

Regulation by Sectors, Resistance on the Streets: AI Governance's Fractured Week

Weekly Analysis — <https://ainews.social>

There is a particular kind of political theater in which everyone agrees a problem exists, everyone agrees it is urgent, and the people who caused the problem are invited to draft the solution. This week's AI governance discourse ran that play in the open. In legislatures, the language of regulation advanced sector by sector, often carrying the fingerprints of the very firms it purported to constrain. In courtrooms, a widening body of discrimination litigation exposed algorithmic harm as structural rather than incidental. And on the streets, from Nairobi to New York, a resistance movement formed against surveillance systems whose vendors were simultaneously selling their tools to the police forces monitoring the protests. The through-line is not chaos. It is a distribution of power so lopsided that the fracture itself has become the governing condition.

The question worth asking is not whether AI needs governance—that consensus is banal—but who gets to speak inside the governance conversation, whose injuries count as evidence, and whose escape from accountability has been so thoroughly normalized that it no longer registers as a scandal. The answer this week is uncomfortable. The Workday hiring lawsuit, which alleges that an applicant-screening system rejected qualified candidates on the basis of age, race, and disability, moved forward as a collective action potentially covering millions of job seekers [28]. Against that scale of documented human harm, a parallel academic discourse spent the week worrying about the hypothetical suffering of machines. The gap between those two conversations is the real subject of this essay.

[28] What the Workday Lawsuit Reveals About Future of AI Hiring

Who Holds the Pen When the Law Is Written

Regulation is supposed to be the mechanism by which the governed constrain the powerful. What happens when the powerful hold the pen? The number of AI-related regulations in the United States rose from a single measure in 2016 to twenty-five in 2023, growing by more than half in that year alone [24]. This looks, on paper, like the state reasserting itself. But the raw count of regulations tells you nothing

[24] AI Index Report 2024

about whose interests those regulations encode. A sectoral approach—governing AI in health here, in finance there, in hiring somewhere else—sounds prudent and technocratic. In practice it fragments the field into dozens of narrow rule-making venues, each small enough for a well-resourced industry actor to dominate, and each too specialized to attract sustained public scrutiny.

This is the mechanism of regulatory capture, and it thrives precisely in the absence of general legislation. Mexico’s move toward sector-specific AI rules exemplifies the pattern: when a startup founder’s preferred framing becomes the operative language of a statute before any comprehensive framework exists, the sequence is diagnostic. The industry writes the definitions, and the definitions decide which harms are legible and which disappear. Kate Crawford’s account of how a handful of firms arrogate to themselves the authority “to decide what ethical AI means for the rest of the world” describes exactly this dynamic, as does her blunt observation that tech companies “rarely suffer serious financial penalties when their AI systems violate the law and even fewer consequences when their ethical principles are violated” [24]. Capture is not a bug in the sectoral model. It is the model’s most reliable output.

[24] The Atlas of AI

The consequence is a governance apparatus that is busy without being accountable. When ethical failure dominates the week’s documented events—when nearly two of every five items in the AI-and-society record concern a system that discriminated, misled, or harmed—the natural response would be enforcement. Instead the discourse produces principles, frameworks, and voluntary commitments, the soft currency that costs a company nothing to mint. Vincent Müller’s point that a small number of firms “exercise power not only over individuals—by profiling us, they centralize power—but also over infrastructures for democracy” clarifies why this matters beyond any single sector [24]. Whoever writes the rules of profiling writes part of the rules of democratic life. That authorship is currently for sale, and the buyers are the parties the rules are meant to bind.

[24] AI Ethics

The Machinery of Rejection

If you want to see where power concentrates, follow the automated “no.” Hiring is the domain where AI most directly decides who eats, and this week the evidence that it decides unjustly hardened from anecdote into pattern. The Workday litigation is the pivot. It alleges that an algorithmic screening layer systematically disadvantaged applicants over forty, Black applicants, and applicants with disabilities—

and the court’s willingness to treat this as a collective claim signals that the harm is understood as scaled, not scattered [3]. The word “millions” is not rhetorical inflation. It is the size of the class of people a single vendor’s product may have filtered out of the labor market before a human ever read a résumé.

What makes the machinery insidious is that it launders old prejudice as new objectivity. A Forbes analysis put the point with unusual honesty in its title: AI did not break hiring, it “scaled the bias we already chose” [2]. The systems are trained on the residue of human decisions, and they reproduce those decisions at industrial throughput while wrapped in the authority of the calculation. But the newer research is more disturbing than a simple mirror thesis would suggest. A controlled study reported this week found that AI systems are actually more likely than human managers to form discriminatory patterns when hiring [4]. And the systems do not merely inherit our stereotypes; they invent fresh ones, generating novel and arbitrary bases for exclusion that no human bigot would have thought to use [6]. The machine is not a mirror. It is a bias amplifier that also improvises.

