its own. One does not escape the validation requirement by selecting an unvalidatable device and then pleading that validation was impossible by one's own procurement choice. Either branch ends in liability. The opacity that looked like a shield is a box with no exit.

E. Cost Is Not Business Necessity

A final defense remains to be foreclosed, and it is the one the original article's own thesis defeats. Pressed to justify the screen, a deployer will say that the automated system is necessary because the volume of applications or requests cannot be handled by human reviewers, and that human review is therefore not an equally effective, less discriminatory alternative. Two answers meet this, and they converge.

The first is doctrinal. The only thing the automated system adds over an individualized human review is cost and scale. It is faster and cheaper; it is not, on the disability axis, more accurate at the thing the law requires, which is an individualized assessment of whether this person, with a reasonable accommodation, can meet a genuine requirement of the position or program. Cost reduction is not business necessity. An employer may not justify a discriminatory screen on the ground that the nondiscriminatory alternative is more expensive, and individualized human review is both the less discriminatory alternative and, independently, what the ADA's individualized-assessment requirement already demands.⁶⁶ The deployer's refusal to provide the human review it could provide, in favor of an unexplainable system it chose for its cheapness, is the actionable choice.

The second answer is supplied by the original article itself. That article's central empirical claim is that the hidden optimization goal driving these systems is the reduction of cost, and that this cost-reduction goal is the engine that produces and conceals the discrimination.⁶⁷ If that is right, then the only "necessity" a deployer can name for its automated screen is the very

66. See *Arline*, 480 U.S. at 287; *Martin*, 532 U.S. at 688–90.

67. *Fairfield*, *supra* note 1, at 542, 592–94.

cost-reduction goal the article identifies as the source of the harm, and which the law has long refused to recognize as a business necessity. The original article names the disease. The disability axis supplies the remedy.

F. Human Review That Launders Rather Than Pierces

One proposed safeguard deserves separate treatment, because it is weaker than it appears and its weakness follows from the same research. The literature, including the article engaged here, endorses a human-in-the-loop requirement: a human who reviews the automated decision provides explainability, supplies a person who can be deposed, and reconnects the decision to a responsible entity.⁶⁸ In the disability setting, a human who reviews only the model's stated rationale does not pierce the opacity; the human launders it. Deposed, that reviewer will testify truthfully that the system reported a particular reason, and that reported reason is, on the developers' own findings, not a reliable account of the actual basis of the decision. A human reviewing a fluent but unfaithful rationale cannot verify that the decision was not driven by disability, which is the one thing the ADA's individualized-assessment duty requires the reviewer to be able to do.⁶⁹ The human-in-the-loop is a genuine improvement against the displacement of responsibility, but as a cure for the validation and access problems it can become a ceremony that certifies a decision no one has actually examined.

VI. Why Injunctive Relief Is the Feature

The existing literature treats the remedial structure of disparate impact as a weakness. Because the remedy is typically an injunction and punitive damages are unavailable for unintentional discrimination, a liability-minimizing system reads injunction-only exposure as a lower cost and steers toward the conduct that produces it.⁷⁰ On the deterrence logic of that account, the point

68. *Id.* at 597–99.

69. See *Arline*, 480 U.S. at 287 (individualized inquiry).

70. *Id.* at 556; 42 U.S.C. §1981a(a)(1).

is sound. On the disability axis, the same remedial fact is the decisive strength rather than the weakness.

The intent-free theories developed in this Article, failure to accommodate and denial of meaningful access, are the equitable theories. They do not require the showing of intentional discrimination, and indeed the absence of intent is irrelevant to them. Money damages under Section 504 and Title II run on a different track and generally require intentional discrimination, which courts have read to mean at least deliberate indifference,⁷¹ a showing that an unexplainable and unintentional system makes hard to mount. That is exactly where the original article's pessimism bites hardest. But the injunctive remedy does not require it. A court can order a covered entity to stop using a screening device it cannot validate, to provide individualized human review, and to afford the accommodation the statute guarantees, without any finding of intent at all.

This is why the enforcement story belongs on the equitable side. Injunctive relief forces the design to change. A cost-minimizing deployer can absorb a damages judgment as a cost of doing business and continue to run the system; it cannot continue to run a system a court has enjoined. The remedy that the deterrence account treats as the system's escape hatch, injunction without damages, is the remedy that closes the escape, because it reaches the design rather than the balance sheet. The original article and this one are therefore not in conflict. That article mourns the damages-and-deterrence track, on which the unexplainable system does indeed read injunction-only exposure as cheap. This Article builds on the accommodation-and-injunction track, on which the same injunctive remedy is the instrument that forces the unexplainable system off the field.

71. See *Barnes v. Gorman*, 536 U.S. 181, 184–89 (2002) (punitive damages unavailable under Section 504 and ADA Title II; compensatory damages available on a showing tied to intentional discrimination). The precise damages standard, including the deliberate-indifference formulation, varies by circuit and should be stated to the governing circuit on the citation pass.

VII. Conclusion

The scholarship on algorithmic discrimination has been mapping the wrong axis for the disability problem. Intent versus impact is the axis on which artificial intelligence wins, because the systems are built to hide intent, manufacture justification, and eliminate the comparisons that prove effect. The accommodation axis is the axis on which artificial intelligence loses, because the duty to provide meaningful access has no intent element to hide, needs no comparator to eliminate, and places the burden of justification on the deployer rather than the plaintiff. And on that axis the opacity of the system, which is an asset to a Title VII defendant, is a liability to a disability defendant, because a selection device the deployer cannot validate fails the business-necessity requirement, and the developers of these systems have published that the deployer cannot cure the failure by pointing to the system's account of itself.

The institutions most exposed are the ones the literature has overlooked. Universities and funded public agencies are paradigmatic recipients of federal financial assistance, covered institution-wide, and they are now routing admissions, accommodation, and eligibility decisions through systems no administrator can explain. The registered disabled student denied an accommodation by an unexplainable model is not a hard case at the frontier of the law. She is a straightforward Section 504 plaintiff whose claim does not depend on solving the detection problem that traps everyone else, because her disability is known and the duty she invokes is affirmative.

The original article ends by asking what comes next, now that the guardrails fail. The answer this Article offers is not a better guardrail and not a new statute. It is the body of law that already exists, applied to the deployers it already reaches: intent-agnostic, design-forcing accommodation and access liability, enforced by injunction, against entities that cannot validate the boxes they have chosen to put between a disabled person and the opportunity the law guarantees her.