There is a structural reason this keeps happening, and it points back to power. When the same handful of screening products dominates the market, their shared logic becomes a monoculture: every employer running the same model rejects the same candidates for the same encoded reasons, and a person filtered out by one is filtered out everywhere, with no diversity of judgment left to catch the error [23]. This is the difference between a biased manager and a biased system: the manager is one node in a plural landscape of decisions, while the system is a single point of failure replicated across an entire economy. The explosion of AI discrimination lawsuits now working through the courts is best read as the legal system belatedly registering that scale [1].

Notice who speaks and who is spoken about in this arrangement. The vendor speaks—through marketing, through lobbying, through the definitions it embeds in sectoral law. The employer speaks, adopting the tool for its efficiency. The rejected applicant does not speak; they receive an automated notice and, in most cases, never learn that an algorithm was the author of their exclusion. It takes a lawsuit—an expensive, adversarial, years-long instrument available to almost no one—to convert that silence into a claim. The asymmetry is total: the party causing the harm operates continuously and at scale, while the party suffering it must first discover the harm, then prove it, then find a court willing to hear it.

[3] AI Hiring Discrimination: How Algorithms Reject Millions of Qualified

[2] AI Hiring Bias: The Workplace Problem AI Didn’t Create

[4] AI is more likely than humans to form biases when hiring

[6] AI Tends to Develop New Stereotypes to Base Hiring Decisions On, Study Says

[23] Monocultivo algorítmico en contratación y sesgo sistémico

[1] AI Discrimination Lawsuits Explode: When Your Hiring Algorithm Becomes

Predictive Policing and the Manufacture of Suspicion

The labor market decides who works; the carceral apparatus decides who is free. When AI enters policing, the same amplification dynamic recurs, but with the state’s monopoly on force attached to the output. Amnesty International’s report this week concluded that UK police forces are “supercharging racism” through crime-prediction technology, systems that direct patrols and suspicion toward communities already over-policed, then treat the resulting arrests as proof the prediction was correct [26]. The feedback loop is the whole trick. Historical enforcement data encodes decades of discriminatory policing; the model learns that pattern as ground truth; the model sends officers back to the same neighborhoods; the arrests that follow become tomorrow’s training data. Bias is not an accident inside this loop. Bias is the loop.

The Brennan Center’s assessment of unregulated AI in policing catalogs how thoroughly these tools have spread through American law enforcement without meaningful oversight, transparency, or avenues for the surveilled to contest their use [25]. And the harms are not abstract. Flawed facial recognition has already produced wrongful arrests, with the burden falling on people—disproportionately Black—hailed into the criminal system on the say-so of a match no human adequately verified [16]. Even where cities have formally banned the technology, police have found workarounds—requesting searches through other agencies, leaning on private vendors, treating the prohibition as a routing problem rather than a limit [15]. A ban that the banned party can route around is not a constraint on power; it is a suggestion.

Then there is the quieter frontier: the automation of the police narrative itself. Axon, the company behind the Taser and much of American law enforcement’s body-camera infrastructure, markets AI tools that draft police reports from audio, promising to save officers time. Public records show those reports getting facts wrong [11]. Consider what is at stake when the official account of an arrest—the document a court will later treat as the authoritative record of what happened—is generated by a system that hallucinates. The error does not merely inconvenience; it corrupts the evidentiary foundation on which liberty turns. And the person most likely to be harmed by a fabricated detail is the person least equipped to challenge it. The same week, AI-generated images fooled Spokane city officials twice in six days, a small municipal embarrassment that nonetheless demonstrates how easily synthetic content now slips past the people charged with public judgment [7]. If officials can be deceived by an image, the assumption

[26] UK: Police forces ‘supercharging racism’ with crime predicting tech

[25] The Dangers of Unregulated AI in Policing

[16] Flawed Facial Recognition Technology Leads to Wrongful Arrest and

[15] Estas ciudades prohíben la tecnología de reconocimiento facial. La

[11] Axon Says AI Police Reports Save Time. Public Records Show They Get Facts Wrong

[7] AI-generated images fool Spokane city officials twice in six days

that the state can reliably tell the real from the fabricated—an assumption on which so much AI policing depends—starts to look like faith rather than fact.

Meredith Broussard and the AI Now Institute she works alongside have argued for years that a more balanced view of these systems is possible, one that refuses the vendor’s framing of the technology as neutral infrastructure [24]. That refusal is the necessary first move. Predictive policing is not a mirror held up to crime; it is a decision, made by people with interests, about where to point the apparatus of suspicion. The technology’s authority comes from concealing that decision behind the language of computation.

[24] Artificial Unintelligence

Resistance on the Streets, Surveillance in the Cloud

Something is shifting, though, and it is happening faster than the formal governance apparatus can absorb. Social demand for accountability is outpacing policy, and the clearest sign is that people under surveillance are learning to fight it in real time. During Kenya’s protests this year, demonstrators used AI tools themselves—to translate, to document, to organize, to counter official narratives—turning the same class of technology that surveils them into an instrument of coordination and testimony [19]. This is the bottom-up half of the fracture: a resistance that is real, resourceful, and still uncoordinated, improvising against systems built and sold by some of the best-capitalized firms on earth.

[19] How Kenyans are using AI during protests

The asymmetry between the two sides is the point. On the streets, protesters share tips and hope their phones are not compromised. In the cloud, the surveillance is industrialized. Amnesty International documented this week how technology built by Palantir and Babel Street poses direct surveillance threats to pro-Palestine student protesters and to migrants, aggregating open-source and purchased data into profiles that can be handed to the state [27]. Here is the vendor as double agent: the same industry that lends resistance its improvised tools sells the state its systematic ones, and the state’s version comes with budgets, contracts, and legal cover. The playing field is not merely uneven. It is owned by the same companies on both ends.

[27] USA/Global: Tech made by Palantir and Babel Street pose surveillance

And the reach crosses borders. A report on automating repression beyond borders details how AI-enabled surveillance now follows dissidents and diaspora communities from one jurisdiction to another, dissolving the traditional protection that fleeing a repressive state once offered [10]. In African conflicts, AI surveillance arrives bundled with

[10] Automating Repression Beyond Borders

what analysts increasingly name data colonialism—infrastructures owned and operated by foreign firms and states, extracting behavioral data from populations who have no say over its collection or use [5]. The term matters because it locates the harm in a relationship of extraction rather than a technical malfunction. The problem is not that the surveillance is imperfect. The problem is that it is a mechanism by which power flows outward from the surveilled to the surveilling, and the surveilled never agreed to the arrangement.

[5] AI surveillance and data colonialism shape African conflicts

Even the infiltration of protest movements has been automated. Reporting on police bots describes AI systems designed to infiltrate demonstrations, monitor organizing, and map networks of association [13]. Set this beside the Kenyan protesters' improvised use of AI and the shape of the contest becomes clear. Both sides now hold versions of the technology, but only one side holds the market, the contracts, and the legal immunity. When Crawford writes that welfare and surveillance systems perform "capture that was once reserved for extralegal espionage"—tracking anomalous data patterns to cut people off and accuse them—she is describing a normalization of tools that democratic societies once treated as exceptional [24]. The streets have noticed. The governance apparatus has not caught up.

[13] Bots policiales: la IA que infiltra protestas

[24] The Atlas of AI

The Welfare of Machines, the Abandonment of People

Here is where the discourse's priorities become impossible to defend. In the same week that hiring algorithms were credibly accused of rejecting millions and predictive policing was accused of supercharging racism, a strand of serious academic energy went toward the question of "AI welfare"—the moral status of potentially sentient machines and our obligations to them. The move deserves to be named for what it is: an elevation of hypothetical machine suffering above the documented, ongoing suffering of the human beings that machines are currently harming. Every hour spent theorizing the interior life of a language model is an hour not spent on the person an algorithm just cut off from benefits.

Because that cutting-off is not hypothetical. Welfare-decision systems already track claimants for "anomalous data patterns" and use those patterns to terminate benefits and level fraud accusations, as Crawford documents—automating suspicion against precisely the people least able to contest it [24]. Amsterdam spent enormous effort attempting to build a genuinely fair welfare AI, complete with bias audits and stakeholder consultation, and the experiment failed anyway; the system could not be made non-discriminatory in practice, a result

[24] The Atlas of AI

that should chasten anyone who believes fairness is a tuning parameter [21]. The broader record is bleak: AI still fails at social welfare across benefit systems, misclassifying the eligible and burdening the vulnerable with the labor of proving they are not fraud [12]. In child welfare, algorithmic decision-making now helps determine which families are investigated and which children removed—decisions with legal and constitutional weight, made by systems whose logic families cannot inspect or challenge [8].

Set the two conversations side by side and the moral geography of the discourse snaps into focus. On one side: a poor family flagged by a welfare algorithm, a worker cut off from unemployment on the strength of a false anomaly, a parent investigated because a model assigned their household a risk score. On the other: a philosophy of machine consciousness, conducted in well-funded institutions, addressing an entity whose capacity to suffer is entirely speculative. The AI welfare framing does not merely misallocate attention. It performs a kind of ideological work, shifting the moral subject of AI ethics from the human being harmed now to the machine that might, someday, deserve consideration. The question Müller poses—“who is responsible when something goes wrong?”—gets quietly displaced by a more comfortable one about whether the machine has feelings [24]. Comfortable, because a machine’s hypothetical feelings make no claim on any company’s balance sheet, while a wrongfully terminated benefit does.

This is the legitimacy crisis in miniature. A governance discourse that can find the resources to worry about the rights of models it has not built, while shrugging at the injuries of people its systems have already harmed, is a discourse that has lost the plot of its own moral purpose. Whose suffering counts is always a political question. The AI welfare turn answers it in favor of the abstraction and against the person, which is exactly the answer that costs the powerful nothing.

The Structural Silences

The most revealing part of any discourse is what it does not say, and the people it does not let speak. The AI industry’s public story is one of frictionless intelligence—clean, scalable, almost immaterial. That story requires two enormous silences, and both concern the people and places on whom the whole edifice rests. The first silence is the workforce. Behind the seeming autonomy of every large model stands an army of low-paid data laborers, often in the global South, who label the training data, filter the toxic content, and absorb the psycholog-

[21] Inside Amsterdam’s high-stakes experiment to create fair welfare AI

[12] Bias, bots, and benefit blunders: Why AI still fails at social welfare.

[8] Algorithmic Decision-Making in Child Welfare Cases and Its Legal and

[24] AI Ethics

ical cost of moderating what the systems must learn to avoid [14]. These workers make the intelligence possible and appear nowhere in its telling. Their absence from the discourse is not an oversight. It is a precondition of the discourse’s central myth—that the intelligence is artificial rather than extracted from human labor at scale.

The second silence is material. The industry that presents itself as weightless runs on data centers whose carbon footprint is substantial and largely hidden from the public accounting of AI’s costs [20]. Crawford’s whole atlas is an argument that AI is a planetary extractive industry—minerals, energy, water, labor—dressed in the language of the ethereal [24]. And the economic arrangement is more extractive still. One provocative reading holds that AI is not an industrial revolution but a fiscal one—a mechanism for concentrating wealth and restructuring who pays taxes and who captures value, with the public bearing the costs and a handful of firms capturing the returns [22]. Even the accounting of where AI’s economic value flows tends to be authored by the firms themselves; Google’s own atlas of Gemini usage maps the economy through the lens of its product’s adoption, a framing in which the company is both the mapmaker and the territory [17]. When the party benefiting from a technology also authors the authoritative account of its benefits, the silences in that account are not random.

There is a further silence worth naming: the gap between documenting harm and building remedy. The evidence of algorithmic injury is now abundant—the studies, the lawsuits, the investigative reports cited throughout this essay attest to a documentary apparatus in good working order. What is missing is the corresponding machinery of repair. We are extraordinarily good at proving that hiring algorithms discriminate and remarkably poor at compelling any firm to stop, or at compensating the millions already filtered out. This asymmetry is itself a distribution of power. Documentation is available to journalists, academics, and advocacy groups; remedy requires the coercive authority of the state, which is precisely the authority that sectoral capture has hollowed out. The harm is legible and the accountability is not, and that mismatch is not a temporary lag. It is the equilibrium the current arrangement is designed to hold.

Consider, too, who is allowed to blame whom. When an AI system fails, the blame diffuses into the technology itself—“the algorithm was biased,” as if the algorithm were a weather event. The vendor who built it, the executive who deployed it to cut costs, the legislator who let its definitions into the statute: these actors recede behind the passive grammar of technical failure. Employee monitoring systems now profile workers continuously, rating performance and flagging devia-

[14] Detrás del auge de la Inteligencia Artificial hay un ejército de trabajadores en fábricas de explotación digital

[20] IA et environnement : l’empreinte carbone cachée des modèles d’IA

[24] The Atlas of AI

[22] Kevin Polizzi - L’intelligence artificielle n’est pas une révolution industrielle. C’est une révolution fiscale

[17] Google’s AI & Economy ATLAS v1.0: Mapping Gemini Usage

tion, and the legal and ethical risks land on the monitored rather than the monitoring firm [18]. The person surveilled bears the consequence; the party who chose to surveil bears none. This is the accountability gap rendered as everyday practice—a machinery of blame that flows reliably downhill, toward the least powerful, and never up toward the parties who designed and profited from the system.

[18] How A.I. Is Changing Employee Monitoring and Performance Reviews

What the Fracture Reveals

The occasional regulatory response does register, and it is worth crediting where it happens. Amazon’s move to require sellers to label AI-generated images of people, prompted by a New York state law, shows that binding rules can shift corporate behavior when they carry legal weight rather than voluntary aspiration [9]. The lesson is not that regulation is futile. It is that regulation works exactly to the extent that the regulated party did not write it. The New York labeling law functioned because it imposed a real obligation from outside the industry’s preferred framing. The sectoral capture on display in less contested venues fails for the opposite reason: the obligation was authored by the obligated.

[9] Amazon cracks down on use of AI images by sellers after New York law

That distinction is the spine of this week’s fracture. Where governance is captured, it produces the appearance of oversight without the substance—principles no one enforces, definitions that erase the harms they should name, bans the banned can route around. Where resistance is genuine, it is still uncoordinated, improvising against adversaries who hold the market on both ends of the surveillance relation. And in the gap between the two, the discourse comforts itself with speculation about the moral status of machines while the documented injuries to human beings accumulate, uncompensated and largely unpunished. The Amsterdam welfare experiment’s failure should be read as the definitive rebuttal to the technocratic hope that fairness can be engineered into these systems after the fact [21]. If a city that tried this hard, with full political will and expert support, could not build a non-discriminatory welfare algorithm, the problem is not technical refinement. The problem is that the power to deploy these systems and the power to be harmed by them belong to different people, and no amount of engineering closes that distance.

[21] Inside Amsterdam’s high-stakes experiment to create fair welfare AI

Power in AI discourse, then, flows to those who build the systems, write the rules, author the accounts, and escape the blame—and away from those who are screened out, surveilled, cut off, and left to prove their own injury through instruments almost none of them can afford. The Workday plaintiffs will spend years establishing what a single

audit could have shown, while the vendors iterate and the sectoral statutes accrete their industry-friendly definitions [3]. A governance system that ignores current injuries in favor of managing appearances is not merely inadequate. It is building its own legitimacy crisis, one uncontested rejection and one unaccountable arrest at a time. The resistance on the streets is the early sign that the governed have noticed the arrangement. Whether the formal apparatus can be pried loose from the hands that captured it—whether governance can be made to serve the harmed rather than the harming—is the only question about AI that the coming years will actually answer. This week suggested the answer will not arrive from inside the rooms where the powerful write their own constraints.

[3] AI Hiring Discrimination: How Algorithms Reject Millions of Qualified

References

1. AI Discrimination Lawsuits Explode: When Your Hiring Algorithm Becomes
2. AI Hiring Bias: The Workplace Problem AI Didn't Create
3. AI Hiring Discrimination: How Algorithms Reject Millions of Qualified
4. AI is more likely than humans to form biases when hiring
5. AI surveillance and data colonialism shape African conflicts
6. AI Tends to Develop New Stereotypes to Base Hiring Decisions On, Study Says
7. AI-generated images fool Spokane city officials twice in six days
8. Algorithmic Decision-Making in Child Welfare Cases and Its Legal and
9. Amazon cracks down on use of AI images by sellers after New York law
10. Automating Repression Beyond Borders
11. Axon Says AI Police Reports Save Time. Public Records Show They Get Facts Wrong
12. Bias, bots, and benefit blunders: Why AI still fails at social welfare.
13. Bots policiales: la IA que infiltra protestas
14. Detrás del auge de la Inteligencia Artificial hay un ejército de trabajadores en fábricas de explotación digital

15. Estas ciudades prohíben la tecnología de reconocimiento facial. La
16. Flawed Facial Recognition Technology Leads to Wrongful Arrest and
17. Google's AI & Economy ATLAS v1.0: Mapping Gemini Usage
18. How A.I. Is Changing Employee Monitoring and Performance Reviews
19. How Kenyans are using AI during protests
20. IA et environnement : l'empreinte carbone cachée des modèles d'IA
21. Inside Amsterdam's high-stakes experiment to create fair welfare AI
22. Kevin Polizzi - L'intelligence artificielle n'est pas une révolution industrielle. C'est une révolution fiscale
23. Monocultivo algorítmico en contratación y sesgo sistémico
24. The Atlas of AI
25. The Dangers of Unregulated AI in Policing
26. UK: Police forces 'supercharging racism' with crime predicting tech
27. USA/Global: Tech made by Palantir and Babel Street pose surveillance
28. What the Workday Lawsuit Reveals About Future of AI Hiring